

**Experimental and theoretical studies of
non-equilibrium systems: Motor-microtubule
assemblies and the human-Earth system**

Thesis by:
Rachel A. Banks

In Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Biochemistry and Molecular Biophysics

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2022
Defended on December 13, 2021

© 2022

Rachel A. Banks

ORCID: 0000-0003-2028-2925

All rights reserved

Acknowledgements

This work would not have been possible without the endless support of my mentors, labmates, friends, and family. However, I probably would not have even ended up at Caltech studying biophysics if not for Phil Nelson, who first introduced me to the field in the third year of my undergraduate studies at the University of Pennsylvania. My perspective on biology was changed by taking his biophysics course and I immediately added a biophysics major.

I met my advisor, Rob Phillips, during my first week at Caltech, during the physical biology bootcamp that he teaches, along with several of his graduate students as TAs. His approach to quantitatively understanding biology reminded me of the course at Penn that had influenced me so much that I knew I wanted to rotate in the lab. From then on, Rob has been an incredible advisor and guide throughout my time in graduate school. I am constantly inspired by his love of science and tireless dedication to his students. He has always been supportive of me and happy to help in any way he can. I have truly enjoyed all the hikes, swims, and bike rides we have gone on together and the time to casually discuss science. He gave me incredible freedom to follow my instincts, even when I told him I wanted to work on something completely new in my fourth year, and develop as a scientist while also always being there for guidance and support, and I am incredibly grateful for that.

From the beginning of my rotation in the lab, I have been working with Heun Jin Lee, who I would like to thank for being a great guide and source of knowledge

on a wide array of topics. We have spent countless hours working together to come up with ideas and make experiments work. Both at Caltech and during summers at the Marine Biological Laboratory he has been a great partner and together we were able to make experiments work that sometimes seemed like they never would. I am also grateful for all the coffee breaks we have taken that helped us push through those hard times.

Also, the person who none of this work on active matter would have been possible without, Tyler Ross. He started the active matter work, and Heun Jin and I joined to help him and then continue studying the system. I have been inspired by his resilience, determination, and creativity in designing and getting experiments to work. I learned so much from him and have enjoyed all the time we have worked together. Tyler has also been a great friend, from trying restaurants to watching soccer games and comedy shows, I am grateful for all the time we have spent together.

The rest of the active matter team has made being in the lab a joy. Soichi Hirokawa, who taught me how to ride a bike when we spent a month at Stanford, has been great to work with and go on bike rides with. Ana Duarte, whose energy and spirit I have always enjoyed being around. She even became my hands when I broke my arm, happily helping me with everything from pipetting to tying my shoes. For about a year and a half during the pandemic, Soichi, Ana, Heun Jin, and I played time series almost every week, which helped me stay sane and was something I constantly looked forward to. Vahe Galstyan was a pleasure to work with over the years, discussing experiments and theory. I have learned a lot working with him and his rigorous approach to science.

For the past year, I have had the joy of working with the human impacts team : Griffin Chure, Avi Flamholz, Ignacio Lopez-Gomez, Nicholas Sarai, and Mason Kamb. Working with this eclectic group of people joined by a desire to understand and quantify how humans are shaping the Earth has been incredibly fun and satisfying. Special thanks to Griffin Chure, our fearless leader who I have had a great time doing estimates and writing with and have learned so much from.

The rest of the Phillips lab has been a constant source of inspiration, support, and fun times. I would like to thank current and past members of the group including: Rebecca Rousseau, Tom Röschinger, Gabe Salmon, Grace Solini, Scott Saunders, Manuel Razo Mejia, Suzy Beeler, Muir Morrison, Niko McCarty, Bill Ireland, Zofii Kaczmarek, Tal Einav, Stephanie Barnes, and Nathan Belliveau. I am very grateful that I have had the opportunity to work with this group.

I also had the opportunity to spend a month with Zev Bryant's lab at Stanford my first summer at Caltech and learn from them how to work with insect cells and purify and characterize motor proteins. I would like to thank Zev, Athena Ierokomos, and the rest of his lab for being so welcoming and willing to teach me. The skills I learned from them were vital for the rest of my work on active matter.

That same summer, I also went to the Marine Biological Laboratory in Woods Hole for the first time as a TA for the physical biology of the cell course. I am grateful to the course directors (Rob Phillips and Hernan Garcia), other faculty, TAs, and students that made that course an amazing experience. I also was able to TA for Phil Nelson during that course; it was great to work with and learn from him again. I went back to MBL the following year (2019) as a TA for the physiology course. Heun Jin Lee and I worked together as experimental TAs for Rob, trying to get new experiments to work in a new environment. I'll always remember the first time we got the "light as a ligand" and iATP experiments to work, and how we couldn't believe our eyes. That experience was incredibly challenging and rewarding and pushed me as a scientist, it would not have been possible without Heun Jin or the students who we worked with. Again, I would like to thank the course directors (Nicole King and Daniel Fletcher), faculty, TAs, and students who defined that experience.

The Biochemistry and Molecular Biophysics option has been a great community throughout my time at Caltech. I would like to thank Bil Clemons, the head of the option, and Courtney Oida, our option admin for organizing the retreats, seminars, and parties that brought the option together. All of the students in the option have been a great group to navigate grad school with and I am deeply grateful to everyone. In particular, I would like to thank Alex Barbato, my roommate for our

first two years, for all her support and many nights making popcorn and watching tv shows.

My thesis committee, David Van Valen, Justin Bois, and Matt Thomson, have also provided considerable guidance. I thank them for giving their time to think critically about my work and provide incredibly useful insight. Justin's classes were among the best I have ever taken, he taught me so much material but also about what it means to do high-quality science.

Outside of the lab, I would like to thank all of my friends and family, who I would not have made it through grad school without. My parents provided every opportunity and constantly supported me to learn and explore my interest in the sciences from summer programs and trips to the science center as a kid to reading my papers and being practice audiences for talks in undergrad and grad school. Soccer has been an outlet and how I have spent a lot of my free time, and I am grateful to everyone who I have had the pleasure to play with during these years. When I couldn't play soccer during the pandemic, bike riding and triathlons became my outlet, and the Pasadena Triathlon Club was a great source of inspiration and pushed me to do things I never thought I could. I would not have made it to this point without the support of my friends and family who have been with me through all the ups and downs of grad school.

They say it takes a village, and I have been incredibly lucky to have you all with me along the way, getting me to this point.

Abstract

Systems out of equilibrium are pervasive around us. In fact, being out of equilibrium is a key property of life, as described by Erwin Schrodinger in his series of essays "What is life?". Through the consumption of energy, i.e. food, living organisms achieve ordered states that would be very unlikely to occur at equilibrium, such as the mitotic spindle during cell division, swarms of bacteria, or flocks of starlings. The Earth system is another example of a non-equilibrium system. The state of the Earth has been evolving for billions of years, often under the influence of life. Today, humanity is a dominant influence forcing the Earth system to new states. Understanding these non-equilibrium systems has posed many challenges; in this thesis, we work towards quantitatively dissecting and gaining an intuition for the functioning of both a molecular scale and planetary scale non-equilibrium system.

Underlying many cellular functions such as cell division and transportation of organelles is the cytoskeleton composed of motor proteins and their constituent filaments. One of the key components are kinesin motors, which consume chemical energy to walk along and reorganize microtubules. Collections of these motors and microtubules are able to form organized structures. Understanding how these structures are formed has remained an open question. In Chapter 2, we develop a system of kinesin motors and microtubules wherein motor activity is controlled by light, thereby gaining spatiotemporal control over the formation of motor-microtubule assemblies. We demonstrate the creation of a variety of structures

of different sizes and geometry, and measure how length and time scales of these assemblies depend on the activated region.

A remaining question was how the microscopic details of the interaction between motors and microtubule affect the dynamics and steady-state structure formed. With our scheme for light-control in hand, we extended the system to a variety of motor proteins that have different speeds, processivities (how many steps they take before unbinding from the microtubule), directionalities (which end of the microtubule they walk towards), and forces they are able to exert in Chapter 3. We found that the size of steady-state structures, distribution of motors within assemblies, and rate of contraction of networks depend on motor properties. Further, we demonstrate that various structures can be formed by combining different motors. This work begins to build a connection between the detailed microscopic interactions of cytoskeletal components to the larger scale structures they form.

Chapter 4 begins our work on understanding the state of the human-Earth system. A major hurdle to quantitatively understanding this system is the difficulty of finding and parsing the relevant data, which is often within long, complicated reports. In order to facilitate access to this data, we created the Human Impacts Database, which houses a collection of > 300 carefully curated values related to human impacts on the Earth, introduced in Chapter 4. In this chapter, we describe the format of the database as well as demonstrate how it can be harnessed to gain a more holistic perspective on humanity's influence on the Earth.

Having this data is only a starting point towards deciphering the ways that humans are altering the state of the Earth, though. In Chapter 5, we combine these quantitative measurements with simple order-of-magnitude estimates to gain an intuition for the magnitude of several of the values. In this way, we show that many of the ways humanity is affecting the Earth can be tied back to how much land, water, and power we use. We further contextualize the magnitude of human influence by comparing human activities to natural analogs, finding that humans currently rival natural processes in influencing the state of the Earth system.

Published Content and Contributions

- [1] Rachel A. Banks et al. *Motor processivity and speed determine structure and dynamics of microtubule-motor assemblies*. en. preprint. BioRxiv, Oct. 2021. DOI: 10.1101/2021.10.22.465381.
R.A.B. designed and implemented experiments, performed data analysis and wrote the manuscript.
- [2] Griffin Chure et al. “The Anthropocene by the Numbers: A Quantitative Snapshot of Humanity’s Influence on the Planet”. In: *arXiv:2101.09620 [physics]* (Jan. 2021). arXiv: 2101.09620.
R.A.B. compiled and analyzed data, assisted with visualizations, and wrote the manuscript.
- [3] Tyler D. Ross et al. “Controlling organization and forces in active matter through optically defined boundaries”. en. In: *Nature* 572.7768 (Aug. 2019), pp. 224–229. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-019-1447-1.
R.A.B. performed and analyzed gliding assays, edited figures, and aided in writing the manuscript.

TABLE OF CONTENTS

Acknowledgements	iii
Abstract	vii
Published Content and Contributions	ix
Table of Contents	ix
List of Figures	xv
List of Tables	xx
Chapter 1: Introduction	1
1.1 Introduction	1
1.2 Various Ways Biology Generates Order in and out of Equilibrium	2
1.3 Towards a Quantitative Understanding of Active Matter	4
1.3.1 The Earth System is out of Equilibrium	7
Chapter 2: Controlling Organization and Forces in Active Matter Through Optically-Defined Boundaries	12
2.1 Abstract	12
2.2 Introduction and Results	13
2.3 Methods and Materials	22
2.3.1 Kinesin Chimera Construction and Purification	22
2.3.2 Microtubule Polymerization and Length Distribution	23
2.3.3 Sample Chambers for Aster and Flow Experiments	24
2.3.4 Reaction Mixture and Sample Preparation for Aster and Flow Experiments	25
2.3.5 Sample Preparation for Gliding Assay	26
2.3.6 Preparation of Tracer Particles	26
2.3.7 Microscope Instrumentation	26
2.4 Data Acquisition, Analysis, and Supplemental Discussion	27
2.4.1 Aster Distribution in 3D	27
2.4.2 Comparisons with Similar Systems	27
2.4.2.1 Microtubule Vortices	27
2.4.2.2 Extensile vs. Contractile	30
2.4.3 Microscopy Protocol	30
2.4.4 Measuring Aster Spatial Distribution with Image Standard Deviation	31
2.4.5 Characteristic Size of an Aster	32
2.4.5.1 Determining Characteristic Size	32

2.4.6	Image Analysis of Asters	32
2.4.6.1	Image Preparation	32
2.4.6.2	Defining the Regions of Interest	33
2.4.7	Reversibility of Aster Formation and Decay	33
2.4.8	Speed and Characteristic Time Scales of Formation and Merging	34
2.4.9	Comparison to Light Activated Actomyosin Networks	36
2.4.10	Analysis of Aster Decay	37
2.4.11	Diffusion Coefficient of a Microtubule	38
2.4.12	Scaling Arguments for Aster Size and Comparison to Data	40
2.4.13	Tracking of Moving Aster	41
2.4.14	Effective Potential of a Moving Aster	42
2.4.15	Mechanism and Stability of a Moving Aster	44
2.4.16	Single Motor Velocity Determination from Gliding Assay	45
2.4.17	Minimum Size Limits of Structures	46
2.4.18	Fluid Flow Patterns from Particle Tracking	49
2.4.19	2D Flow Field	52
2.4.20	Time Stability of Flow Patterns	52
2.4.21	Generation of Streamline Plots	52
2.4.22	Correlation Length	55
2.4.23	Theoretical Model of the Fluid Flow Field	55
2.4.24	Calculating Fluid Viscosity	57
2.4.25	Comparison to Optically Controlled Bacteria	60
Chapter 3: Motor Processivity and Speed Determine Structure and Dynamics of Microtubule-Motor Assemblies 61		
3.1	Abstract	61
3.2	Introduction	62
3.3	Results	63
3.3.1	Aster Size Depends on Motor Used	63
3.3.2	Spatial Distribution of Motors in Asters	66
3.3.3	Contraction Rate Scales with Motor Speed	71
3.3.4	Motor Properties Determine Outcomes of Motor Competitions	72
3.4	Discussion	75
3.5	Materials and Methods	77
3.5.1	Cloning of Motor Proteins	77
3.5.2	Protein Expression and Purification	77
3.5.3	Microtubule Polymerization	78
3.5.4	Microtubule Length	79
3.5.5	Sample Chamber Preparation	79
3.5.6	Reaction Mixture Preparation	80

3.5.7	Microscope Instrumentation	80
3.5.8	Activation and Imaging Protocol	80
3.5.9	Gliding Assay	81
3.5.10	Disordered Aster Core	82
3.5.11	Aster Size	83
3.5.12	Merger Analysis	85
3.5.13	Model of Network Contraction	87
3.5.14	Motor Competition	87
3.5.15	Motor distributions	90
3.5.15.1	Model formulation	90
3.5.15.2	Extraction of concentration profiles from raw images	95
3.5.15.3	Model fitting	97
3.5.15.4	Expected ratio of K_d values for K401 and Kif11 motors	99
3.5.15.5	Accounting for finite MT lengths	101
3.5.15.6	Broader spread of the tubulin profile	103
3.6	Supplementary Movies	107
Chapter 4: The Human Impacts Database: A Quantitative Repository of Humanity's Influence on the Planet		109
4.1	Abstract	109
4.2	Introduction	110
4.3	Results	112
4.3.1	Finding and Compiling Numbers from Scientific Literature, Governmental and Non-Governmental Reports, and Indus- trial Datasets	112
4.3.2	Global Magnitudes	115
4.3.3	Regional Distribution	117
4.3.4	Timeseries of Human Impacts	120
4.4	Discussion	123
4.5	Supplementary Material	127
4.5.1	Curation of Human Impacts Database Entries	127
4.5.2	References and Explanations for Values Reported in Figure 4.2	127
4.5.2.1	A. Surface Warming	128
4.5.2.2	B. Annual Ice Melt	129
4.5.2.3	C. Sea Ice Extent	130
4.5.2.4	D. Annual Material Production	130
4.5.2.5	E. Livestock Population	131
4.5.2.6	F. Annual Synthetic Nitrogen Fixation	132
4.5.2.7	G. Ocean Acidity	132
4.5.2.8	H. Land Use	132

4.5.2.9	I. River Fragmentation	133
4.5.2.10	J. Human Population	133
4.5.2.11	K. Greenhouse Gas Emissions	134
4.5.2.12	L. Water Withdrawal	135
4.5.2.13	M. Sea Level Rise	135
4.5.2.14	N. Total Power Use	136
4.5.2.15	O. Tree Coverage Area Loss	136
4.5.2.16	P. Power From Fossil Fuels	137
4.5.2.17	Q. Power From Renewable Resources	137
4.5.2.18	R. Fossil Fuel Extraction	138
4.5.2.19	S. Ocean Warming	138
4.5.2.20	T. Power From Nuclear Fission	139
4.5.2.21	U. Nuclear Fallout	139
4.5.2.22	V. Contemporary Extinction	139
4.5.2.23	W. Earth Moving	140
Chapter 5: Numeracy for a Human-Dominated Planet: An Order of Magnitude View of the Anthropocene 141		
5.1	Abstract	141
5.2	Introduction	142
5.3	Human Land Use	143
5.3.1	Urban Land Use	144
5.3.2	Rural Land Use	145
5.3.3	Agricultural Land Use	147
5.3.4	The Terra Number	150
5.4	Sizing up Human Livestock Cultivation	151
5.4.1	Chicken	153
5.4.2	Cattle	155
5.4.3	Fish	158
5.4.4	The Barnyard and Fishery Numbers	162
5.5	Human Water Use	164
5.5.1	Domestic Water Use	166
5.5.2	Agricultural Water Use	168
5.5.3	Thermoelectric Water Use	169
5.5.4	Hydroelectric Water Use	175
5.5.5	The Water Number	178
5.6	Human Impacts on the Nitrogen Cycle	181
5.6.1	Natural Sources of Reactive Nitrogen	185
5.6.2	Anthropogenic Nitrogen Fixation	187
5.6.3	The Nitrogen Number	190
5.7	Human Impacts on Methane Emissions	191

5.7.1	Assessing Atmospheric Methane Concentrations	193
5.7.2	Methane Emissions Due to Agriculture	194
5.7.3	Total Anthropogenic Methane Emissions	198
5.7.4	Natural CH ₄ Sources	198
5.7.5	Natural Methane Sinks	200
5.7.6	The Methane Number	201
5.8	Human Power Use	202
5.8.1	Transportation Power Use	204
5.8.2	Industrial Power Use	204
5.8.3	Residential Power Use	206
5.8.4	Energy Losses	207
5.8.5	Sizing up Incident Solar Power	208
5.8.5.1	Caveats in principle and in practice	208
5.8.6	The Solar Number	210
5.9	The Mass of Human-Made Materials	211
5.9.1	Buildings and Roads	212
5.9.2	Plastics	213
5.9.3	The Anthropomass Number	214
5.10	Human Impact on CO ₂ Emissions	215
5.10.1	Estimating Anthropogenic CO ₂ Emissions	219
5.10.2	Estimating CO ₂ Emissions from Fossil Fuels	220
5.10.3	Estimating CO ₂ Emissions from Land-Use Change	222
5.10.4	How Much CO ₂ Remains in the Atmosphere?	224
5.10.4.1	Plant CO ₂ Uptake	224
5.10.5	Oceanic CO ₂ uptake	225
5.10.6	How do Human CO ₂ Emissions Compare to Natural Processes?	227
5.11	Conclusions	228
5.12	Supplementary Information	229
5.12.1	Choice of dimensionless ratios	229
	Chapter 6: Conclusions and Future Work	234
	Bibliography	237

LIST OF FIGURES

<i>Number</i>	<i>Page</i>
1.1 Creation of order in biological systems	3
1.2 Active structures formed from <i>in vitro</i> motor-microtubule assemblies.	7
1.3 The planetary boundary framework. Several aspects of the human-Earth system are shown along with an assessment of whether human influence is currently "safe" (green) or "unsafe" (red). From [37].	9
1.4 Human water consumption from various perspectives.	11
2.1 Light-switchable active matter system enables optical control over aster formation, decay and size.	14
2.2 Moving and merging operations of asters with dynamic light patterns.	16
2.3 Operations for creating and moving asters are composed to make different desired patterns and trajectories.	18
2.4 Advective fluid flow is created and controlled with patterned light.	19
2.5 Kinesin motor coding regions.	22
2.6 Length distribution of microtubules.	24
2.7 3D projections of asters from Z-stacks imaged with a 20x objective.	28
2.8 Analysis of microtubule distribution in 3D.	29
2.9 Time series of light induced aster formation, decay, then formation.	34
2.10 A comparison of model fittings for a contracting aster experiment.	35
2.11 Characteristic times for contraction and merger as functions of activation length scales.	36
2.12 Plot of mean variance of image intensity as a function of time for different initial aster sizes. The shaded region is treated as part of the linear regime. The measure of time is relative to the beginning of aster decay.	39
2.13 Measuring the conservation of labeled fluorescent microtubules in the excitation region during aster formation.	40

2.14	The speed at which an aster returns to the center of the light pattern once the pattern stops moving. Red line is a plot of $y = x$	42
2.15	Aster following a 50 μm disk moving at 200 nm/s from right to left. Image is integrated across z	44
2.16	Intensity of an aster for a light pattern moving at 200 nm/s.	45
2.17	Velocity distribution of gliding microtubules.	47
2.18	Minimum length experiment for a $L \times 20 \mu\text{m}$ excitation pattern.	48
2.19	Minimum height experiment for a $350 \times H \mu\text{m}$ excitation pattern.	49
2.20	Minimum angle experiment for two $350 \times 20 \mu\text{m}$ excitation pattern.	50
2.21	Minimum aster size experiment for disk patterns.	51
2.22	Flow velocity field generated with a 350 μm activation bar measured with PTV of tracer particles. Vector data is used to calculate streamline plot in Fig. 4c.	51
2.23	A flow field measured at three different z -positions separated by 20 μm	53
2.24	Velocity field averaged over 7.5 minute intervals in a single experiment.	54
2.25	The average maximum speed for four different 7.5 minute time windows.	54
2.26	The correlation coefficient as a function of distance.	55
2.27	Flow field generated by 14 Stokeslets, indicated by green circles, to model the 350 μm activation bar case.	57
2.28	Demonstration of the linearity of the flow field.	58
2.29	Theoretical simulation of fluid flows under complex light patterns using Stokeslets.	59
3.1	Aster size depends on motor used.	64
3.2	Modeling the motor distribution in an aster.	69
3.3	Contractile speeds in motor-microtubule networks scale with network size and motor speed.	73
3.4	Motor competition modulates structure formation.	75
3.5	Cumulative distribution of microtubule lengths.	79
3.6	Histogram of calculated instantaneous speed of microtubules glided by K401 motors.	82
3.7	Aster centers are disordered and the arms are aligned radially.	83

3.8	Example images of microtubule fluorescence of asters made with each motor used and each excitation diameter.	84
3.9	Merger speeds depend on motor stepping speed.	86
3.10	Aster identification and speed calculation in aster mergers.	88
3.11	Distributions of Kif11 and Ncd in competition experiments.	89
3.12	Procedure for extracting protein concentration profiles demonstrated on an example aster.	98
3.13	Demonstration of the model fitting procedure for average as well as separate wedge profiles.	99
3.14	Collection of all fits to motor profiles. The green and blue dots represent the radially averaged motor and tubulin concentrations. The solid black lines represent the model fits.	100
3.15	Schematic representation of initial motor binding, advection and stalling at the microtubule end.	101
3.16	Relationship between motor and microtubule distributions.	104
4.1	A representative entry in the Human Impacts Database.	116
4.2	Human impacts on the planet and their relevant magnitudes.	118
4.3	Regional distribution of anthropogenic effects.	121
4.4	Timeseries of human impacts and their relationship to the human population growth.	124
5.1	Back-of-the-envelope estimate for the Earth's terrestrial area.	144
5.2	Back-of-the-envelope estimate for the extent of urban land use.	146
5.3	Estimate for the rural population density and global rural land area.	148
5.4	Estimating agricultural area from food consumption.	150
5.5	The Terra Number	151
5.6	Little's Theorem and calculating standing populations.	153
5.7	Estimates for the annual production of eggs and poultry from chicken.	156
5.8	Estimates for the annual production of beef and dairy and their corresponding cattle populations.	159
5.9	Estimate for annual seafood consumption standing populations of aquacultural animals.	161

5.10	Livestock population and biomass distribution by type and animal category.	163
5.11	The Barnyard and Fishery Numbers.	164
5.12	Estimating domestic water use.	167
5.13	Estimating agricultural water use for livestock drinking and irrigation.	169
5.14	Thermoelectric power generation and estimates for required water volumes.	173
5.15	Water withdrawal and consumption volumes for different thermoelectric cooling systems.	174
5.16	Estimating hydroelectric water use.	177
5.17	Estimating water consumption by hydroelectricity by evaporation.	178
5.18	The distribution of human water use.	179
5.19	The Water Number.	180
5.20	A simplified view of the global nitrogen cycle.	182
5.21	Estimate of the mass of nitrogen in the atmosphere and land.	184
5.22	Back-of-the-envelope estimates for NO _x produced by lightning and N released by weathering.	188
5.23	Estimate of nitrogen fixation for fertilizer.	189
5.24	The global natural and anthropogenic sources of reactive nitrogen.	191
5.25	The nitrogen number.	192
5.26	Sizing up the mass of methane in the atmosphere.	194
5.27	Estimate of methane emissions from cattle.	196
5.28	Estimate of methane emissions from rice paddies.	197
5.29	Anthropogenic sources of methane emissions.	198
5.30	Non-anthropogenic sources of methane emissions.	200
5.31	The Methane Number.	202
5.32	Breakdown of human power consumption by category.	203
5.33	Order-of-magnitude estimate of the amount of power consumed by driving.	205
5.34	Earth's radiative energy balance and an estimate of incident solar power.	209

5.35	Decomposition of the incoming solar radiation into energy fluxes that fuel the dynamics of the atmosphere, ocean and life on Earth.	211
5.36	The Solar Number.	212
5.37	Order-of-magnitude estimate of the mass of buildings and roads.	214
5.38	Order-of-magnitude estimate of plastic consumption.	215
5.39	Comparison of anthropomass and biomass.	216
5.40	The Anthropomass Number.	216
5.41	The concentration and mass of atmospheric CO ₂	218
5.42	Major fluxes in the global carbon cycle.	220
5.43	Energy and CO ₂ produced from burning common fuel sources.	221
5.44	Estimating total human energy production and anthropogenic CO ₂ emissions.	223
5.45	Estimation of CO ₂ emissions from land-use change.	224
5.46	Estimating oceanic carbon uptake using the Revelle factor.	226
5.47	The CO ₂ number.	228

LIST OF TABLES

<i>Number</i>		<i>Page</i>
3.1	Motor proteins used and their properties	65
3.2	Plasmids used	77
3.3	Processivities of motors, decay length scales of motor profiles, and corresponding ratios of these two length scales. Step size of ≈ 10 nm corresponding to the length of a tubulin dimer was used for estimating the motor processivities in μm units. For Ncd motors, the upper limit in processivity corresponds to that of oligomeric motor assemblies. Estimates for the decay length scales λ were made based on the motor profiles in Fig. 3.14.	95

Chapter 1

Introduction

1.1 Introduction

Many of the behaviors characteristic of life, such as replication, ordered patterns, and reliable processes have long fascinated and puzzled scientists. These characteristics were so different from behaviors typically observed in the non-living world that it was proposed that another "vital" force existed in biological systems that was outside the realm of the previously established physical laws [1]. However, over the past century, many of these mysterious processes have been shown to be governed by similar physical laws as the non-living world, negating the need for a new "vital" force. One key characteristic of these systems is the consumption of energy, enabling them to escape the rules of equilibrium and create patterned, organized structures.

Many examples of non-equilibrium systems exist outside of the living world as well. For example, weather patterns are non-equilibrium, dissipating energy in the form of kinetic energy such as winds. In fact, the Earth system as a whole is dynamic and out of equilibrium, being pushed to new states by biological and non-biological forces. Currently, humanity is a dominant force influencing the state of the Earth, affecting it in many distinct but interconnected ways.

In this thesis, I discuss our work to investigate and characterize both a biological non-equilibrium system and the human-Earth system. While these systems

differ in length scale by many orders of magnitude, in both cases, measuring various properties and then trying to understand what sets them is a step towards understanding. That is what we attempt to do in this thesis, but before diving into the work done for this thesis, we will provide context by briefly exploring the role of non-equilibrium systems in biology and the dynamic nature of the Earth system.

1.2 Various Ways Biology Generates Order in and out of Equilibrium

There are a variety of ways in which ordered states can be favored, both at equilibrium and out of equilibrium. Self assembly of structures at equilibrium has long been recognized, from protein folding to crystal structures of molecules and proteins to viral assembly. In these cases, organized structures are energetically favorable despite having low entropy due to the strength of inter- and intra-molecular interactions such as hydrogen bonds and hydrophobic/hydrophilic interactions. In these cases, the molecular composition, concentration of components, and solvent used are usually enough to determine the equilibrium state. A notable exception is the class of intrinsically disordered proteins, which tend to form flexible, dynamic structures with multiple possible conformations.

Alan Turing recognized another method of order creation prevalent in biological systems now known as Turing patterns. How patterns of gene expression vital to correct development during morphogenesis such as distinguishing head from tail or left from right are generated has long been an open question. It was hypothesized that patterns of concentration of molecules that modulate gene expression, termed morphogens, could be responsible for establishing differences in gene expression. Turing showed that these patterns could be established through simple chemical reactions and diffusion [2]. For example, Turing showed that a gradient of concentration could be formed just from diffusion, which resembles the observed gradient in Bicoid concentration in *Drosophila* embryos (Fig. 1.1, top). He went on to show that more complicated patterns such as oscillations could also arise from reaction-diffusion networks. These ideas have since been built upon to

1.2. Various Ways Biology Generates Order in and out of Equilibrium R. A. Banks

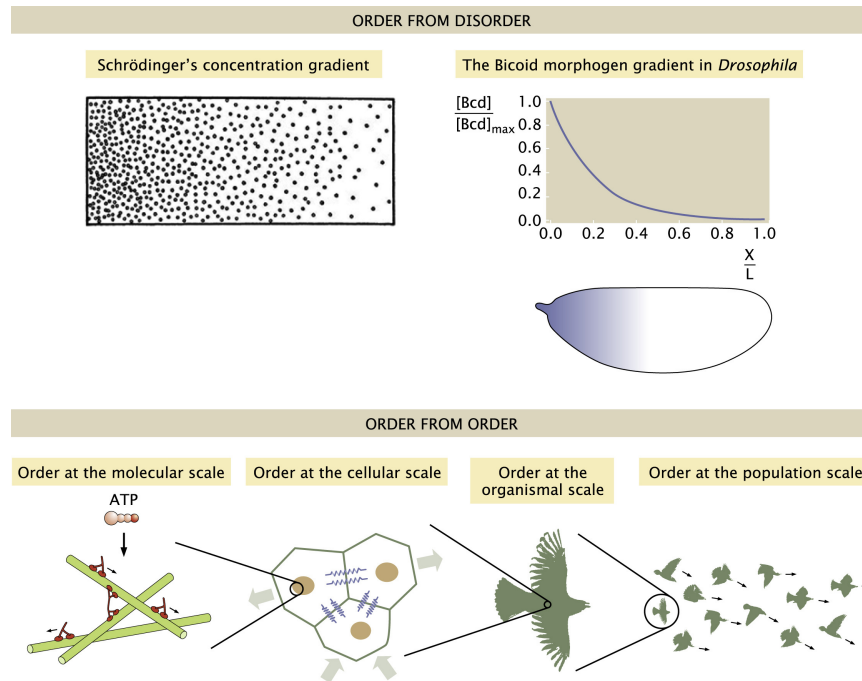


Figure 1.1: Creation of order in biological systems. Top: (Left) Schrodinger predicted that order such as concentration gradients could arise from diffusion. (Right) The Bicoid concentration gradient, plotted on the top and shown schematically on the bottom, is an example of the type of gradient from diffusion that Schrodinger predicted. Bottom: Consumption of chemical energy in the form of ATP generates ordered structures from the molecular to organismal scale. Adapted from [4].

engineer ever more intricate gene regulatory circuits [3].

The stable states that these reaction-diffusion networks adopt are distinct from the equilibrium states described previously in a number of ways. First, while stable, these states are still dynamic. The chemical reactions and diffusion of molecules continue to occur, generating fluxes of materials, but they are balanced so that the overall patterns are maintained. Second, there are often multiple possible stable states that could be produced; external symmetry breaking or random fluctuations are needed to push the systems away from one stable state and towards another. Thus, these systems depend on the initial conditions unlike equilibrium systems described above.

Active matter systems are another way through which life achieves order.

Active systems are composed of collections of components that consume energy in order to produce mechanical work on the local level. Collections of these components exhibit ordered behavior on length scales that can be hundreds to thousands of times larger than the individual actors. Examples of this type of organization are shown in the bottom panel of Figure 1.1. Consumption of chemical fuel, adenosine triphosphate (ATP) by molecular motor proteins allows them to walk along and organize filamentous proteins, microtubules (left). These networks underlie order at the cellular scale, which then enables order at the organismal and population scales. In this way, active matter creates the organization observed from molecular scale structures to flocks of birds.

While active systems can form steady structures that may appear to be at equilibrium at first glance, similar to the patterns described by Turing, there are still fluxes of materials. It is hypothesized that the consumption of energy enables these systems to undergo error correction and respond to outside forces. For example, the components of the mitotic spindle, molecular motors and microtubules, consume energy in the form of ATP and GTP, respectively. It has been shown that if this structure is perturbed chemically or by applying external forces, adapt and heal [5, 6]. The energy consumed by the motors and microtubules enables this healing process, making the critical process of cell division robust to external forces. We dive deeper into active matter systems in the next section, and how they have been studied experimentally and theoretically.

1.3 Towards a Quantitative Understanding of Active Matter

Active matter systems are composed of collections of self-propelled components that consume energy at the local level to perform mechanical work, often to create motion. The interactions of many of these particles, and with the medium they are in, can lead to organized, collective motion and structure formation. As mentioned previously, these systems are pervasive in biology across length scales. These phenomena have gained the attention of biologist and physicists alike, and much experimental and theoretical work has been done to build an understanding of them.

Some of the earliest studies of active matter centered around flocking behavior. This type of organized, directional motion without the influence of an external force is a unique phenomena of active systems. Vicsek first proposed a model for flocking, demonstrating through simulations that directional motion could be achieved through simple local interactions wherein particles assume the average direction of motion of those around them, with some noise [7]. Using a similar model, he was able to simulate bacterial colony growth that closely aligned with experimental data [8]. This model served as the base for further models of flocking behavior [9, 10, 11].

Similar self organized clustering behaviors have been shown more recently in non-living systems of particles that asymmetrically catalyze chemical reactions on their surfaces to produce motion [12, 13, 14]. These simple systems have proved fruitful towards developing a theoretical understanding of active systems, building general frameworks and testing which principles from thermodynamics can be applied in self-propelled non-equilibrium systems [13, 14].

The cellular cytoskeleton is another class of active matter that has inspired much theoretical and experimental work. The cytoskeleton is composed of collections of filamentous proteins (actin and microtubules) that undergo polymerization and depolymerization through the consumption of energy, and associated motor proteins (myosin and kinesin/dynein, respectively) that consume chemical energy to walk along and move these filaments. Collections of these molecules produce the structures and forces necessary for processes such as cellular migration and division. *In vitro* experiments with purified filaments and motors have demonstrated various behaviors depending on the concentration of components, medium they are in, and confinement of the system, from aligned traveling waves and flows to asters and vortices [11, 15, 16, 17, 18]. A few examples of the types of spontaneous structures formed are shown in Figure 1.2 (A) – (C).

These simplified systems again enabled progress towards theoretical modeling of these dynamic structures [19, 20, 21, 22]. For example, Surrey et al. developed a model for motor-microtubule interactions and performed simulations to attempt to recreate the types of structures observed in [15] (asters and vortices as in Fig. 1.2

(A) - (B)). With this theory experiment interplay, they were able to test the effects of varying motor and microtubule density on the steady-state structures formed, thus beginning to reveal the important parameters for dictating which structures are formed. A continuum model that described the evolution of motor and microtubule distributions during the formation of asters and vortices was also proposed [20]. This model was further developed by Sankaraman, Menon, and Kumar, and they began to explore what the effects of changing motor properties such as processivity would be on the steady-state organization [21]. However, further progress and testing of these models has been limited by the lack of spatiotemporal control over the formation of motor-microtubule assemblies and the limited ability to tune experimental parameters.

In the work described here, we begin to rectify these by developing an experimental system of motors and microtubules in which we are able to control the organization of motor-microtubule assemblies. Chapter 2 of this thesis describes and characterizes an active matter system of kinesin motors and microtubules wherein regions of motor activity are dictated by light. In this way, we are able to dictate when and where motor-microtubule assemblies are created. This deviates from previous experimental cytoskeletal systems that relied on spontaneous structure formation. We demonstrate the formation of asters [Fig. 1.2 (D)] and other structures, such as contractile bars, that generate flow patterns in the surrounding fluid [Fig. 1.2 (E)]. Gaining control over the creation of active structures enabled us to study how their shapes, sizes, and dynamics depend on the region of motor activity.

In Chapter 3, we utilize this system to dig into the question of how to connect the details of the motor-microtubule interactions to mesoscopic properties of the assemblies, such as sizes, organization, and contractile rates. Our approach was to perform similar experiments with a variety of motor proteins that have different speeds, processivities (how far they walk before falling off the microtubule), and directionality. We find that several aspects of the assemblies can be connected to these microscopic properties and also how the motors work in a collective.

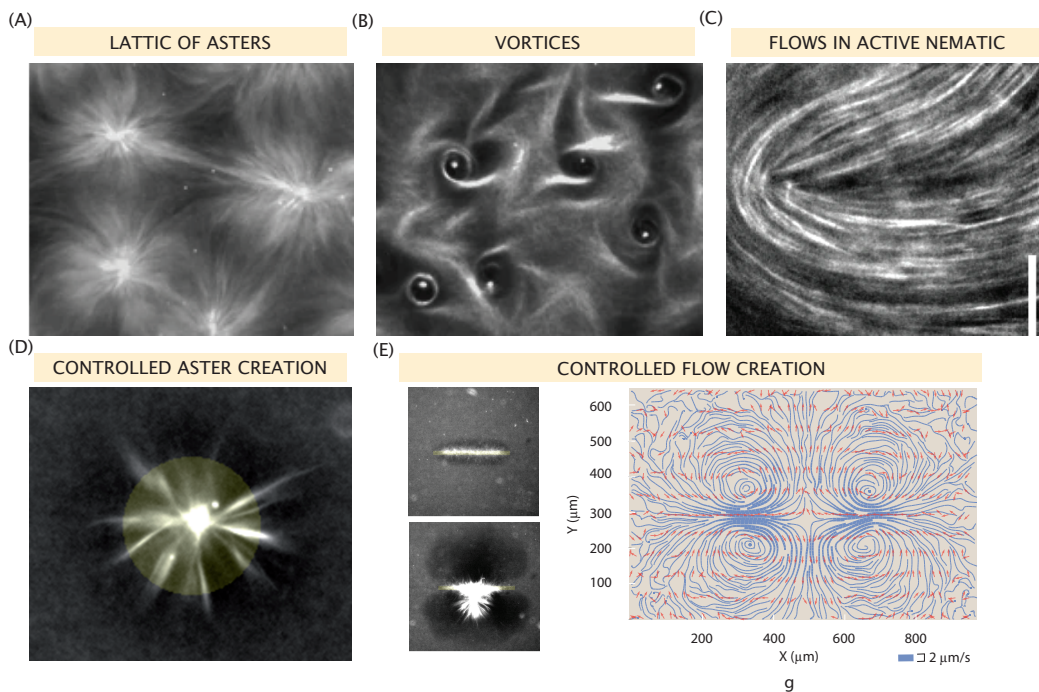


Figure 1.2: Active structures formed from *in vitro* motor-microtubule assemblies. (A) Lattice of asters formed spontaneously. (B) Vortices formed spontaneously. (A) and (B) are from [15]. (C) Disclination formed in an active nematic with spontaneous flows, from [17]. (D) Aster formed under light control, the region of light activation is shown with a yellow circle. (E) Controlled flow creation from light activated contractile network formation. Image of the microtubule fluorescence is shown on the left, the flow pattern is shown on the right. (D) and (E) are from [23].

1.3.1 The Earth System is out of Equilibrium

Many non-equilibrium systems exist in the non-living world as well. At the planetary scale, the Earth is constantly being shaped by forces both living and non-living, and is perpetually out of equilibrium. In fact, one of the key scientific insights of recent history is the realization that the evolution of the Earth is deeply intertwined with the evolution of life. A famous example of this is the rise of photosynthetic organisms around 3 billion years ago [24]. These organisms produce oxygen as a byproduct of their metabolism, and eventually produced so much oxygen that it began to fill the atmosphere, dramatically altering the chemistry, termed the Great Oxygenation Event. It has been hypothesized that this change

resulted in a mass extinction since the organisms occupying the Earth at that time were adapted to living off of other sources, such as methane and sulfur. To these organisms, oxygen would have acted as a poison, killing them and paving the way for the evolution of new organisms suited to an oxygenic atmosphere. Today, humanity is a dominant life force altering the Earth, from the chemistry of the atmosphere [25, 26, 27, 28] to the surface temperature and pH of the oceans [29, 30] to the composition of the land [31, 32], to name just a few examples.

The dominant influence of humanity on the Earth is underlined by the coining of the current geological epoch the "Anthropocene", or characterized by the action of humans [33]. Crutzen estimated the beginning of this epoch around the Industrial Revolution, the action of humans can be seen in the geological record through the rapid increase in atmospheric CO₂ concentration, which is estimated to have been stable for the previous two thousand years [34]. James Watts' invention of the steam engine enabled the Industrial Revolution and also drove the burning of fossil fuels to power the steam engines that are now used to power transportation, the generation of power, and many industrial processes. Burning fossil fuels results in CO₂ as a byproduct, which has been accumulating in the atmosphere as a fingerprint of human development.

While the previous geological epoch, the Holocene, was a relatively stable period in the Earth's history, humanity is now rapidly altering the Earth, pushing the system further from equilibrium. It has been hypothesized that several aspects of the Earth system are now nearing "tipping points," that could shift the system to a new state. Examples include the extent of Arctic sea ice, the strength of the El-Niño Southern Oscillation, and the dieback of the Amazon rainforest [35, 36].

Previously, studies have been done to attempt to set safe boundaries for human action to try to keep the Earth system in the stable Holocene-like conditions [37]. They defined several "planetary boundaries" that, if crossed, could have drastic consequences. The boundaries they defined encompassed axes such as climate change, ocean acidification, biodiversity loss, freshwater use, ozone depletion, and others. By comparing the boundary with current measurements, they determined whether human action currently lies within "safe" operating boundaries, as shown

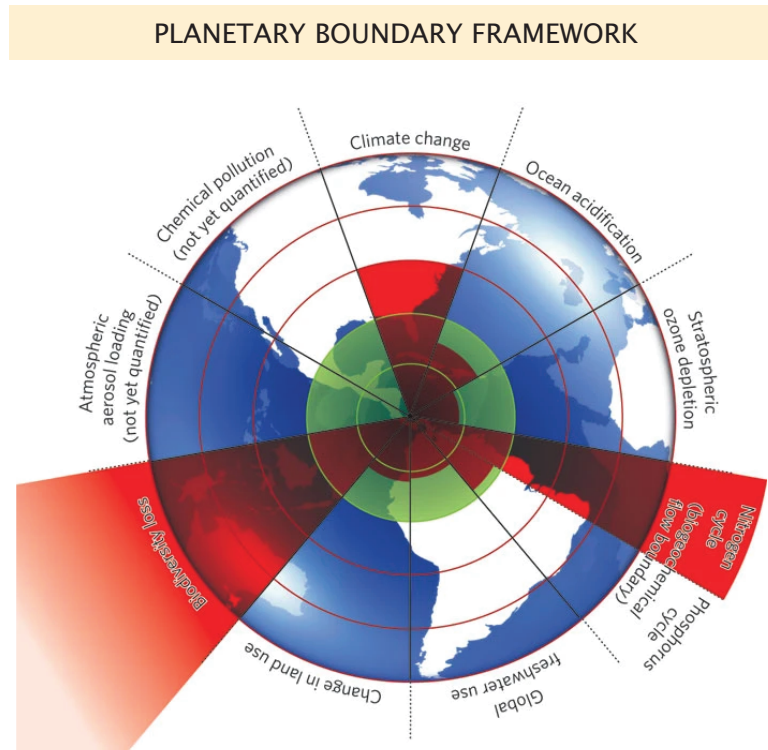


Figure 1.3: The planetary boundary framework. Several aspects of the human-Earth system are shown along with an assessment of whether human influence is currently "safe" (green) or "unsafe" (red). From [37].

in Figure 1.3. However, drawing distinct boundaries between safe and unsafe with respect to these quantities is not simple and where those boundaries would lie is widely contested among experts.

In our work, we take a more modest "just the facts" approach, attempting to compile and present the facts about the quantitative ways that humans are impacting the Earth system. Chapter 4 of this thesis describes the creation of the "Human Impacts Database," a searchable online resource housing > 300 numbers quantifying human influence. These numbers span the land, water, atmosphere, biogeochemistry, power use, and flora & fauna. While these values exist in other reports and studies, finding and understanding them can be a time consuming and frustrating process. We hope that the Human Impacts Database will facilitate this process and make quantitative facts about the human-Earth system accessible to

all. In addition, we compile several of the values in the database into a snapshot that presents the global magnitudes, regional breakdowns, and timeseries of several axes of human impacts. The regional breakdown and timeseries of human water use are shown as an example in Figure 1.4. From this presentation of the data, it becomes evident that while Asia uses a majority of the water used for agriculture, on a per capita basis, North America uses the most water for agriculture, industry, and domesticly [Fig. 1.4 (A)]. The timeseries data shows that agricultural water use has dominated the increase in water use as the population has grown, but that per capita agricultural and domestic use have had little change.

In Chapter 5, we give context for the magnitudes of human impacts through simple order-of-magnitude estimates. With these estimates, and comparisons to the data, we build an intuitive feel for the sometimes astoundingly large magnitude of human impacts. An example for agricultural water use is shown in Figure 1.4 (C) and (D). We estimate consumption by livestock, assuming that they drink about as much as humans per kilogram. We also estimate the amount of water used for irrigation from the extent of cropland and our intuition about how much water crops require [Fig. 1.4 (C)]. In Figure 1.4 (D), we compare this estimate with data on agricultural water consumption, finding that our simple estimate correctly accounts for the amount of agricultural water used by humans. We formulate these types of estimates for several axes of human impacts, revealing some of the dominant processes that make humanity's influence take the magnitude that it does. Finally, we also contextualize the magnitude of human impacts by comparing them to natural analogs, defining a series of dimensionless ratios which reveal that humans now rival natural forces in shaping the Earth.

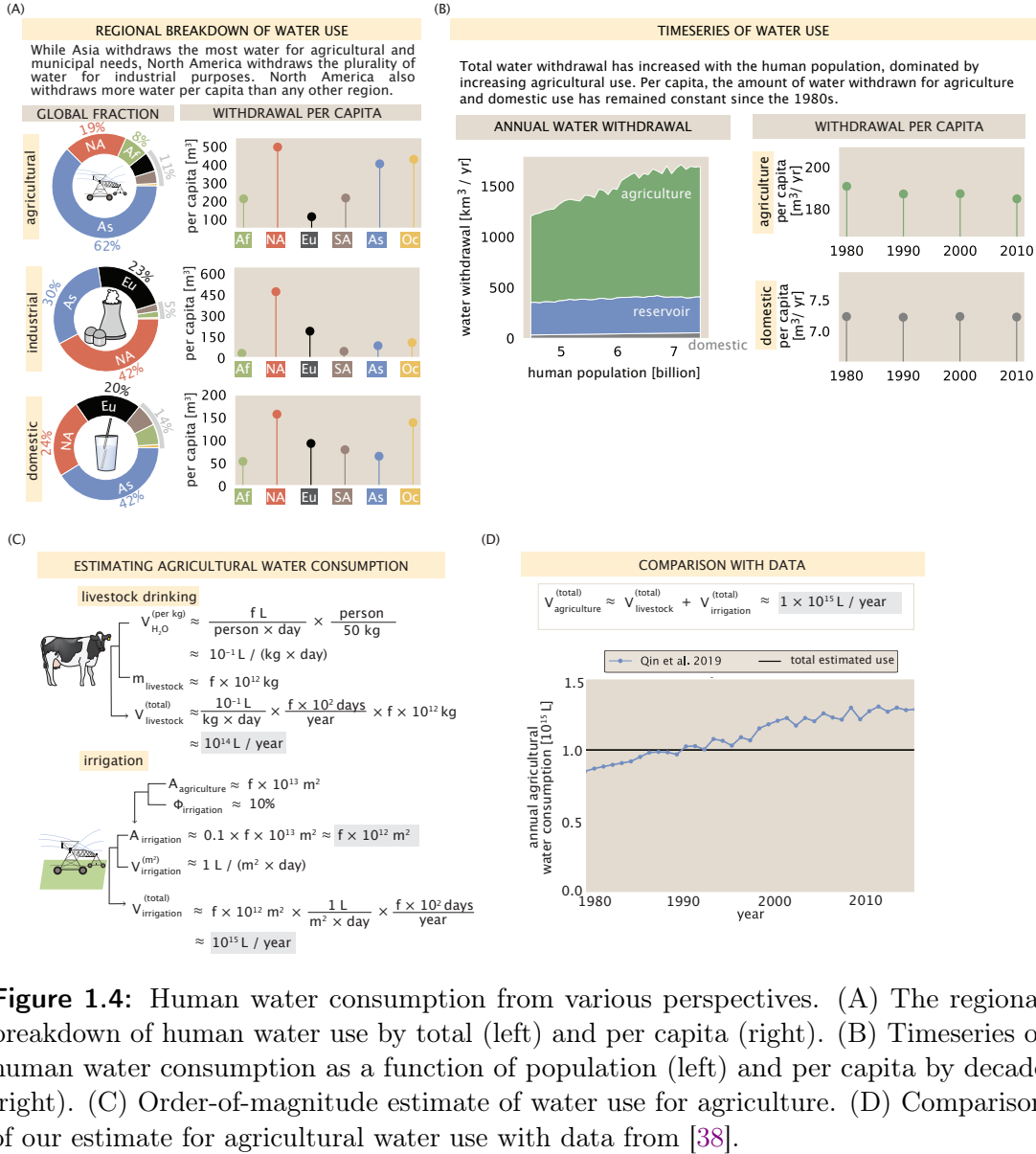


Figure 1.4: Human water consumption from various perspectives. (A) The regional breakdown of human water use by total (left) and per capita (right). (B) Timeseries of human water consumption as a function of population (left) and per capita by decade (right). (C) Order-of-magnitude estimate of water use for agriculture. (D) Comparison of our estimate for agricultural water use with data from [38].

Chapter 2

Controlling Organization and Forces in Active Matter Through Optically-Defined Boundaries

This chapter is derived from the journal publication : Ross, T.D., et al. Controlling organization and forces in active matter through optically defined boundaries. *Nature* **572**, 224-229 (2019). DOI: 10.1038/s41586-019-1447-1

2.1 Abstract

Living systems are capable of locomotion, reconfiguration, and replication. To perform these tasks, cells spatiotemporally coordinate the interactions of force-generating, “active” molecules that create and manipulate non-equilibrium structures and force fields that span up to millimeter length scales [39, 40, 18]. Experimental active matter systems of biological or synthetic molecules are capable of spontaneously organizing into structures [15, 19] and generating global flows [41, 17, 42, 43]. However, these experimental systems lack the spatiotemporal control found in cells, limiting their utility for studying non-equilibrium phenomena and bioinspired engineering. Here, we uncover non-equilibrium phenomena and principles by optically controlling structures and fluid flow in an engineered system of active biomolecules. Our engineered system consists of purified microtubules

and light-activatable motor proteins that crosslink and organize microtubules into distinct structures upon illumination. We develop basic operations, defined as sets of light patterns, to create, move, and merge microtubule structures. By composing these basic operations, we are able to create microtubule networks that span several hundred microns in length and contract at speeds up to an order of magnitude faster than the speed of an individual motor. We manipulate these contractile networks to generate and sculpt persistent fluid flows. The principles of boundary-mediated control we uncover may be used to study emergent cellular structures and forces and to develop programmable active matter devices.

2.2 Introduction and Results

Our scheme is based on a well-studied active system composed of stabilized microtubule filaments and kinesin motor proteins [15, 19, 44, 20, 41, 17, 45, 42]. In the original biochemical system, kinesin motors are linked together by practically irreversible biotin-streptavidin bonds. As linked motors pull on microtubules, a variety of phases and structures spontaneously emerge, such as asters, vortices, and networks. However, spatial and temporal control of these structures is limited [19, 46].

We engineered the system so that light activates reversible linking between motors (Fig. 3.1a) by fusing Kinesin I motors to optically-dimerizable iLID proteins [47]. Light patterns are projected into the sample throughout its depth and determine when and where motors link (see Supplementary Information for details). Outside of the light excitation volume, microtubules remain disordered, while inside the light volume, microtubules bundle and organize. The reversibility of the motor linkages allows structures to remodel as we change the light pattern. For a cylinder pattern of light excitation, microtubules contract into a 3D aster (Fig. 3.1b) (Supplementary Information 2.4.1, Video 1, Video 2). We use the projection of a cylinder of light as an operation for creating asters. We note that vortices, spirals, and extensile behavior are not observed under our conditions (Supplementary Information 2.4.2).

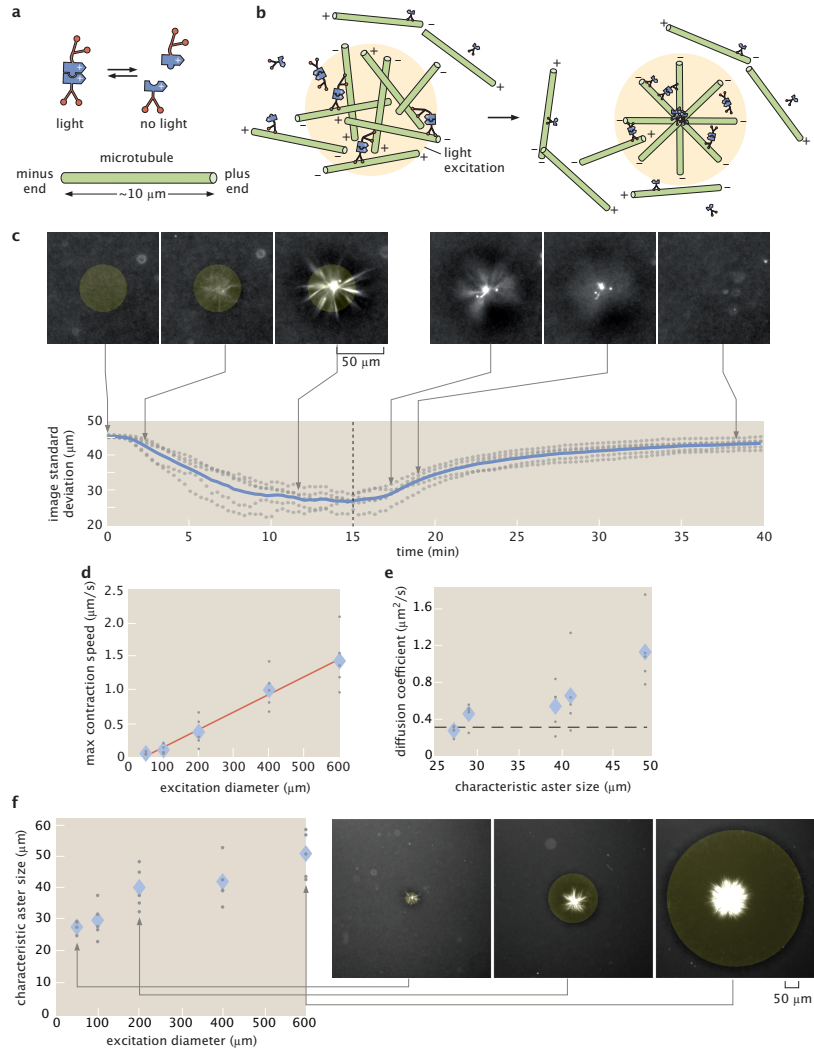


Figure 2.1: Light-switchable active matter system enables optical control over aster formation, decay, and size. **a**, Schematic of light-dimerizable motors. **b**, Schematic of light-controlled reorganization of microtubules into an aster. **c**, Images of labeled microtubules during aster assembly and decay and corresponding image spatial standard deviation versus time. The blue line is the mean of 5 experiments and the gray dots represent individual experiments. The dashed line is when the activation light is removed, transitioning from creation to decay. **d**, Max contraction speed versus excitation diameter. The red line is a linear fit. **e**, Diffusion coefficients versus characteristic aster size. The characteristic size is the image spatial standard deviation at the 15 minute time point shown in **(c)**. The dashed line represents the diffusion coefficient of a 7 μm microtubule (Supplementary Information 2.4.11). **f**, Aster characteristic size versus excitation diameter with representative images. In **(d, e, f)** the diamonds represent the mean of 5 experiments and the gray dots represent individual experiments. In **(c, f)**, the yellow shaded disks represent the light pattern.

Our temporal control over aster formation allows us to study the dynamics of their creation and decay (Fig. 3.1c) (Video 3) through time lapse imaging (Supplementary Information 2.4.3). We characterize these dynamics by measuring the spatial width of the distribution of fluorescently-labeled microtubules using image standard deviation (Supplementary Information 2.4.4). During aster formation, the distribution of microtubules within a cylinder pattern contracts. After 10-15 min, the distribution reaches a steady state, indicating that the aster is fully formed. To quantify a characteristic aster size (Supplementary Information 2.4.5), we measure the image standard deviation at 15 min (Supplementary Information 2.4.6). Once the excitation light is removed, asters begin to decay into free microtubules. The spatial distribution of microtubules widens over time, returning to the initial uniform distribution. Further, aster decay is reversible (Supplementary Information 2.4.7).

To understand scaling behavior, we investigate how the dynamics of aster formation and decay depend on excitation volume. During formation, microtubule distributions contract. The contraction speed (Supplementary Information 2.4.8) grows with the diameter of the excitation cylinder (Fig. 3.1d). Similar scaling of contraction speed has been observed for actomyosin systems [48] (Supplementary Information 2.4.9) and modeled for generic networks [49]. Alternatively, contraction can be measured by a characteristic contraction timescale [50] (see Supplementary Information 2.4.8). During decay, microtubule distributions spread in a manner consistent with diffusion (Supplementary Information 2.4.10). The effective diffusion coefficient is independent of characteristic aster size (Fig. 3.1e) and is consistent with what is expected for free microtubules (Supplementary Information 2.4.11). Further, we manipulate aster size through the diameter of the excitation volume (Fig. 3.1f) and find a scaling dependence (Supplementary Information 2.4.12) that shows similarities to the dependence of spindle size on confining volumes [51].

Moving activation patterns are responsible for dynamically re-positioning structures and forces within a cell [52]. We are able to similarly move asters by re-positioning light patterns relative to the sample slide by moving the slide stage

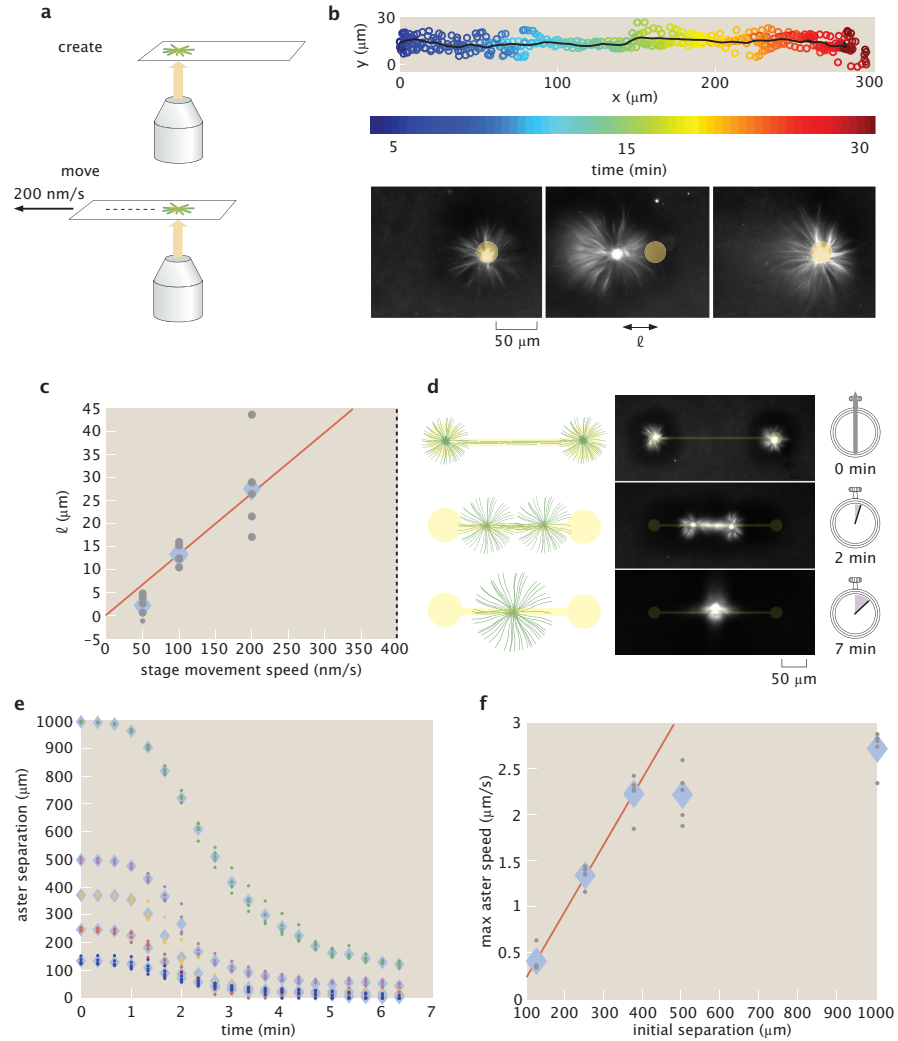


Figure 2.2: Moving and merging operations of asters with dynamic light patterns. **a**, Asters are moved relative to the slide by repositioning the microscope stage. **b**, Overlay of five individual trajectories of aster movement relative to slide moving at 200 nm/s. The line represents the mean trajectory. Time lapse images show the position of the aster relative to the light pattern. ℓ is the displacement of the aster from center of the light pattern. **c**, ℓ versus stage speed. The dotted line at 400 nm/s represents the escape velocity. The red line is a linear fit. **d**, Illustration of the aster merge operation by a connected excitation pattern and the corresponding time series of images. **e**, Distance between merging asters over time for different initial separations. **f**, Maximum speeds of asters as measured from (e). The red line is a linear fit to the first three data points. In (c, e, f) the diamonds represent the mean of 5 experiments and the dots represent individual experiments.

(Fig. 2.2a). We are also able to move asters by directly moving the light pattern, however, moving the stage allows for a greater range of travel. As the stage moves, the asters track with the light pattern, traveling up to hundreds of microns relative to the slide (Fig. 2.2b) (Video 4) (Supplementary Information 2.4.13). The aster maintains a steady state distance ℓ between itself and the light pattern (Fig. 2.2c). We find that asters are always able to track the pattern for stage speeds up to 200 nm/s. At 400 nm/s asters are not able to stay with the pattern, setting an "escape velocity" that is comparable to the motor speeds measured in gliding assays (Supplementary Information 2.4.16). When the stage stops moving, the aster returns to the center of the light pattern, indicating that the aster is experiencing a restoring force. We can characterize aster movement as caused by an effective potential (Supplementary Information 2.4.14), and observe mesoscopic phenomena that may inform the underlying mechanisms of aster motion (Supplementary Information 2.4.15).

Intriguingly, we find that asters formed near each other interact by spontaneously merging. To study this interaction, we construct an aster merger operation, where asters are connected with light (Fig. 2.2d) (Video 5). At the beginning of the merging process, a network of bundled microtubules forms, which connects the asters. The connecting network begins to contract and the asters move towards each other (Fig. 2.2e). The speed at which asters merge (Supplementary Information 2.4.8) increases as a function of linking distance up to a speed of roughly 2.5 $\mu\text{m/s}$ (Fig. 2.2f). The scaling of aster merger speed as a function of distance is similar to the observed relationship of contraction speed as a function of the excitation cylinder size discussed above. We note that the maximum observed merger speed is about an order of magnitude higher than the speeds observed during gliding assays (Supplementary Information 2.4.16), which is analogous to how cell migration speeds can exceed single motor speeds [53]. Our ability to move and merge microtubule asters reveals that they are not steady state structures as previously observed [19], but are dynamic and constantly remodeling.

The capability to perform successive operations remains a fundamental step towards engineering with active matter. Our ability to form dynamic light-defined

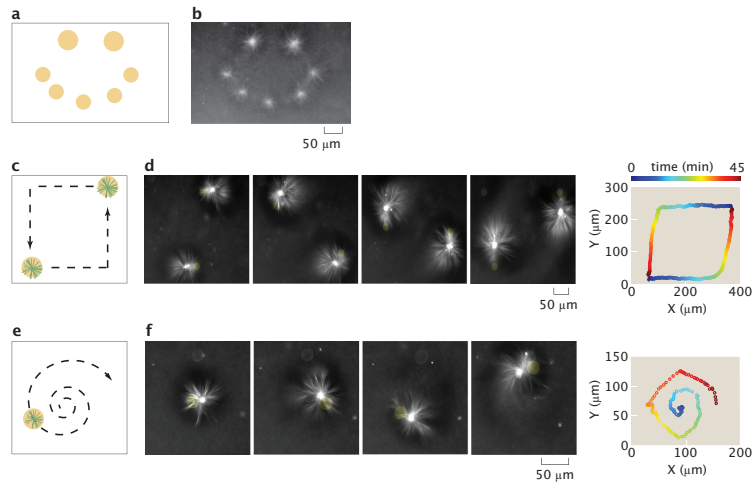


Figure 2.3: Operations for creating and moving asters are composed to make different desired patterns and trajectories. **a**, Sketch for using excitation cylinders to simultaneously pattern asters of different sizes. **b**, Resultant pattern of asters corresponding to **(a)**. **c**, Illustration of simultaneous control of two different aster trajectories, as indicated by the dashed arrows. **d**, Time lapse and the 2D trace of the aster trajectories corresponding to **(c)**. The trajectory trace is color-coded to represent progression in time. **e**, Dynamically projected spiral to illustrate curvilinear motion. **f**, time lapse and the 2D trace of the aster trajectory. Time is color coded as in **(d)**.

compartments of active molecules enables us to execute multiple aster operations. By composing aster creation operations, we are able to form asters of differing sizes and place them at prescribed positions in parallel (Fig. 2.3a, b) (Video 6). Once asters are created, they can be simultaneously moved by using multiple dynamic light patterns (Fig. 2.3c, d) (Video 7). Further, aster trajectories are not limited to rectilinear motion but can be moved along complex trajectories (Fig. 2.3e, f) (Video 8). During movement, there are inflows of microtubule bundles created in the light pattern, which feed into the aster. There are also outflows of microtubules, which appear as comet-tail streams following the asters (Fig. 2.3d, f). These mass flows illustrate some of the complex non-equilibrium dynamics that are introduced by moving boundaries of molecular activity. The new capability to simultaneously generate and manipulate asters provides a basis for “programming” complex systems of interacting non-equilibrium structures.

In our aster merging, moving, and trajectory experiments, we observe fluid flow

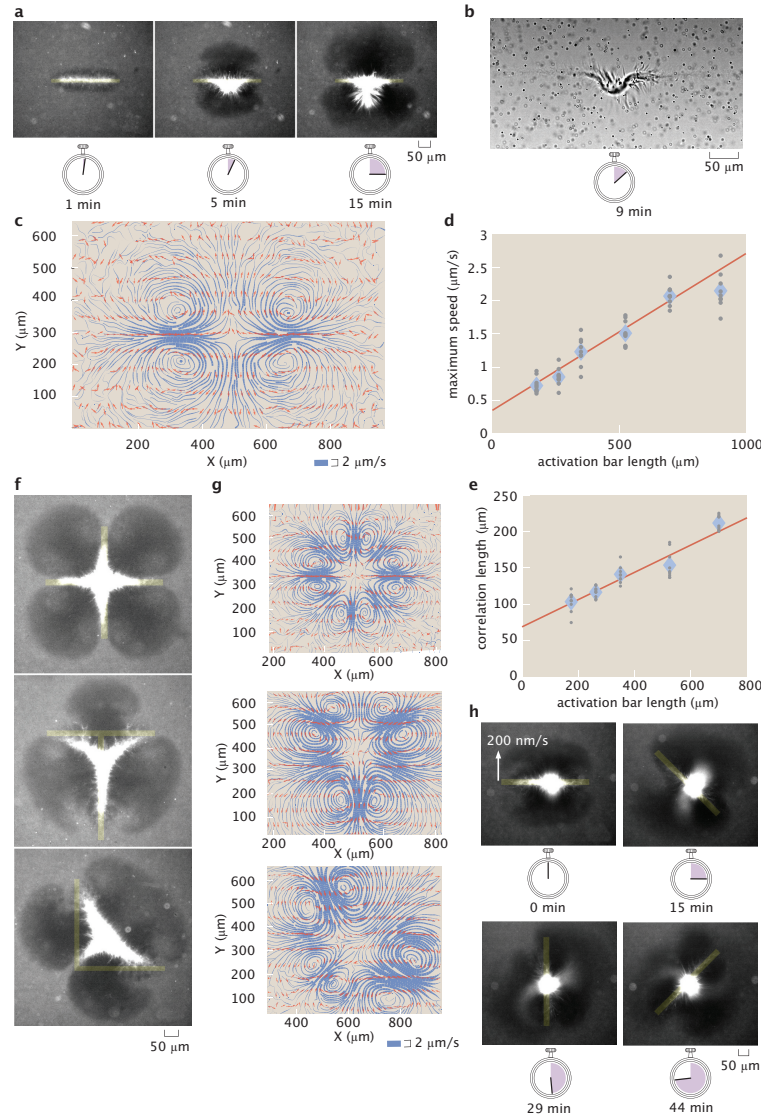


Figure 2.4: Advective fluid flow is created and controlled with patterned light. **a**, Microtubule organization created by an activation bar that is a $350\ \mu\text{m} \times 20\ \mu\text{m}$ rectangular light pattern. Time series demonstrate continuous contraction of microtubules towards the pattern center along the major axis. **b**, Brightfield image of **(a)** shows a contracting microtubule network and tracer particles used to measure fluid flow. **c**, Streamline plots of background buffer flow from **(a)**. The streamline thickness represents the flow speed. The arrows indicate the flow direction. **d**, Averaged maximum flow speed versus activation bar length. **e**, Averaged correlation length (size) of flow field versus activation bar length. **f**, Superposition of activation bars generate different patterns of contractile microtubules. **g**, Corresponding streamline plots. **h**, Time lapse of a light pattern rotating with an edge speed of $200\ \text{nm/s}$. In **(d, e)** the diamonds represents the mean of 9 experiments and the gray dots represent individual experiments. The red line is a linear fit to the data.

of the buffer, as inferred by the advection of microtubules and small fluorescent aggregates. Similar cytoskeletal-driven flow is critical for the development and morphogenesis of various unicellular and multicellular organisms [54, 55, 56, 57, 58, 59, 60].

Based on these observations, we seek to generate and tune flows in our engineered system with light, which may also provide insight into the mechanics of cellular fluid flow. Recent work has used light to thermally induce cytoplasmic flows [61]. Here, we can generate fluid flows with light by activating contractile microtubule networks with the rectangular bar pattern used during aster merging (Fig. 2.4a) (Video 9). Brightfield images reveal a structurally changing microtubule network (Fig. 2.4b) (Video 10), which appears to drive the fluid flow. We observe that there are minimum size and angle limits for these microtubule structures, as well as for asters (Supplementary Information 2.4.17).

We measure the flow fields with tracer particles (Supplementary Information 2.4.18). The pattern of the flow is 2D (Supplementary Information 2.4.19) and stable throughout the experiment (Supplementary Information 2.4.20), consisting of inflows and outflows of microtubules, as illustrated by streamline plots (Fig. 2.4c)(Supplementary Information 2.4.21). The competition of these flows ensures that microtubules do not continuously accumulate in the illuminated region and that the surrounding medium is not completely depleted of microtubules.

We manipulate the properties of the flow field through the geometry of the activation volume. The size (Supplementary Information 2.4.22) and speed of the flow field depend linearly on the length of the activation bar (Fig. 2.4d, e). The scaling of the flow speed is similar to the relationships for both the formation rate versus activation diameter and the aster merging speed versus separation. The positioning and number of inflows, outflows, and vortices are determined by the extrema of the light pattern geometry (Fig. 2.4f, g) (Video 11, Video 12, Video 13). A model that uses a series of point forces following the observed microtubule networks is able to recreate similar inflows and outflows (Supplementary Information 2.4.23), suggesting that forces from microtubule bundles drive the flow.

Furthermore, the shape of the flow field has a temporal dependence on the light pattern. We modulate the flow field to create an “active stir bar” by applying a rotating light pattern (Fig. 2.4h) (Video 14). While simplified active matter systems are able to spontaneously generate global flows [41, 42], *in vivo* cytoskeletal-driven fluid flows can be controlled and highly structured [54, 55, 59]. Our results demonstrate the creation and dynamic manipulation of localized, structured fluid flow in an engineered active matter system for the first time.

In this work, we uncover active matter phenomena through the creation and manipulation of non-equilibrium structures and resultant fluid flows. Our ability to define boundaries of protein activity with light enables unprecedented control of an active matter system’s organization (Supplemental Information 2.4.25). We find scaling rules of contractile networks, movement of non-equilibrium structures, and modulation of flow fields. This framework may be built upon to create active matter devices that control fluid flow. Future work will explore spatiotemporal limits of non-equilibrium structures, the interplay of mass flows and structural changes, and develop new theories of non-equilibrium mechanics and dynamics. Our approach of understanding through construction creates a path towards a generalizable theory of non-equilibrium systems, engineering with active matter, and understanding biological phenomena.

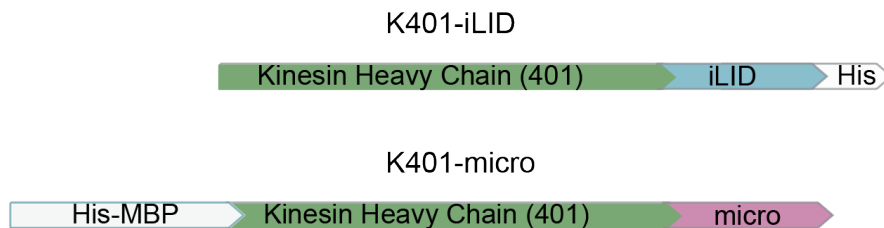


Figure 2.5: Kinesin motor coding regions.

2.3 Methods and Materials

2.3.1 Kinesin Chimera Construction and Purification

To introduce optical control, we implemented the light-induced hetero-dimer system of iLID and SspB-micro [47]. We constructed two chimeras of *D. melanogaster* kinesin K401: K401-iLID and K401-micro (Fig S1).

To construct the K401-iLID plasmid (Addgene 122484), we PCR amplified the coding region of iLID from the plasmid pQE-80L iLID (gift from Brian Kuhlman, Addgene 60408) and used Gibson assembly to insert it after the C-terminus of K401 in the plasmid pBD-0016 (gift from Jeff Gelles, Addgene 15960). To construct the K401-micro plasmid (Addgene 122485), we PCR amplified the coding region of K401 from the plasmid pBD-0016 and used Gibson assembly to insert it in between the His-MBP and micro coding regions of plasmid pQE-80L MBP-SspB Micro (gift from Brian Kuhlman, Addgene 60410). As reported in [47], the MBP domain is needed to ensure the micro domain remains fully functional during expression. Subsequent to expression, the MBP domain can be cleaved off by utilizing a TEV protease site.

For protein purification, we used the His tags that were provided by the base plasmids. For protein expression, we transformed the plasmids into BL21(DE3)pLysS cells. The cells were induced at OD 0.6 with 1 mM IPTG and grown for 16 hours at 18°C. The cells were pelleted and then resuspended in lysis buffer (50 mM sodium phosphate, 4 mM MgCl₂, 250 mM NaCl, 25 mM imidazole, 0.05 mM MgATP, 5 mM BME, 1 mg/ml lysozyme and 1 tablet/50 mL of Complete Protease Inhibitor). After an hour, the lysate was passed through a 30 kPSI cell disruptor to lyse any remaining cells. The lysate was then clarified by an ultra-centrifuge spin at 30,000

g for 1 hour. The clarified lysate was incubated with Ni-NTA agarose resin (Qiagen 30210) for 1 hour. The lysate mixture was loaded into a chromatography column, washed three times with wash buffer (lysis buffer without lysozyme and protease inhibitor), and eluted with 500 mM imidazole. For the K401-micro elution, we added TEV protease at a 1:25 mass ratio to remove the MBP domain. Protein elutions were dialyzed overnight using a 30 kDa MWCO membrane to reduce trace imidazole and small protein fragments. Protein was concentrated with a centrifugal filter (EMD Millipore UFC8030) to 8-10 mg/ml. Protein concentrations were determined by absorption of 280 nm light with a UV spectrometer.

2.3.2 Microtubule Polymerization and Length Distribution

We polymerized tubulin with the non-hydrolyzable GTP analog GMP-CPP, using a protocol based on the one found on the Mitchison lab homepage [62]. A polymerization mixture consisting of M2B buffer (80 mM K-PIPES pH 6.8, 1 mM EGTA, 2 mM MgCl₂), 75 μ M unlabeled tubulin (PurSolutions 032005), 5 μ M tubulin-AlexaFluor647 (PurSolutions 064705), 1 mM DTT, and 0.6 mM GMP-CPP (Jenna Biosciences NU-405S) was spun at \approx 300,000 g for 5 minutes at 2°C to pellet aggregates. The supernatant was then incubated at 37°C for 1 hour to form GMP-CPP stabilized microtubules.

To measure the length distribution of microtubules, we imaged fluorescently labeled microtubules immobilized onto the cover glass surface of a flow cell. The cover glass was treated with a 0.01% solution of poly-L-lysine (Sigma P4707) to promote microtubule binding. The lengths of microtubules were determined by image segmentation. To reduce the effect of the non-uniformity in the illumination, we apply a Bradley adaptive threshold with a sensitivity of 0.001 and binarize the image. Binary objects touching the image border and smaller than 10 pixels in size were removed. To connect together any masks that were “broken” by the thresholding, a morphological closing operation was performed with a 3 pixel \times 3 pixel neighborhood. Masks of microtubules are then converted into single pixel lines by applying a morphological thinning followed by a removal of pixel spurs. The length of a microtubule is determined by counting the number of pixels that

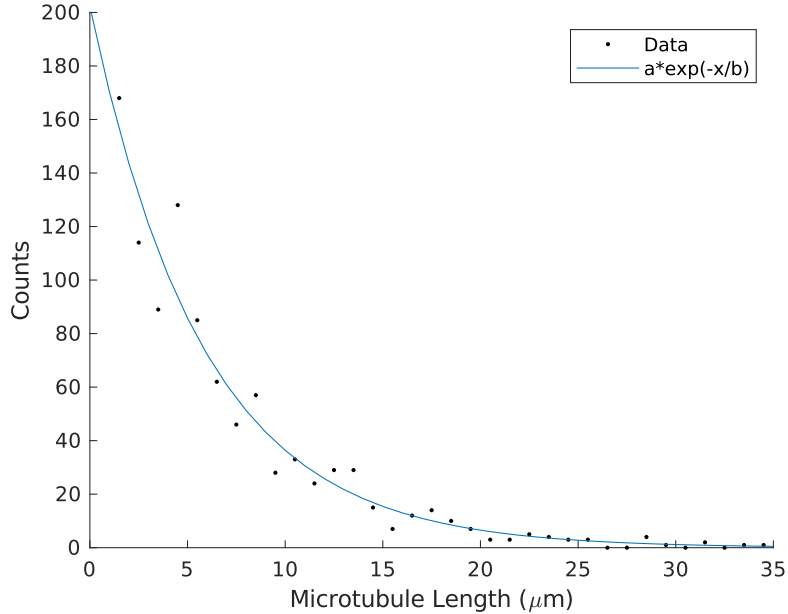


Figure 2.6: Length distribution of microtubules. The mean length given by the data histogram is $7 \pm 0.2 \mu\text{m}$, where the \pm indicates the standard error of the mean. This mean length is similar to the $\approx 6 \mu\text{m}$ mean length given by a fit to an exponential distribution.

make up each line and multiplying by the interpixel distance. For the characteristic microtubule length, we report the mean of the measured lengths (Fig. 2.6). For comparison, we also fit an exponential distribution to the observed histogram. We note that a full distribution of microtubule lengths does not, in general, follow an exponential decay, however, the exponential has been shown to be appropriate for limited length spans [63].

2.3.3 Sample Chambers for Aster and Flow Experiments

For the aster and flow experiments, microscope slides and cover glass are passivated against non-specific protein absorption with a hydrophilic acrylamide coating [64]. The glass is first cleaned in a multi-step alkaline etching procedure that removes organics and the surface layer of the glass. The slides and cover glass are immersed and sonicated for 30 minutes successively in 1% Hellmanex III (Helma Analytics) solution, followed by ethanol, and finished in 0.1 M KOH

solution. After cleaning, the glass is immersed in a silanizing solution of 98.5% ethanol, 1% acetic acid, and 0.5% 3-(Trimethoxysilyl)propylmethacrylate (Sigma 440159) for 10-15 min. After rinsing, the slides are immersed overnight in a degassed 2 % acrylamide solution with 0.035% TEMED and 3 mM ammonium persulfate. Just before use, the glass is rinsed in distilled water and nitrogen dried. Parafilm M gaskets with pre-cut 3 mm wide channels are used to seal the cover glass and slide together, making a flow cell that is $\approx 70\mu\text{m}$ in height. After the addition of the reaction mixture, a flow cell lane is sealed with a fast setting silicone polymer (Picodent Twinsil Speed).

2.3.4 Reaction Mixture and Sample Preparation for Aster and Flow Experiments

For the aster and flow experiments, K401-micro , K401-iLID , and microtubules were combined into a reaction mixture, leading to final concentrations of $\approx 0.1\ \mu\text{M}$ of each motor type and 1.5-2.5 μM of tubulin. Concentrations refer to protein monomers for the K401-micro and K401-iLID constructs and the protein dimer for tubulin. To minimize unintended light activation, the sample was prepared under dark-room conditions, where the room light was filtered to block wavelengths below 580 nm (Kodak Wratten Filter No. 25). The base reaction mixture provided a buffer, an energy source (MgATP), a crowding agent (glycerol), a surface passivating polymer (pluronic F-127), oxygen scavenging components to reduce photobleaching (glucose oxidase, glucose, catalase, Trolox, DTT), and ATP-recycling reagents to prolong motor activity (pyruvate kinase/lactic dehydrogenase, phosphoenolpyruvic acid). The reaction mixture consisted of 59.2 mM K-PIPES pH 6.8, 4.7 mM Mg Cl₂, 3.2 mM potassium chloride, 2.6 mM potassium phosphate, 0.74 mM EGTA, 1.4 mM MgATP (Sigma A9187), 10% glycerol, 0.50 mg/mL pluronic F-127 (Sigma P2443), 0.22 mg/ml glucose oxidase (Sigma G2133), 3.2 mg/ml glucose, 0.038 mg/ml catalase (Sigma C40), 5.4 mM DTT, 2.0 mM Trolox (Sigma 238813), 0.026 units/ μl pyruvate kinase/lactic dehydrogenase (Sigma P0294), and 26.6 mM phosphoenolpyruvic acid (Beantown Chemical 129745).

We note that the sample is sensitive to the ratio of motors and microtubules and the absolute motor concentration. When the motor concentration is below 0.1

μM for K401-micro and K401-iLID, light patterns are able to create microtubule bundles or lattices of small asters, similar to the phases observed as functions of motor concentration described in [19]. If this motor concentration is above $\approx 2 \mu\text{M}$, however, the number of binding events between inactivated K401-micro and K401-iLID proteins is sufficient to cause the spontaneous microtubule bundling and aster formation.

2.3.5 Sample Preparation for Gliding Assay

For the gliding assay experiments, microscope slides and cover glass are coated with antibodies to specifically bind motor proteins. First, alkaline cleaned cover glass and ethanol scrubbed slides were prepared and $5 \mu\text{L}$ flow chambers were prepared with doubled sided tape. Motors were bound to the surface by successive incubations of the chamber with $400 \mu\text{g}/\text{mL}$ penta-His antibody (Qiagen 34660) for 5 min, $10 \text{ mg}/\text{ml}$ whole casein (Sigma C6554) for 5 min, and finally motor protein ($1 \text{ mg}/\text{mL}$ in M2B) for 5 min. Unbound motors were washed out with M2B buffer, then AlexaFluor 647 labeled GMP-CPP stabilized microtubules in M2B with 5 mM MgATP and 1 mM DTT were flowed in.

2.3.6 Preparation of Tracer Particles

To measure the fluid velocity, we used $1 \mu\text{m}$ polystyrene beads (Polysciences 07310-15) as tracer particles. To passivate the hydrophobic surface of the beads, we incubated them overnight in M2B buffer with $50 \text{ mg}/\text{ml}$ of pluronic F-127. Just before an experiment, the pluronic coated beads are washed by pelleting and resuspending in M2B buffer with $0.5 \text{ mg}/\text{ml}$ pluronic to match the pluronic concentration of the reaction mixture.

2.3.7 Microscope Instrumentation

We performed the experiments with an automated widefield epifluorescence microscope (Nikon TE2000). We custom modified the scope to provide two additional modes of imaging: epi-illuminated pattern projection and LED gated transmitted light. We imaged light patterns from a programmable DLP chip (EKB

TEchnologies DLP LightCrafter™ E4500 MKII™ Fiber Couple) onto the sample through a user-modified epi-illumination attachment (Nikon T-FL). The DLP chip was illuminated by a fiber coupled 470 nm LED (ThorLabs M470L3). The epi-illumination attachment had two light-path entry ports, one for the projected pattern light path and the other for a standard widefield epi-fluorescence light path. The two light paths were overlapped with a dichroic mirror (Semrock BLP01-488R-25). The magnification of the epi-illuminating system was designed so that the imaging sensor of the camera (FliR BFLY-U3-23S6M-C) was fully illuminated when the entire DLP chip was on. Experiments were run with Micro-Manager [65], running custom scripts to controlled pattern projection and stage movement. For the transmitted light path, we replaced the standard white-light brightfield source (Nikon T-DH) with an electronically time-gated 660 nm LED (ThorLabs M660L4-C5). This was done to minimize light-induced dimerization during bright field imaging.

2.4 Data Acquisition, Analysis, and Supplemental Discussion

2.4.1 Aster Distribution in 3D

From Z-stack imaging, we observe that asters are complex 3D structures (Fig. 2.7). By analyzing the microtubule density in Z, we find that asters form near the midpoint of the sample plane (Fig. 2.8a). Further, we show that these are symmetric structures by fitting the intensity profiles in the Y plane and Z plane to Gaussians (Fig. 2.8b, c).

2.4.2 Comparisons with Similar Systems

2.4.2.1 Microtubule Vortices

The original microtubule-motor system [15, 19] is contractile and shows the formation of microtubule vortices in addition to asters. Microtubule vortices have not been observed in our experiments, however. This is likely due to the substantial differences between the boundary conditions. Experiments where

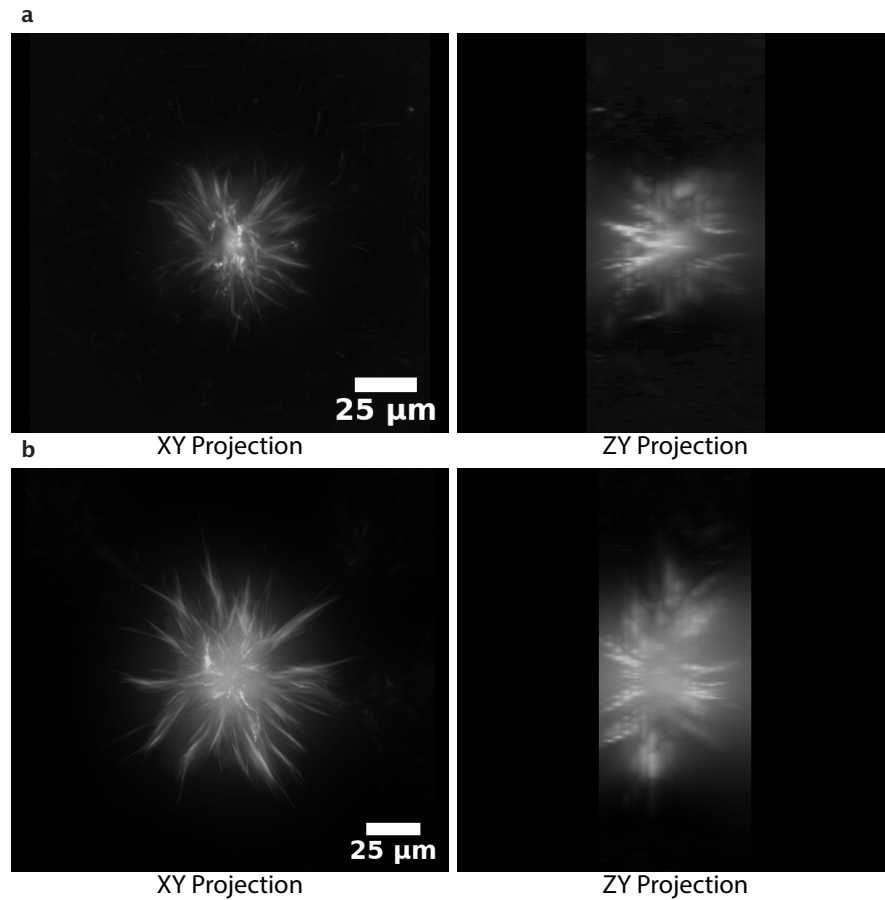


Figure 2.7: 3D projections of asters from Z-stacks imaged with a 20x objective. **a**, Aster generated with a 100 μm disk (Video 1). **b**, Aster generated with a 300 μm disk (Video 2). The XY plane is along the plane of the sample slide. The ZY plane is orthogonal to the sample slide and the image is constructed by interpolating over 18 Z-slices spaced by 4 μm .

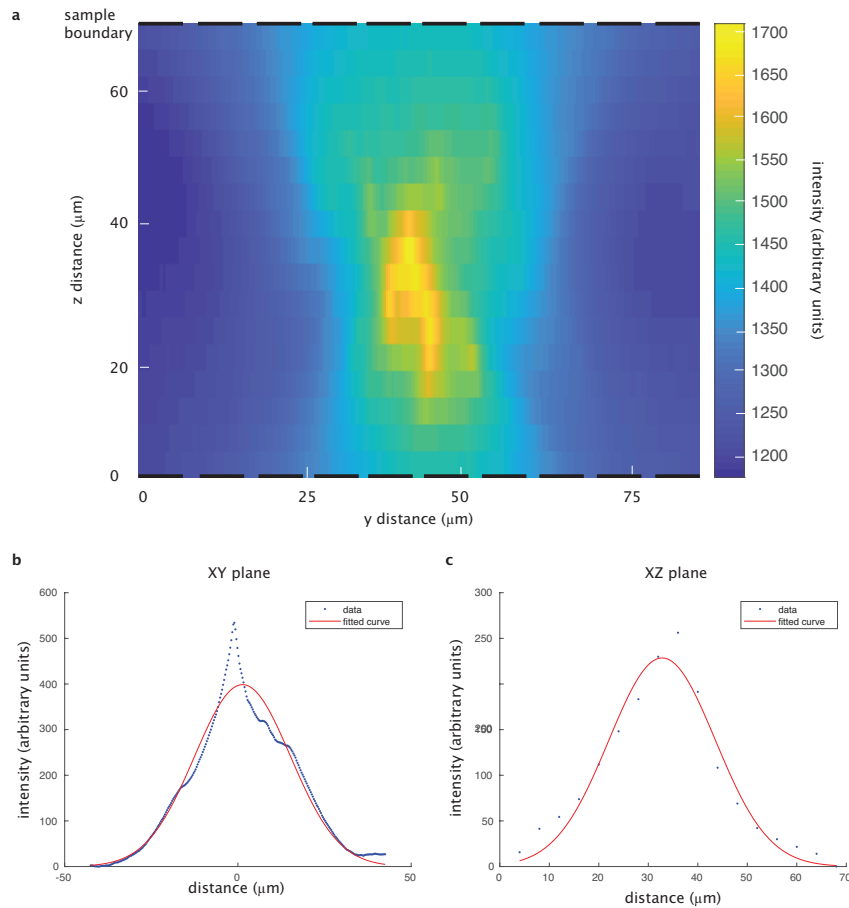


Figure 2.8: Analysis of microtubule distribution in 3D. **a**, Heatmap of microtubule distribution in the YZ plane shown in Fig. 2.7a. Sample boundaries, defined by the coverslips, are denoted by the dashed lines. **b**, Gaussian fit to the microtubule density in the middle slice of the Y plane. **c**, Gaussian fit to the microtubule density along the Z plane.

vortices are reported have a channel spacing of 5 μm , while our experiments have a channel spacing of $\approx 70 \mu\text{m}$. A large microtubule vortex forms with a boundary that is 90 μm in diameter [15], however, our boundaries are 18 mm x 3 mm. Further, our experiments use GMPCPP stabilized microtubules with an average length of 7 μm , while the work reporting vortices uses taxol stabilized filaments with length range of $\approx 10\text{-}100 \mu\text{m}$. There may also be a significant difference between the acrylamide surface chemistry we use and the agarose chemistry used in the other work.

2.4.2.2 Extensile vs. Contractile

We note that our experimental system results in a contractile network rather than an extensile gel. Recent works have shown that conditions leading to a contractile system require long flexible filaments that are capable of buckling and that undergo limited steric interactions [49, 66]. In contrast, the extensile active gel or the active nematic relies on high concentrations of depletion agents to preform bundles of short and stiff filaments, unlike in our system. This suggests that the lack of extensile behavior we observe is unrelated to the optically-dimerizable motors but rather the parameters of the microtubule length and depletion agent. Therefore, there is no inherent limitation in the application of optically-dimerizable motors under extensile conditions.

2.4.3 Microscopy Protocol

Samples were imaged at 10X (Fig 1c, 1e, 1f, 2d, 4a, 4f, and 4h) or 20X (Fig. 1d, 2b, 3b, 3d, and 3e). For Figures 2e and 2f, the distance span of the merger experiments required us to pool data taken at 10X (500 μm and 1000 μm separations) and 20X magnifications (175 μm , 250 μm , and 350 μm separations). For the formation, merging, and movement experiments represented in Figures 1-3, the images of the fluorescent microtubules were acquired every 20 s. For each time point, a Z-stack of 5 slices spaced by 10-15 μm is taken. For the flow experiments represented in Figure 4, a brightfield image and subsequent fluorescent image were acquired every 4 seconds to observe the tracer particles and microtubules,

respectively, without Z-stack imaging. The increased frame rate was needed to ensure sufficient accuracy of the particle velocimetry. For all experiments, we activated light-induced dimerization in the sample every 20 s with a brief 300 msec flash of 2.4 mW/mm² activation light from a \approx 470 nm LED. The rate of activation was based on the estimated off-rate of the iLID-micro complex [47] of \approx 30 s. The duration of the activation light was empirically determined, by gradually increasing the time in 50 msec increments until we observed the formation of an aster. We note that higher frequencies of activation or longer pulse duration result in contractile activity outside of the light pattern. Typically, one experiment was run per sample. Individual samples were imaged for up to 1 hour. We placed the time limitations on the sample viewing to minimize effects related to cumulative photobleaching, ATP depletion, and global activity of the light-dimerizable proteins. After several hours, inactivated "dark" regions of the sample begin to show bundling of microtubules.

2.4.4 Measuring Aster Spatial Distribution with Image Standard Deviation

We interpret the pixel intensity from the images as a measure of the microtubule density. Image standard deviation σ is a measure of the width of an intensity-weighted spatial distribution over a region of interest, ROI. We use σ to characterize how the spatial distribution of microtubules evolves in time. For each time point, we first normalize each pixel value $I(x, y)$ by the total pixel intensity summed across the ROI

$$I_{\text{norm}}(x, y) = \frac{I(x, y)}{\sum_{x, y \in \text{ROI}} I(x, y)} \quad (2.1)$$

where $I(x, y)$ is the raw intensity of the pixel at position (x, y) after background subtraction. To find σ , we define the image variance σ^2 of the intensity-weighted spatial distribution as

$$\sigma^2 = \sum_{x, y \in \text{ROI}} [(x - \bar{x})^2 + (y - \bar{y})^2] I_{\text{norm}}(x, y), \quad (2.2)$$

where coordinates \bar{x} and \bar{y} are the center of the intensity distribution

$$\bar{\mathbf{x}} = \sum_{\mathbf{x} \in \text{ROI}} \mathbf{x} I(\mathbf{x}). \quad (2.3)$$

2.4.5 Characteristic Size of an Aster

2.4.5.1 Determining Characteristic Size

As seen in Fig. 2.7, the irregularity of aster arm spacings and lengths presents very challenging segmentation issues for the detailed modeling of the microtubule distribution. Instead, we chose to determine a single characteristic size to represent the spatial distribution of the aster. First, we perform a maximum projection over the Z-stack for each time point to create a 2D image in the XY plane. To represent the projected 2D image, we chose the image standard deviation approach (Supplementary Information 2.4.4) to integrate over the variations in the XY plane. We define the characteristic aster size as the image standard deviation σ after ≈ 15 min of activation. The characteristic size is used to compare with order-of-magnitude scaling arguments (Supplementary Information 2.4.12).

2.4.6 Image Analysis of Asters

2.4.6.1 Image Preparation

At each time point, each Z-stack of images is summed into a single image in the XY plane. We process each XY image to correct for the non-uniformity in the illumination and background intensity. We “flatten” the non-uniformity of the image with an image intensity profile found in the following process. We take the first frame of the experiment and perform a morphological opening operation with an 80 pixel disk followed by a Gaussian smoothing with a 20-pixel standard deviation. The resulting image is then normalized to its maximum pixel intensity to generate the image intensity profile. Images are flattened by dividing them by the intensity profile. We note that this strategy depends on there being a uniform density of microtubules in the first frame.

Once images are flattened, the background is found by taking the last frame of aster formation and calculating the mean intensity of the activated region that is devoid of microtubules. Images are subtracted by this background intensity and thresholded so that any negative values are set to zero.

2.4.6.2 Defining the Regions of Interest

As mentioned in Supplementary Information 2.4.4, we determine the image standard deviation over a region of interest (ROI). For the formation experiments, we define the region of activation as the disk encompassing the aster and the region devoid of microtubules around the aster, after ≈ 15 min of activation, when formation is complete. To identify this region, we segment the low intensity region around the aster. The low intensity region around the aster is found by subtracting the final frame of aster formation from the first frame of the image acquisition. After subtraction, the void region is the brightest component of the image. We segment this region by performing an intensity and size threshold to create a mask. The aster-shaped hole in the mask is then filled. Using the perimeter of the mask, we calculate the diameter of the disk region of activation.

For analyzing the images for the decay process, we alternatively take a region of interest centered on the aster position (from the last frame of aster formation and found using the intensity weighted center) and proportional to the size of the aster in order to reduce the contribution of microtubules diffusing in from the boundary. This proportionality constant was chosen as the ratio of the ROI diameter to the aster diameter for the aster formed with the 50 μm disk, which is 1.63.

2.4.7 Reversibility of Aster Formation and Decay

To show that aster decay is driven by motors reverting to monomers as opposed to irreversible events such as ATP depletion or protein denaturation, we provide an illustrative experiment of aster formation followed by decay followed again by aster formation. Imaging for this experiment was performed at 20X to increase the

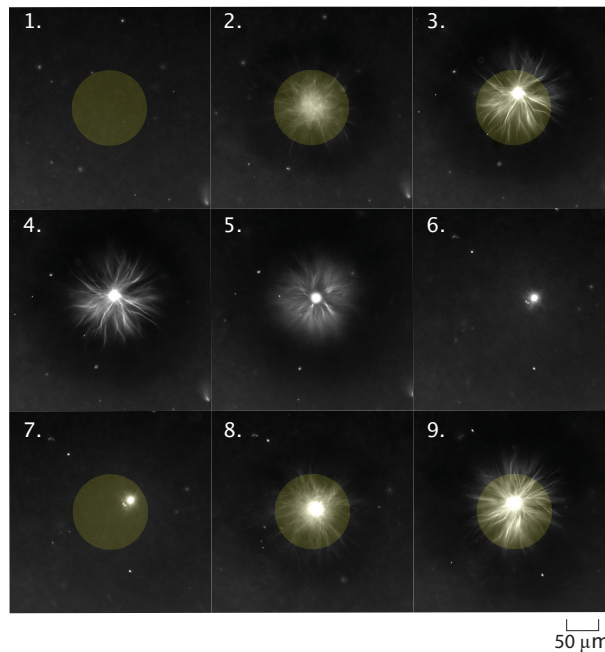


Figure 2.9: Time series of light induced aster formation, decay, then formation. First formation frames are at time points $t = (1) 0, (2) 6.7, \text{ and } (3) 16.3$ min. Aster decay frames are for $t = (4) 16.7, (5) 25, \text{ and } (6) 112.7$ min. Second aster assembly frames are $t = (7) 113, (8) 120, \text{ and } (9) 129.3$ min.

spatial resolution. We note that asters do not completely decay, as it is observed in panel 6 of Fig. 2.9 that the central core of the aster persists.

2.4.8 Speed and Characteristic Time Scales of Formation and Merging

In order to compare the boundary dependence of our contraction behavior to other contractile networks, we calculate the max speeds and characteristic times of contraction and aster merger as described in [48, 49, 50]. We first find the characteristic time by fitting a model to our experimental data and then use this value to calculate the maximum speed. As in [48], we fit to a model of a critically damped harmonic oscillator,

$$L(t) = L_{\text{fin}} + (L_{\text{init}} - L_{\text{fin}}) \left(1 + \frac{t}{\tau}\right) e^{-\frac{t}{\tau}}, \quad (2.4)$$

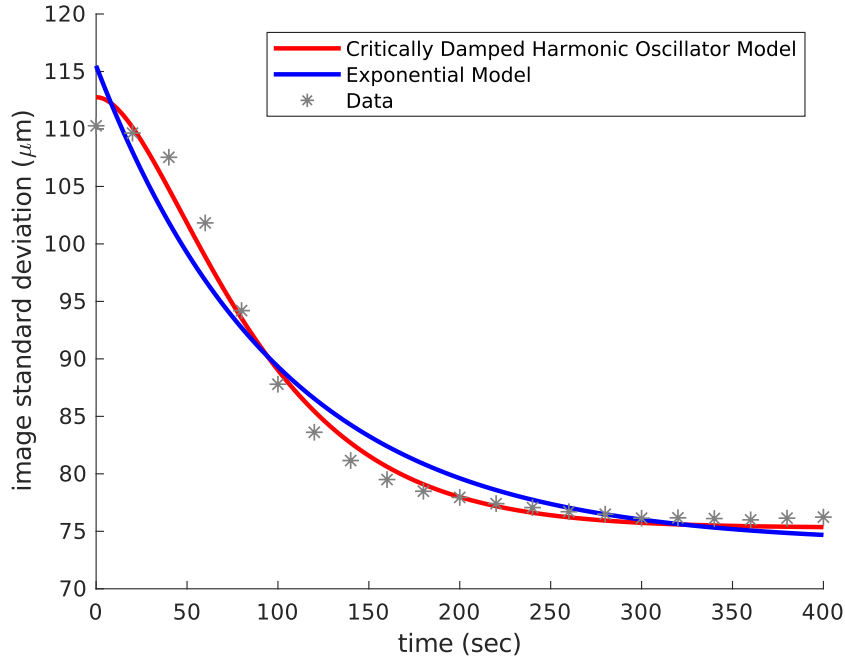


Figure 2.10: A comparison of model fittings for a contracting aster experiment.

where L_{init} is the initial size of the network, L_{fin} is the final network size, and τ is the characteristic time of contraction. This model was developed to describe a contractile actomyosin gel, which shares similar dynamics with our own system. We apply this fit on time points after the initial lag phase, which was empirically determined to end at one minute. While we tried fitting to an exponential function, we found that the harmonic oscillator model was more robust across excitation length scales (Fig. 2.10).

We find that the characteristic times show a general lack of sensitivity to system size for our range of lengths (Fig. 2.11), similar to [50]. The characteristic time is roughly 1 to 2 minutes, comparable to the times reported in [50].

We calculate the maximum speed of contraction or merger, $v_{\text{max}} = -\frac{dL(t_{\text{max}})}{dt}$, by finding the time $t = t_{\text{max}}$ that satisfies $\frac{d^2L(t_{\text{max}})}{dt^2} = 0$. First, we calculate the second derivative of Eq. 2.4,

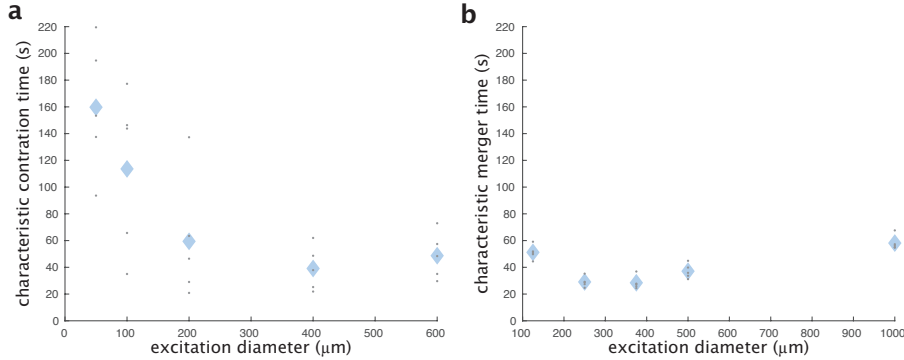


Figure 2.11: Characteristic times for contraction and merger as functions of activation length scales. **a**, Characteristic time for aster formation as a function of the excitation diameter. **b**, Characteristic time for aster merging as a function of the initial distance between asters.

$$\frac{d^2L(t)}{dt^2} = \frac{(L_{\text{init}} - L_{\text{fin}})(t - \tau)}{\tau^3} e^{-\frac{t}{\tau}}. \quad (2.5)$$

Based on this equation, it is apparent that the maximum speed occurs at $t_{\text{max}} = \tau$. The maximum speed is then defined as $v_{\text{max}} = -\frac{dL(\tau)}{dt}$. We calculate $-\frac{dL(t)}{dt}$ by taking the first derivative of Eq. 2.4,

$$\frac{dL(t)}{dt} = \frac{t(L_{\text{init}} - L_{\text{fin}})}{\tau^2} e^{-\frac{t}{\tau}}, \quad (2.6)$$

then set $t = \tau$ to find the maximum speed,

$$v_{\text{max}} = \frac{dL(\tau)}{dt} = \frac{L_{\text{init}} - L_{\text{fin}}}{e\tau}. \quad (2.7)$$

This v_{max} is the reported contraction or merger speed.

2.4.9 Comparison to Light Activated Actomyosin Networks

A system that shows some similar behavior to ours is the light activated actomyosin network in [48]. Here, we note the similarities and differences between the two systems. In the actomyosin network, the actin filaments are globally and permanently crosslinked by the myosin motors in both the dark and the light. In

the light, motors are permanently activated. Light patterns generate a localized contraction of the global actomyosin network. Since the contracting region is still linked to the rest of the actomyosin network, deformations are propagated throughout the entire network.

In contrast, our system starts with unlinked microtubule filaments. Light patterns activate linkages of motors to create a localized contractile network with a free boundary. Thus, there are no connections to an external network, unlike the actomyosin system. Further, the reversibility of these links allows the networks to remodel and to resolve after contraction.

A key similarity between the two systems is the observation that contraction speed increases linearly with the size of the excitation region. A recent theoretical treatment [49] provides a generic model for this observation. Their results in Box 1 Panel C predict a linear scaling of contraction speed versus size for 1D, 2D, and 3D networks. For a 1D network, the contraction speed $\frac{dL}{dt}$ is related to the length L of the network by the contractility constant χ as

$$\frac{dL}{dt} \approx \chi L. \quad (2.8)$$

2.4.10 Analysis of Aster Decay

When the activation light is removed, the iLID-micro dimer begins to dissociate, leading to un-crosslinked microtubules. The original work where iLID is designed and characterized show that the formation and reversion half-lives of individual iLID-micro heterodimers are on the order of 30 seconds [47]. Our empirical determination that sharp localization of contractile forces within the light pattern requires pulsing the light pattern every 20 seconds (Supplementary Information 2.4.3), in addition to the characterization of other iLID and LOV domain based systems [67, 68, 69, 70], supports the notion that the reversion rate of kinesin-fused iLID proteins is similarly on the tens of seconds time scale. We note that the motor density has been predicted and observed to increase exponentially towards the aster center [44]. We therefore expect the central region of the aster

to decay more slowly than an individual motor link. This may explain why asters appear to decay on the order of tens of minutes (Fig. 1c), rather than tens of seconds.

For an ideal 2D Gaussian spatial distribution of diffusing particles starting with a finite radius of w , we expect

$$p(r, t) = \frac{1}{\pi(4Dt + w^2)} e^{-r^2/(4Dt+w^2)}, \quad (2.9)$$

where D is the diffusion coefficient.

The variance σ_{Gauss}^2 of this distribution as a function of time t is given by

$$\sigma_{\text{Gauss}}^2(t) = 4Dt + w^2. \quad (2.10)$$

The variance σ_{Gauss}^2 increases linearly with t with a slope of $4D$.

We characterize the aster decay process by measuring the image variance σ^2 , as a function of time, as described in (SI. 2.4.4). Images are first processed as described in (SI. 2.4.6). Although our spatial distributions are not strictly Gaussian, we observe that for our data that σ^2 increases linearly with t (Fig. 2.12), which suggests that the decay process is described by the diffusion of unbound microtubules. By analogy to the 2D ideal Gaussian case, we calculate an effective diffusion coefficient of our distributions by a linear fit of σ^2 versus time and finding the diffusion coefficient from the slope. This gives us a diffusion coefficient in units of $\mu\text{m}^2/\text{s}$.

We find the diffusion coefficient by applying a linear fit to time points that occur after 200 seconds.

2.4.11 Diffusion Coefficient of a Microtubule

We estimate the diffusion coefficient for a single microtubule to compare with the effective diffusion coefficient we estimate for aster decay. The diffusion coefficient D for an object in liquid media can be calculated from the drag coefficient

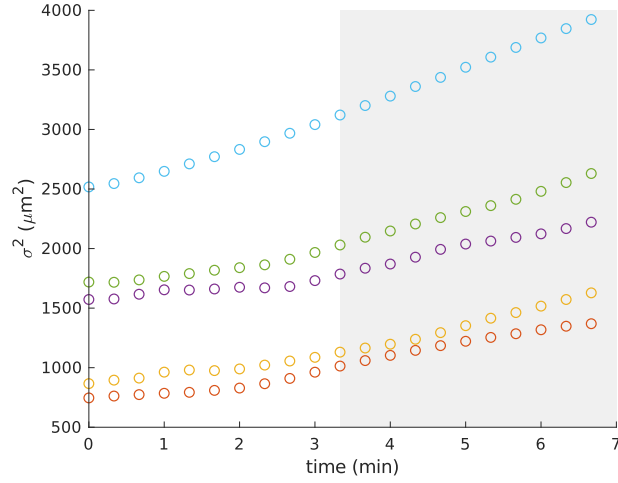


Figure 2.12: Plot of mean variance of image intensity as a function of time for different initial aster sizes. The shaded region is treated as part of the linear regime. The measure of time is relative to the beginning of aster decay.

γ

$$D = \frac{k_B T}{\gamma}, \quad (2.11)$$

where k_B is the Boltzmann constant and T is the temperature, for which we use 298 K. We model a microtubule as a 7 μm long cylinder (SI. 2.3.2) with a radius of 12.5 nm. The drag coefficients for a cylinder have been found previously [71] for motion either parallel γ_{\parallel} or perpendicular γ_{\perp} to the long axis of the cylinder

$$\begin{aligned} \gamma_{\parallel} &= \frac{2\pi\eta L}{\ln(L/2r) - 0.20}, \\ \gamma_{\perp} &= \frac{4\pi\eta L}{\ln(L/2r) + 0.84}. \end{aligned} \quad (2.12)$$

Here, L is the length of the cylinder, r is its radius, and η is the viscosity of the fluid, which we estimate to be $2 \times 10^{-3} \text{ Pa} \cdot \text{s}$ (SI 2.4.24). Using the parameters detailed above, we calculate $D_{\parallel} = 0.3 \mu\text{m}^2/\text{s}$ and $D_{\perp} = 0.2 \mu\text{m}^2/\text{s}$. We assume that the larger diffusion coefficient dominates and thus use D_{\parallel} , the longitudinal diffusion coefficient, as the diffusion coefficient for a single microtubule in Fig. 1e.

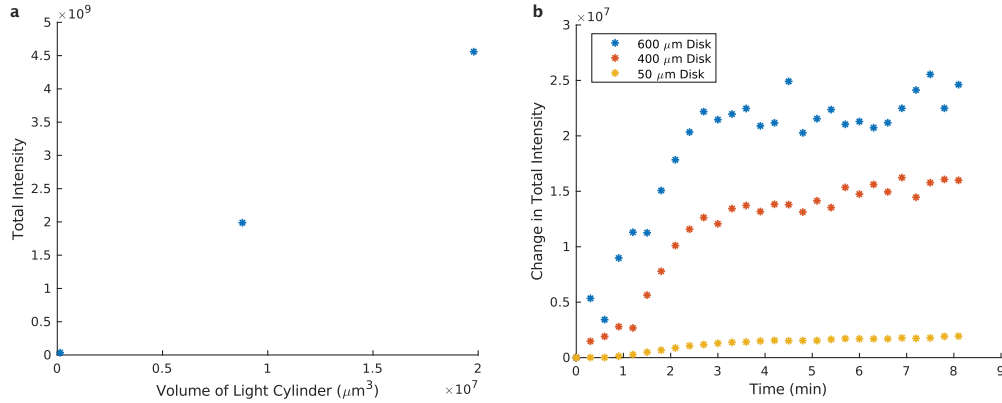


Figure 2.13: Measuring the conservation of labeled fluorescent microtubules in the excitation region during aster formation. **a**, Total intensity of excitation region as a function of volume of light cylinder averaged during aster formation. Measurements are for light disks with diameters 50, 400, and 600 μm . **b**, Change in total intensity inside of the excitation region as a function of time

2.4.12 Scaling Arguments for Aster Size and Comparison to Data

We consider how the total number of microtubules in an aster relates to the volume of the projected light pattern. We are projecting a disk pattern of light on the sample from below. The channel is a constant height, $z \approx 70 \mu\text{m}$. We therefore treat the light excitation volume as a cylinder $V_{\text{light}} = \frac{1}{4}\pi z d_{\text{light}}^2$ where d_{light} is the diameter of the excitation disk. If we look at experimental data, we see evidence of a linear relationship between the light volume and the number of microtubules that are present during aster formation (Fig. 2.13a). The implication of this observation is that the density ρ of microtubules is uniform. Furthermore, we see that after the initial contraction event, the total integrated fluorescence of the excited region remains constant (Fig. 2.13b), indicating that the total number of microtubules N is constant during aster formation.

Based on these observations, we assume that the number of microtubules N in the aster is given by

$$N \approx \rho V_{\text{light}}. \quad (2.13)$$

From Supplementary Information 2.4.1, we observe that asters have a roughly spheroidal symmetry. For an order-of-magnitude estimate of how aster size scales

with the volume of light, we assume the characteristic length of the aster L_{aster} is given by the diameter of an effective sphere which scales with microtubule number as

$$L_{\text{aster}} \propto N^{1/3}. \quad (2.14)$$

and thus

$$L_{\text{aster}} \propto V_{\text{light}}^{1/3}. \quad (2.15)$$

As noted above, the volume defined by the activation light is a cylinder, then

$$V_{\text{light}} \propto d_{\text{disk}}^2. \quad (2.16)$$

From these last two equations, we arrive at the scaling relationship between aster size and excitation disk size

$$L_{\text{aster}} \propto d_{\text{disk}}^{2/3}. \quad (2.17)$$

We made a power law fit with a fixed exponent of 2/3 to the data shown in Fig. 1f. Though we cannot strictly rule out other exponents, we show the fit to demonstrate that the scaling argument determined exponent is at least consistent with the data.

2.4.13 Tracking of Moving Aster

For each time point, we sum over the z-stack to form a single image. The image is then passed through a morphological top-hat filter with a structure element of a 100 pixel disk to “flatten” non-uniformities in the illumination. The image is then projected into a 1D intensity profile. We project onto the x-axis by summing along the line that passes through the center of the excitation disk with a 100 pixel window in y. Aster centers are then found at each frame by fitting the intensity profiles to Gaussian functions.

For 2D tracking, the movement of the aster is found by comparing the centroid of the aster in each frame. The raw images are processed using a Gaussian filter with a standard deviation of 1 pixel, followed by thresholding to eliminate the background noise.

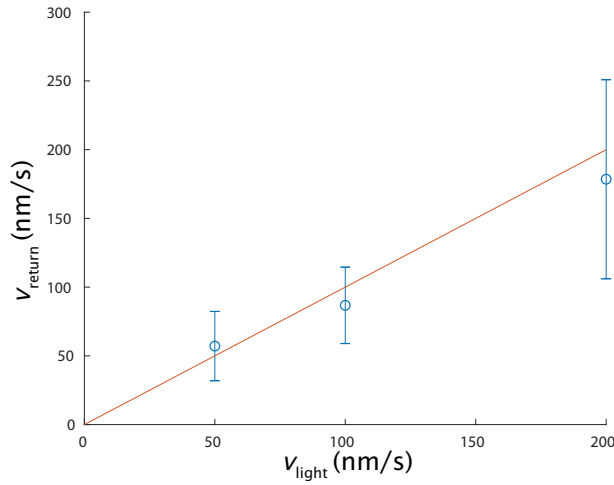


Figure 2.14: The speed at which an aster returns to the center of the light pattern once the pattern stops moving. Red line is a plot of $y = x$.

2.4.14 Effective Potential of a Moving Aster

When the light pattern moves, we observe that the aster appears to be pulled in tow behind the light pattern, perhaps by the aster arms or newly-formed microtubule bundles in the light pattern. Further, when the light pattern stops moving at speed v_{light} , we observe that the aster immediately returns to the center of the light pattern at speed v_{return} . From the Fig. 2.14, we see that

$$v_{\text{return}} \approx v_{\text{light}}. \quad (2.18)$$

This is the behavior expected for an object under the influence of a potential at low-Reynolds-number, where the aster has negligible momentum and the forces are essentially instantaneous. These observations support the notion that a moving aster can be modeled as being in an effective potential. First, we model the observed behavior with a generic potential without any assumption of the mechanistic cause of the potential and then numerically compare these results to the estimated optical tweezer effects of the excitation light pattern.

We estimate the potential and the forces acting on a moving aster from the viscous drag of the background fluid, in an analogous way to how this is done for objects trapped in an optical tweezer [72]. If we assume that the aster is a

spherical object of radius a and is moving with speed v_{light} , it will experience a viscous drag force F_{drag} :

$$F_{\text{drag}} = 6\pi\eta av_{\text{light}}, \quad (2.19)$$

where η is the fluid viscosity. F_{drag} is equal to the force F_{pull} that is pulling the aster towards the light pattern. From the results of Fig. 2c, we note the observed distance shift ℓ of the aster from the center of the moving light pattern is roughly linear with excitation disk movement speed v_{light} . The linearity of ℓ versus v_{light} implies that F_{pull} acts like a spring:

$$F_{\text{pull}} \approx k_{\text{spring}}\ell, \quad (2.20)$$

where k_{spring} is the spring constant. Setting these two forces equal gives a spring constant of

$$k_{\text{spring}} \approx \frac{6\pi\eta av_{\text{light}}}{\ell}. \quad (2.21)$$

The effective potential U_{pull} for this force is

$$U_{\text{pull}} = \frac{1}{2}k_{\text{spring}}\ell^2. \quad (2.22)$$

The aster in Fig. 2c is $\approx 25 \mu\text{m}$ in diameter. Assuming that $\eta \approx 2 \times 10^{-3} \text{ Pa}\cdot\text{s}$ (SI 2.4.24), we find that $k_{\text{spring}} \approx 3 \times 10^{-15} \text{ N}/\mu\text{m}$. For the maximum observed displacement of $\ell \approx 30 \mu\text{m}$, the energy stored in the potential, or equivalently, the work done by the system to return the aster back to the center of the light pattern is $\approx 300 k_B T$.

The spring constant of an optical tweezer trapping polystyrene spheres is $\approx 1 \times 10^{-9} \text{ N}/\mu\text{m}$ for a $\approx 1000 \text{ mW}$ laser beam focused to $\approx 1 \mu\text{m}$ diameter [73]. Accounting for light intensity, we estimate the spring constant to be $\approx 1 \times 10^{-12} \text{ N}/\mu\text{m}$ per $\text{mW}/\mu\text{m}^2$. In comparison, our light pattern has intensity of $2.4 \text{ mW}/\text{cm}^2$. The light is on only for 0.3 sec every 20 sec (SI 2.4.3), giving a time averaged intensity of $0.036 \text{ mW}/\text{cm}^2$. The estimated upper bound spring constant from the light pattern due to optical tweezing effects is $\approx 3.6 \times 10^{-22} \text{ N}/\mu\text{m}$, roughly a factor of 10^7 weaker than the spring constant we observe. Further, we note that it is a generous assumption that a microtubule aster is refractile as a polystyrene

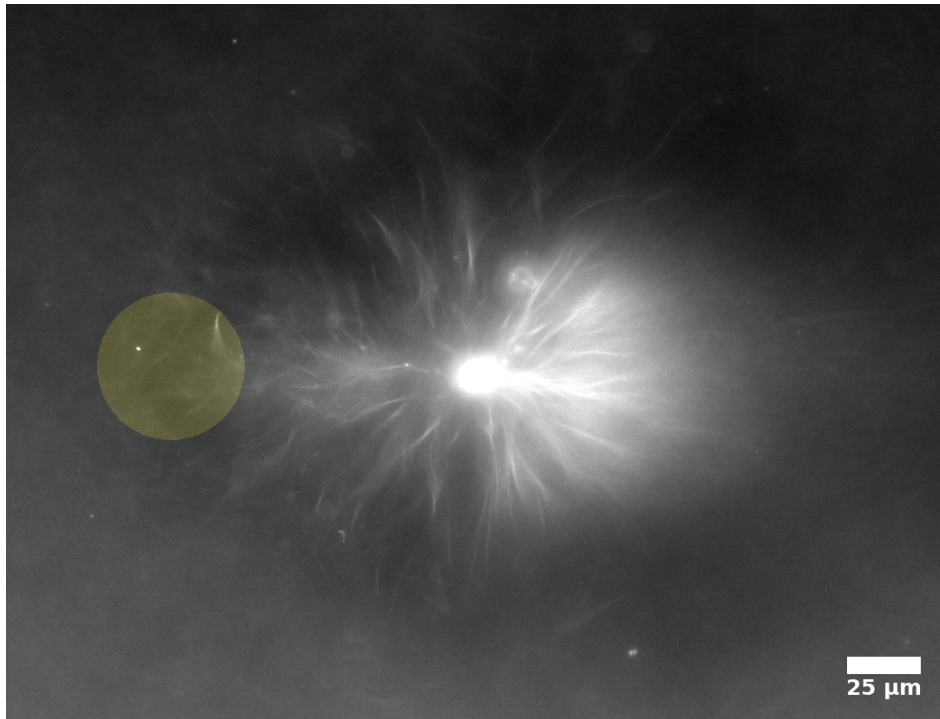


Figure 2.15: Aster following a 50 μm disk moving at 200 nm/s from right to left. Image is integrated across z .

sphere. Given the unlikelihood of optical tweezing being related to the potential we observe, we attribute the effective potential other effects such as the remodeling of the microtubule field.

2.4.15 Mechanism and Stability of a Moving Aster

While the molecular details of aster movement remains a topic of future study, there are mesoscopic phenomena that we observe. When the light pattern activates a region adjacent to the aster, microtubule bundles form. As the light pattern moves, a stream of bundles spans from the light pattern towards the aster. This behavior can be most clearly seen at the highest stage speeds of 200 nm/s and with larger disk sizes (Fig. 2.15).

The stream of bundles appears to pull against the arms of the aster towards a new contractile center.

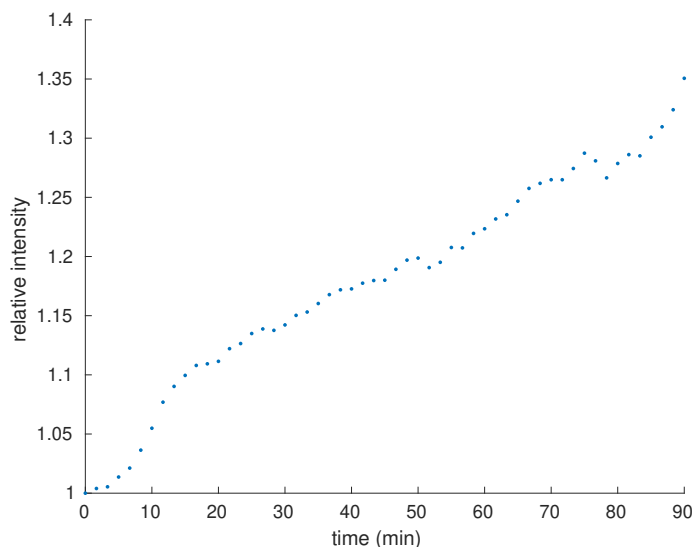


Figure 2.16: Intensity of an aster for a light pattern moving at 200 nm/s. The y-value is normalized to the intensity at $t = 0$. Intensity is measured for an ROI with a fixed diameter and tracks with the aster center.

During aster movement, we observe a cloud of unbundled microtubules are left in the wake of a moving aster, indicating that there is a decay process occurring. At the same time, however, we also observe that microtubules are incorporated into the aster, as demonstrated by the increase in the aster intensity over time (Fig. 2.16), which starts to occur after a few minutes. The increase in intensity also indicates that the incorporation rate is greater than the aster decay rate. We speculate that the newly added microtubules deliver linked motor proteins that maintain some of the bonds between filaments, allowing the aster to persist outside of the light pattern.

2.4.16 Single Motor Velocity Determination from Gliding Assay

Gliding assay images were acquired every second with total internal reflection fluorescence (TIRF) microscopy. Motor speeds were determined by tracking individual microtubules. Single microtubules were identified by edge detection followed by size thresholding to remove small particles on the glass and large objects that are overlaying microtubules. The centroid of each object is identified

and paired with the nearest-neighbor in the next frame. The Euclidian distance between the paired centroids is calculated and used to determine the microtubule velocity. The mean motor speed was determined from the mean frame-by-frame velocities (excluding those less than 75 nm/s, which is our typical sample drift).

2.4.17 Minimum Size Limits of Structures

Here we explore the minimum feature sizes that we can generate. To test the limits for flow generation, we vary the length and height of the excitation bar. We observe that the minimum excitation bar length that is able to generate flows is between 87.5-175 μm Fig. 2.18, which corresponds to a microtubule network of $\approx 100 \times 30 \mu\text{m}$. We note that this length is similar to the bundle buckling length observed in Fig. 4b. We speculate that the limits of the minimum length pattern for generating flow may be related to this buckling length scale.

In addition, we find that the minimum height of an excitation bar that can generate flow is $\approx 2\mu\text{m}$ Fig. 2.19. We observe that the network that forms is $\approx 300 \times 20 \mu\text{m}$. Below this excitation limit we observe the formation of unstable microtubule bundles that do not persist long enough to form a more ordered structure. While the excitation bar extends 350 μm , we speculate that below the minimum height, the density of active motors is too low to completely drive organization. This may be a result of the diffusivity and speed of the motor proteins.

We determine the angle resolution by taking two overlapping bars, as in the “+” shape shown in Fig. 4f, and rotating them relative to each other. When the bars are orthogonal to each other, there are four distinct inflows at the corners. We decrease the angle between the bars until the flow pattern appears to be that of a single bar (two inflows). The minimum angle between two bar patterns for which there remain 4 distinct inflows and outflows is between $\frac{\pi}{16} - \frac{\pi}{8}$ (Fig. 2.20). The angle that sets this limit may in part be determined by the average length of the filament bundles that form orthogonal to the major axis of each bar pattern, which are $\approx 20 \mu\text{m}$ in length. For a sufficiently shallow angle, these orthogonal bundles may interact with each other and cause the two microtubule networks to

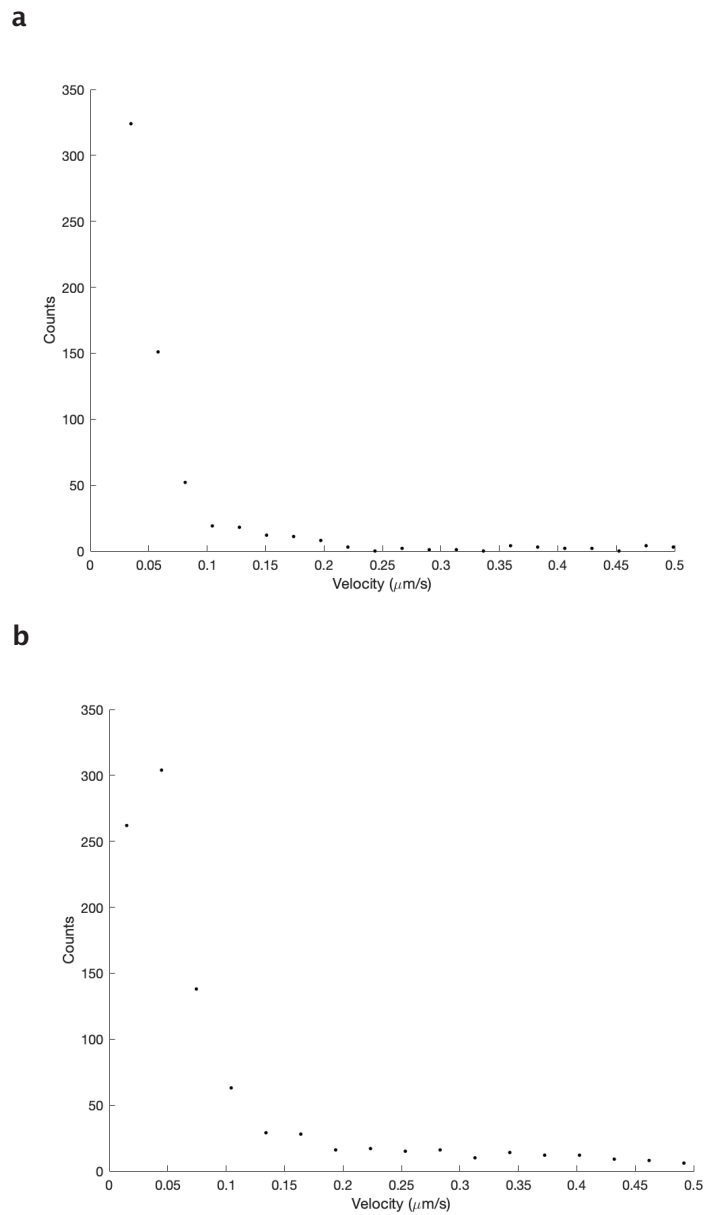


Figure 2.17: Velocity distribution of gliding microtubules. **a**, Binned velocities for K401-iLID motors, the mean of the data is 230 nm/s with a standard deviation of 200 nm/s. **b**, Binned velocities for K401-micro motors, the mean of the data is 300 nm/s with a standard deviation of 250 nm/s.

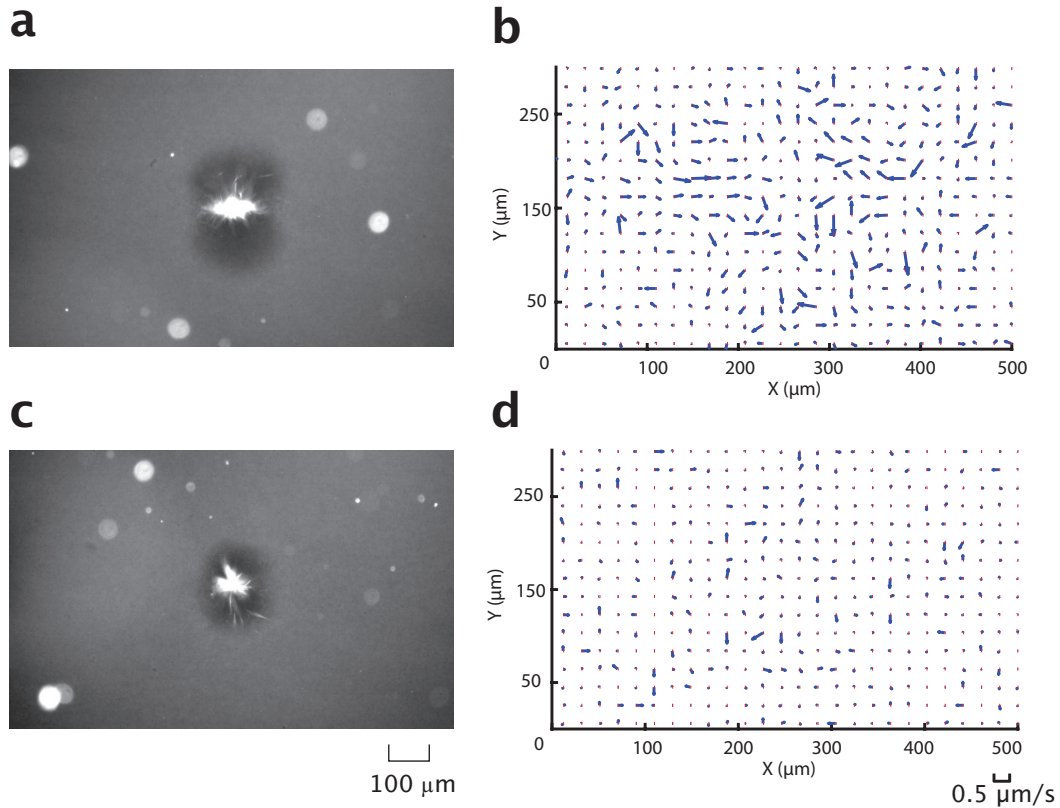


Figure 2.18: Minimum length experiment for a $L \times 20 \mu\text{m}$ excitation pattern. **a**, Fluorescent microtubule channel for $L = 175 \mu\text{m}$. **b**, Corresponding flow field to **(a)**. **c**, Fluorescent microtubule channel for $L = 87.5 \mu\text{m}$. **d**, Corresponding flow field to **(c)**.

be pulled into each other, merging into a single linear structure. The flow pattern and microtubule distribution of Fig. 2.20c and **d** closely resemble those produced by a single rectangular bar of light.

We find that the minimum disk diameter to form an aster is between $6.25\text{--}12.5 \mu\text{m}$ Fig. 2.21. The arms of the smallest aster we are able to form appear to be $\approx 20 \mu\text{m}$. We note that below this limit, small microtubule bundles form transiently and remain disordered. Due to the similarity of the minimum excitation length scale to the average microtubule length, we hypothesize that the smallest aster we can form may in part be determined by the microtubule length distribution.

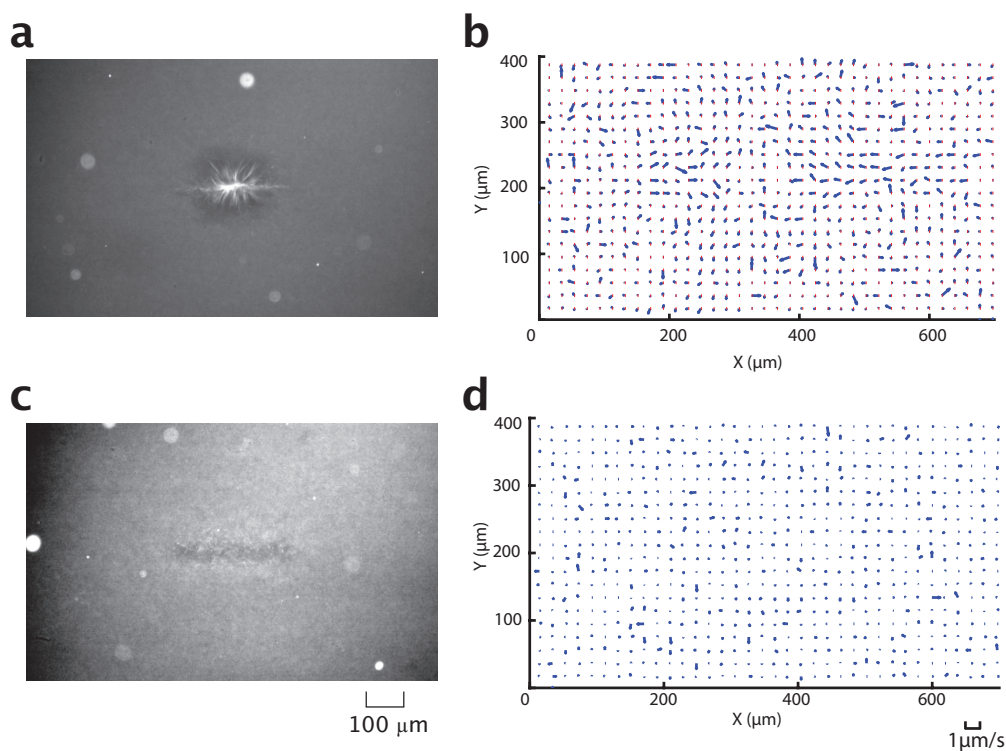


Figure 2.19: Minimum height experiment for a $350 \times H \mu\text{m}$ excitation pattern. **a**, Fluorescent microtubule channel for $H = 2 \mu\text{m}$. **b**, Corresponding flow field to **(a)**. **c**, Fluorescent microtubule channel for $H = 1 \mu\text{m}$. **d**, Corresponding flow field to **(c)**.

2.4.18 Fluid Flow Patterns from Particle Tracking

The fluid flow generated by the movement of microtubule filaments is measured using Particle Tracking Velocimetry (PTV) [74] of fiducial tracer particles. Inert $1 \mu\text{m}$ diameter microspheres (SI 2.3.6) are added to the reaction buffer and imaged with brightfield microscopy. The images are pre-processed using a Gaussian filter with a standard deviation of 1 pixel, followed by thresholding to eliminate the background noise. After filtering, the centroid of each particle is measured and tracked.

A nearest-neighbor algorithm [75] is applied to find particle pairs within a square search window (30 pixels). Displacement vectors are then calculated by comparing the position of particle pairs in consecutive frames. The same process

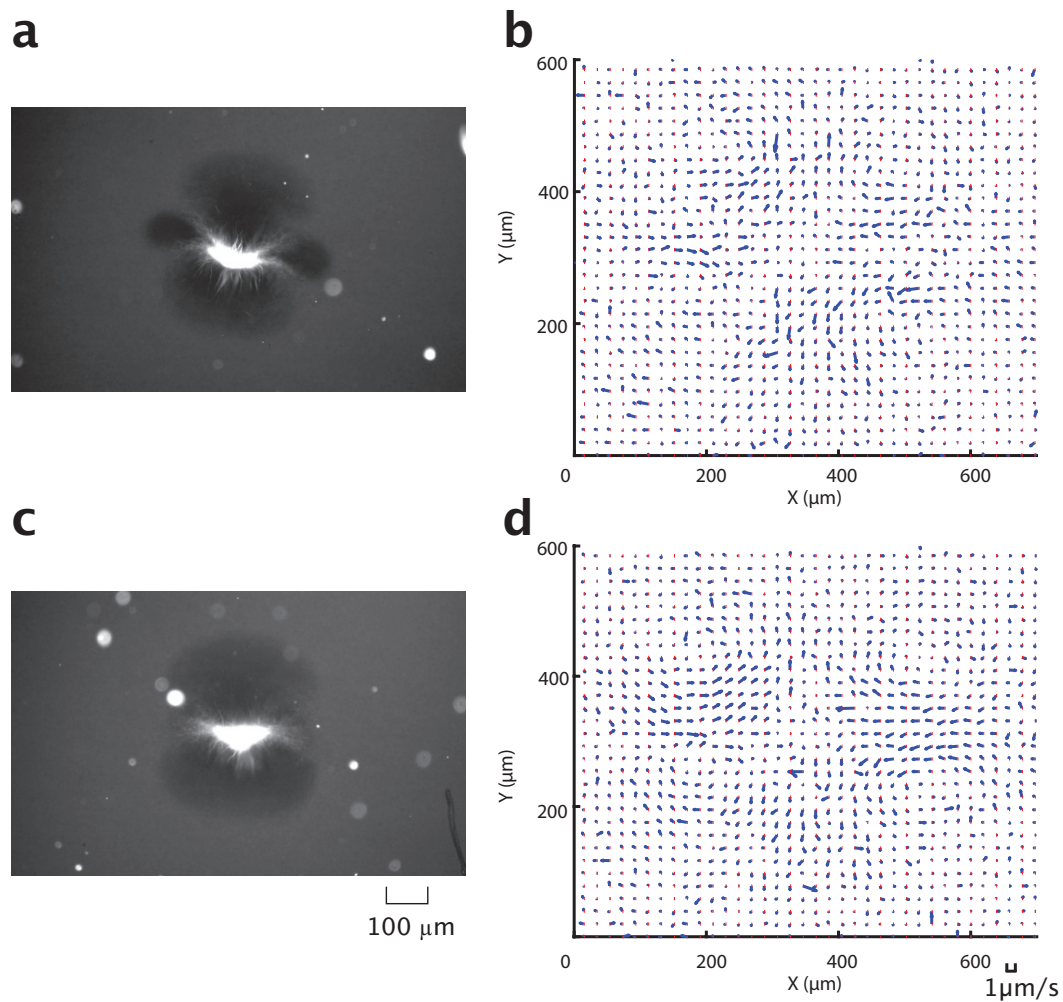


Figure 2.20: Minimum angle experiment for two $350 \times 20 \mu\text{m}$ excitation pattern. **a**, Fluorescent microtubule channel for an angle of $\frac{\pi}{8}$. **b**, Corresponding flow field to **(a)**. **c**, Fluorescent microtubule channel for an angle of $\frac{\pi}{16}$. **d**, Corresponding flow field to **(c)**.

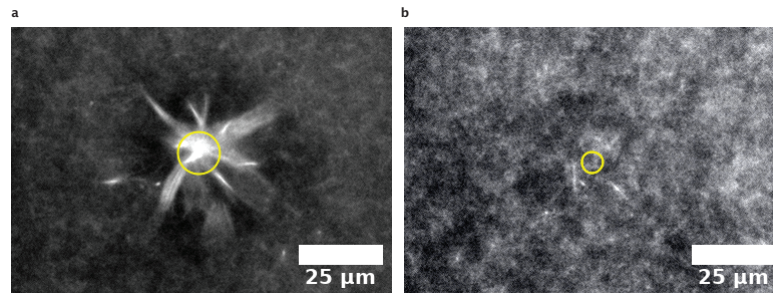


Figure 2.21: Minimum aster size experiment for disk patterns. **a**, Fluorescent microtubule channel for an excitation disk $12.5\ \mu\text{m}$ excitation disk. **b**, Fluorescent microtubule channel for a $6.25\ \mu\text{m}$ excitation disk. The yellow circle represents the perimeter of the excitation disk.

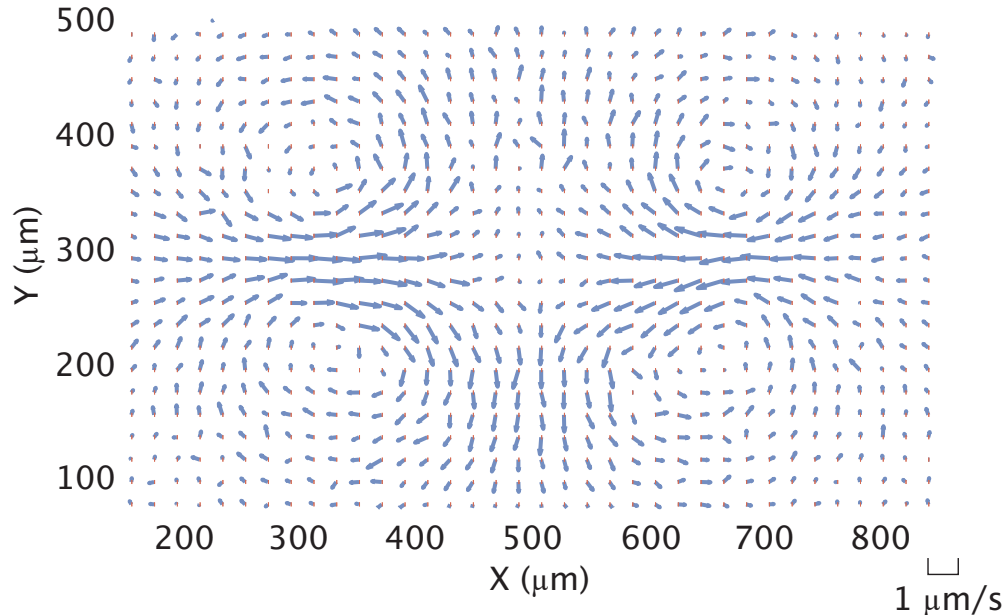


Figure 2.22: Flow velocity field generated with a $350\ \mu\text{m}$ activation bar measured with PTV of tracer particles. Vector data is used to calculate streamline plot in Fig. 4c.

is repeated for the entire image sequence (30 min). The velocity field is generated by dividing the displacement vector field by the time interval between frames. The averaged velocity field shown in Fig. 2.22 is carried out by grouping and averaging all velocity vectors within a $30\ \text{pixel} \times 30\ \text{pixel}$ window.

2.4.19 2D Flow Field

We measure the flow field at different focal planes to determine its z-dependence. The flow fields are generated from PTV, as previously described (SI 2.4.18). We image a z-stack of 3 planes separated by 20 μm , where the sample typically extends $\approx 70\mu\text{m}$ in the z-direction. Following the same particle tracking algorithm, we retrieve the flow fields (Fig. 2.23) averaged over a 20 min time window. We do not observe significant differences in the flow field's structure or speed at the various z-positions. Therefore, for all subsequent flow measurements we image a single focal plane. Further, when we model the flow field (SI 2.4.23), we assume that it is a 2D pattern.

2.4.20 Time Stability of Flow Patterns

In order to understand how the flow field changes in time, we divide the 30 minute experiment into four 7.5 minute time windows and calculate the flow field for each window. The resulting velocity fields are shown in (Fig. 2.24). We note that the structure of the flow field remains similar throughout the experiment. In addition, the maximum speed of the velocity field is constant over time (Fig. 2.25), which further confirms that the fluid flow is stable over the experiment.

2.4.21 Generation of Streamline Plots

Streamlines are the spatial path traced out by fiducial points moving with the fluid flow. They can be numerically generated from a velocity vector field. To generate the streamlines shown in Fig. 4c, g, we use the streamplot function found in the Matplotlib Python library. First, the streamplot function maps a user-defined grid onto the velocity vector field, which determines the density of the streamlines. Next, streamplot creates trajectories from a subset of velocity vectors by performing an interpolation from the current position $x(t)$ of the streamline to the next position $x(t+dt)$ based on the velocity $v(x(t))$ by a 2nd-order Runge-Kutta algorithm. To prevent streamlines from crossing, a mask is defined around each interpolated trajectory, which excludes other trajectories from entering into the mask.

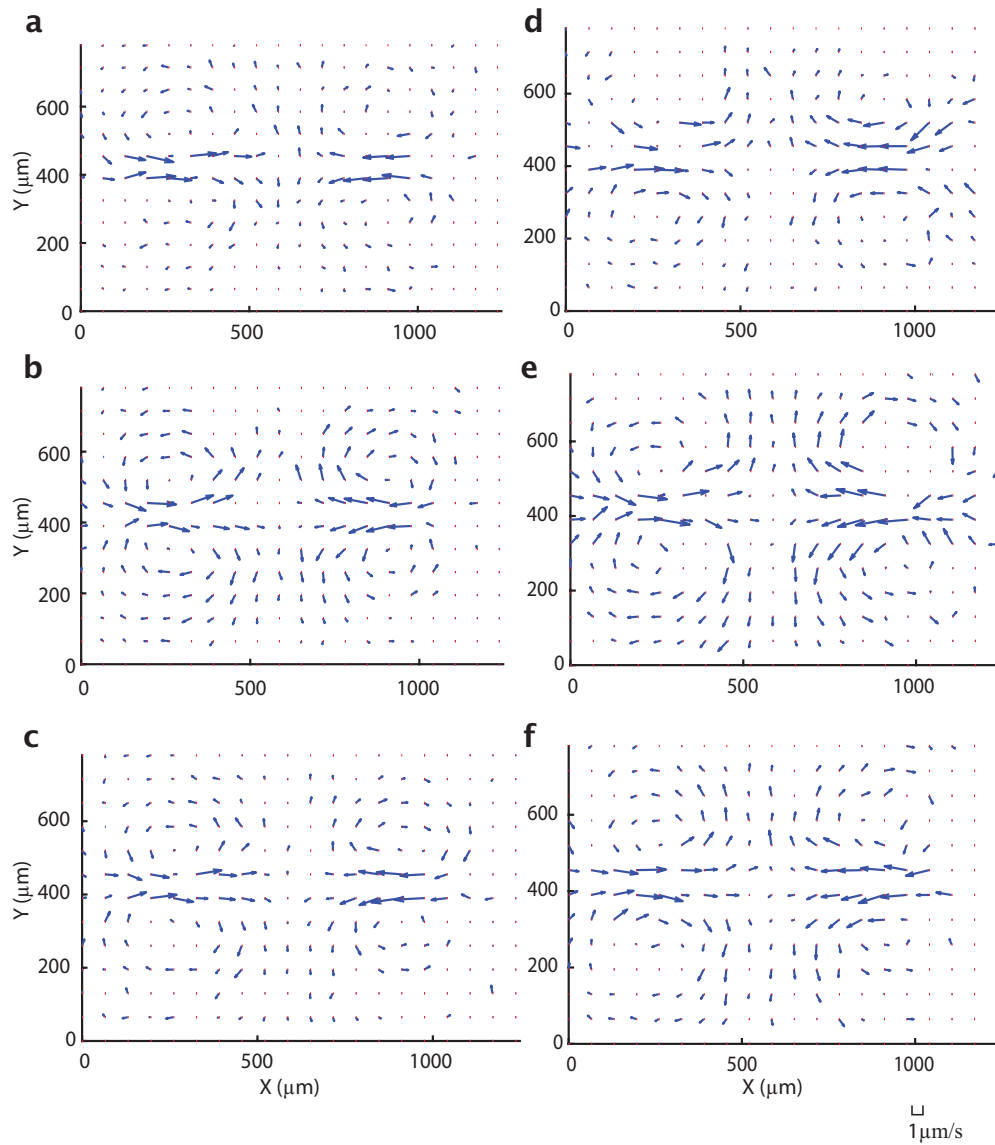


Figure 2.23: A flow field measured at three different z-positions separated by $20 \mu\text{m}$. The field is generated with a $700 \mu\text{m}$ activation bar. **a**, Highest z-position, **b**, middle z-position, **c**, lowest z-position. **d**, **e**, **f**, are from another experiment following the same order.

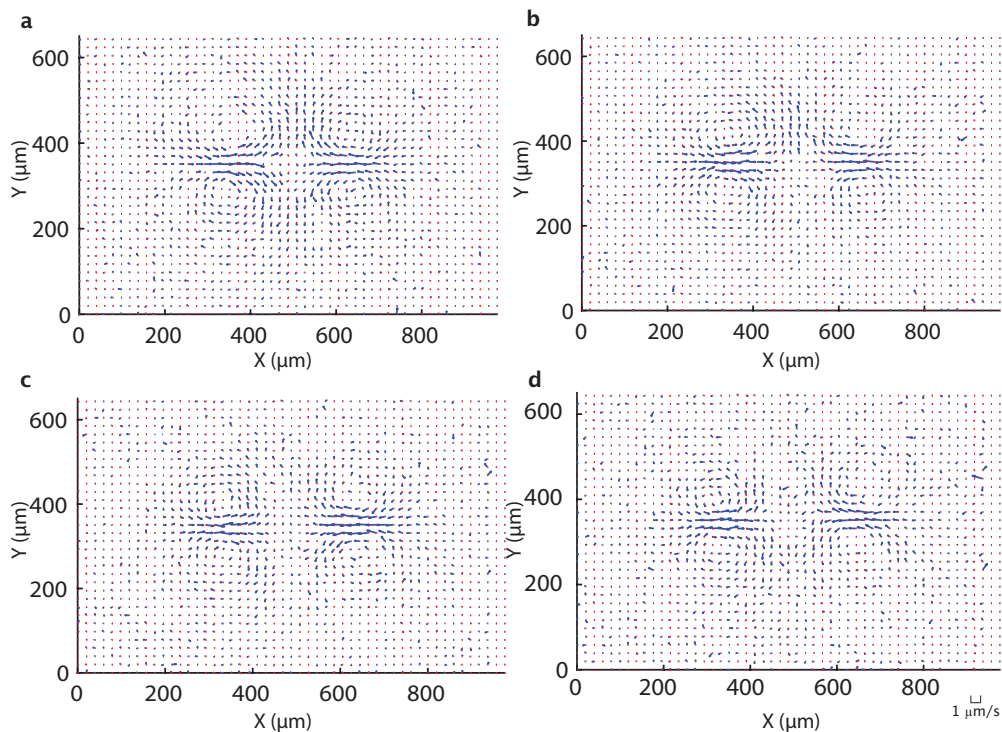


Figure 2.24: Velocity field averaged over 7.5 minute intervals in a single experiment. Time windows are **a**, $t = 0 - 7.5$ min **b**, $t = 7.5 - 15$ min **c**, $t = 15 - 22.5$ min **d**, $t = 22.5 - 30$ min

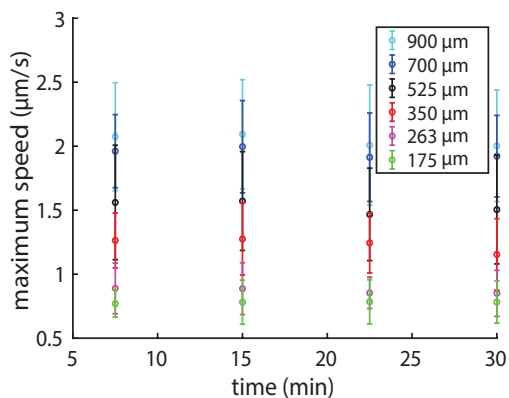


Figure 2.25: The average maximum speed for four different 7.5 minute time windows. The data points represent the average of nine experiments. The error bars are the associated standard deviation.

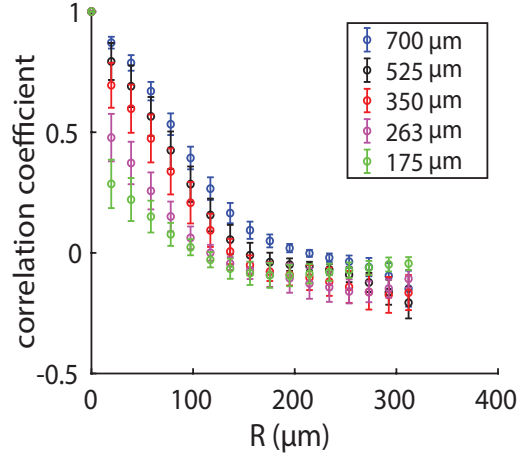


Figure 2.26: The correlation coefficient as a function of distance. Each marker shows the mean over nine individual experiments, and error bars are the associated standard deviation.

2.4.22 Correlation Length

The flow patterns that we observe have vortices. We can characterize the spatial extent of patterns like vortices by the velocity–velocity correlation coefficient $C(R)$ [76, 41]:

$$C(R) = \frac{\langle V(R) \cdot V(0) \rangle}{\langle |V(0)|^2 \rangle} \quad (2.23)$$

where V is the fluid velocity vector, R is the distance between velocity vectors, $\langle \rangle$ denotes assemble average and $||$ is the magnitude of the vector. The correlation length L_c is defined as the distance when $C(L_c) = 0$. This is the length scale where velocities vectors change to an orthogonal direction. By definition, $C(0) = 1$. The correlation coefficient as a function of R is calculated to determine L_c for each bar length (Fig. 2.26).

2.4.23 Theoretical Model of the Fluid Flow Field

We use solutions of the Stokes equation, the governing equation for fluid flow at low-Reynolds-number [77], to model our induced flow fields. One of the simplest solutions of the equation is the Stokeslet, which describes the flow field induced by a point force [78]. Here, we attribute the flow-generating point forces

to contracting microtubule bundles. Since the microtubules at the center of the activation bar appear to contract much more slowly than in other regions of the light pattern, we do not model Stokeslets in the central 120 μm of the activation bar. We superimpose the solutions for two series of Stokeslets, one for each side of the bar. Each series of Stokeslets is composed of 7 point forces with identical magnitude ($|\mathbf{f}| = 2 \text{ nN}$), separated by 20 μm (Fig. 2.27) to model the 350 μm activation bar case.

The velocity field $\mathbf{u}(\mathbf{x})$ generated by a point force \mathbf{f} located at \mathbf{x}' in a 2D plane is given as

$$\mathbf{u}(\mathbf{x}) = \frac{1}{4\pi\eta} \left(-\mathbf{f} \log(r) + \frac{(\mathbf{f} \cdot (\mathbf{x} - \mathbf{x}'))(\mathbf{x} - \mathbf{x}')}{r^2} \right) \quad (2.24)$$

where η is the fluid viscosity and r is the absolute distance, defined as

$$r = |\mathbf{x} - \mathbf{x}'|. \quad (2.25)$$

We estimate that $\eta = 2 \times 10^{-3} \text{ Pa} \cdot \text{s}$ (Supplementary Information 2.4.24).

Comparing Fig. 2.27 to Fig. 2.22, for the rectangular bar experiment, we see our model recovers the general pattern of inflows and outflows in magnitude and direction. In both figures, the inflows along the X direction and the outflows along the Y direction are asymmetric in magnitude, with the inflows being greater than the outflows. However, in the experiments there can be additional asymmetries not captured by the model. For example in Fig. 2.22, outflows in the downward direction (Y-axis, $Y < 300 \mu\text{m}$) appear greater in magnitude than the outflows in the upward direction (Y-axis, $Y > 300 \mu\text{m}$). This may be related to the microtubule buckling shown in Fig. 4b, which leads to asymmetry of the microtubule network density in the last panel of Fig. 4a. Further, we note that we do not observe vortices for our model parameters. It is possible that the presence of vortices may lead to additional effects not generated by the current model.

There are various candidate mechanisms for vortex generation - boundary conditions, zones of depleted microtubules, and non-Newtonian fluid properties, to list a few. Further investigation will be needed to determine which of these effects, if any, cause the observed vortices.

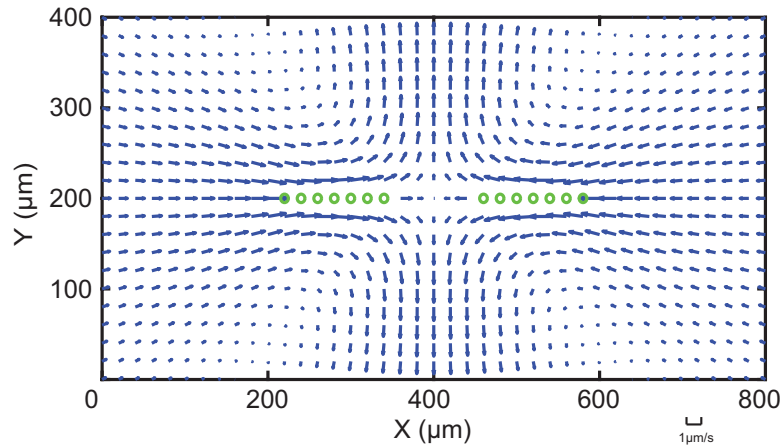


Figure 2.27: Flow field generated by 14 Stokeslets, indicated by green circles, to model the 350 μm activation bar case. This theoretical model recovers the general pattern of inflows and outflows observed in the experiment (Fig. 4a), but not the vortices and asymmetries in flow magnitudes.

Due to the linear nature of low-Reynolds-number flow [79], we expect that the velocity field generated by a complex light pattern can be retrieved by superposition of simple patterns. To confirm this, we superimpose flow fields from single bars to mimic the flow field generated by “L”, “+” and “T”-shaped light patterns (Fig. 2.28). For the “+” case, the superimposed fields closely resembles the experimentally observed field (Fig. 2.28c). The “L” and “T”-shaped cases are roughly similar to the experimental results, but direction of the inflows do not match (Fig. 2.28b, d).

To model the “L” and “T” flow fields more accurately, we generate the flow field for a series of Stokeslets following the geometry of the microtubule structure, rather than the light pattern itself. Using this method, the modeled flow fields are a good approximation of the observed flow fields. The inflows and outflows match the experimentally observed positions and orientations (Fig. 2.29). This result implies that the observed flow patterns are set by the microtubule structure rather than the light pattern.

2.4.24 Calculating Fluid Viscosity

To find the viscosity of the background buffer, we used a similar approach to finding the flow fields. We used PTV of fiducial tracer particles (Supplemental

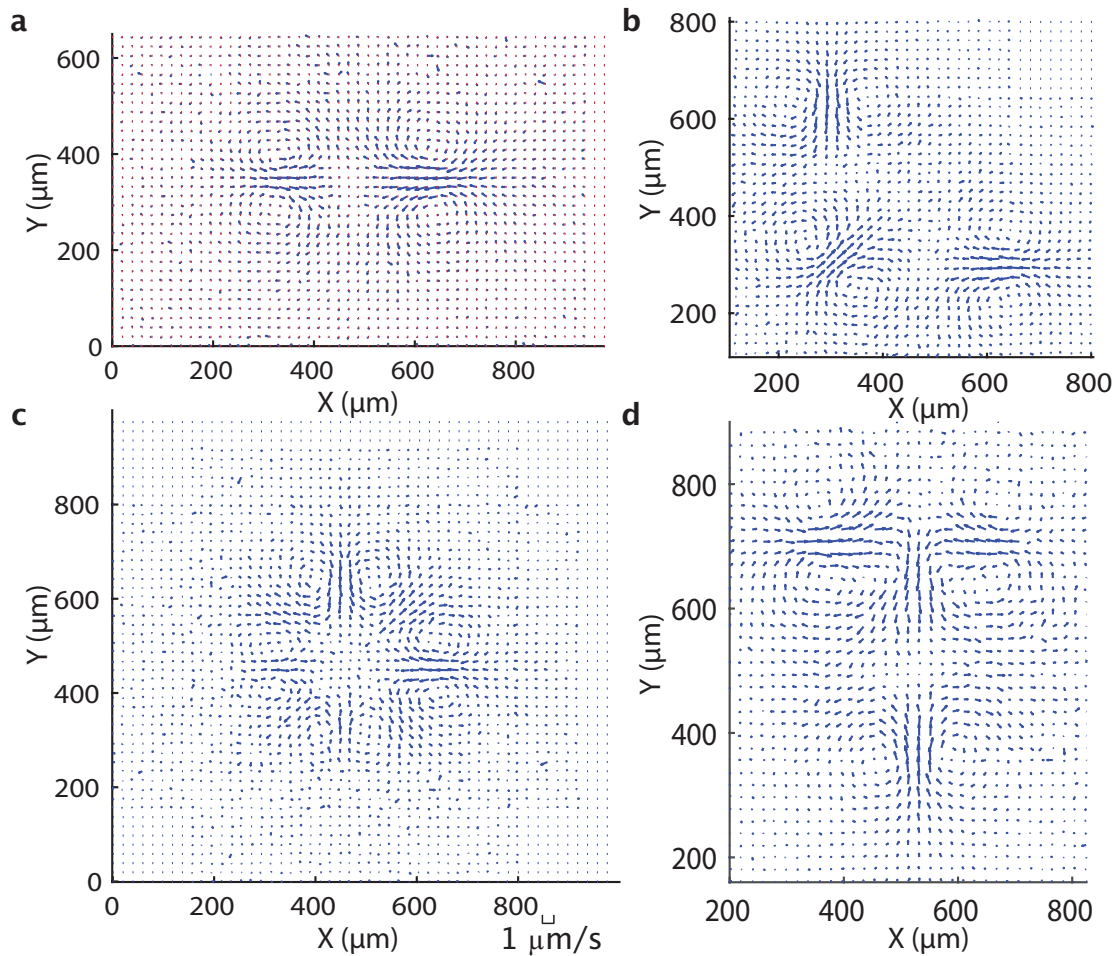


Figure 2.28: Demonstration of the linearity of the flow field. **a**, A time averaged flow field generated by a $350 \mu\text{m}$ rectangular bar. Flow fields generated by the rotation and superposition of the flow field in **(a)** to retrieve flow fields for **b**, “L” **c**, “+”, and **d**, “T”-shaped light patterns.

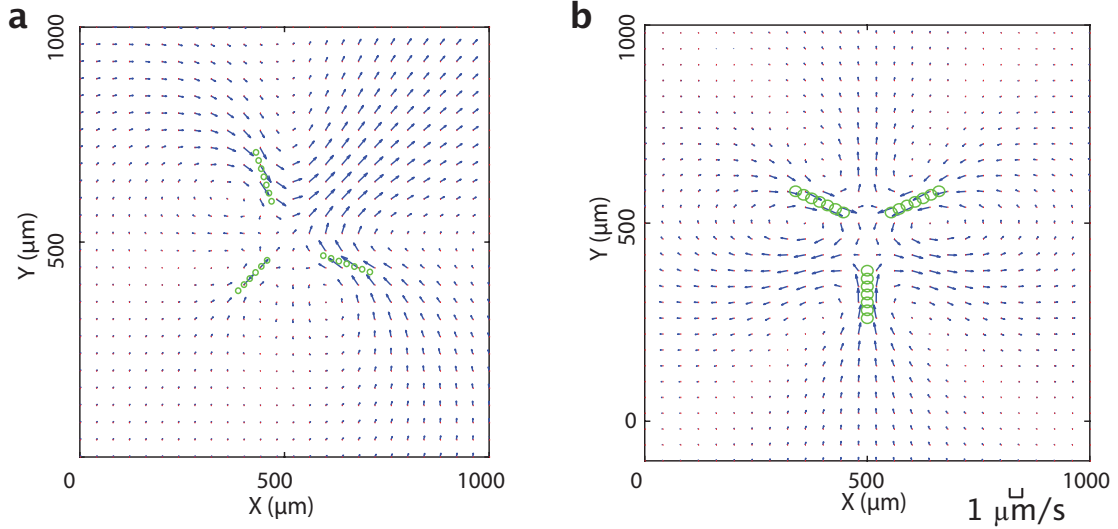


Figure 2.29: Theoretical simulation of fluid flows under complex light patterns using Stokeslets. The Stokeslets are positioned following the shape of the microtubule network observed in Fig. 4f. Green circles denote the Stokeslets. **a**, Flow field for “L”-shaped light pattern. **b**, Flow field for “T”-shaped light pattern.

Information 2.4.18) in inactivated regions of the sample of the 175 μm activation bar experiment. Assuming the buffer is Newtonian [80], the inert tracer particles diffuse freely due to thermal fluctuations. From the tracking results, we measure the mean-squared displacement $\text{MSD}(t)$ of the particles:

$$\text{MSD}(t) = \langle (x(t) - x(0))^2 + (y(t) - y(0))^2 \rangle, \quad (2.26)$$

where $x(t)$ and $y(t)$ are the position of a given particle at time t and $\langle \rangle$ denotes ensemble average. For this calculation, each frame is $t = 4$ s apart. The $\text{MSD}(t)$ of a freely diffused particle in 2D follows the Stokes-Einstein equation

$$\text{MSD}(t) = 4Dt = \frac{2k_B T}{3\pi\eta r} t, \quad (2.27)$$

where $r = 0.5 \mu\text{m}$ is the radius of the particle. Then, the viscosity of the buffer solution is estimated as

$$\eta = \frac{8k_B T}{3\pi r \text{MSD}(t)}. \quad (2.28)$$

The same process is repeated through nine individual experiments and the average estimated viscosity η is $2 \times 10^{-3} \text{ Pa} \cdot \text{s}$.

2.4.25 Comparison to Optically Controlled Bacteria

The polarity of the motors and microtubules makes them distinct from systems based on optically controlled bacteria [81, 82]. In our work, the localization of motor linkages causes microtubules to collectively reorganize into contracting networks. Due to the organization of microtubules and resulting dipolar stresses on the surrounding medium, we are able to create coherent flows. In contrast, localization of the activity of bacterial swimmers results in a change in the bacterial density, but lacks structural order and therefore does not generate coherent flows. However, bacterial densities can form arbitrary patterns that directly correspond to the optical projections analogous to photolithography. The resolution of the patterns we can create (Supplementary Information 2.4.17) is generally lower than the reported $\approx 2 \mu\text{m}$ resolution achievable with bacterial swimmers. Light in our system does not directly pattern microtubules, but rather defines an effective reaction volume where certain reorganizing motifs can occur.

Chapter 3

Motor Processivity and Speed Determine Structure and Dynamics of Microtubule-Motor Assemblies

This chapter is derived from the preprint: Banks, R.A., *et al.* Motor processivity and speed determine structure and dynamics of microtubule-motor assemblies. *BioRxiv* (2021). DOI: 10.1101/2021.10.22.465381

3.1 Abstract

Active matter systems can generate highly ordered structures, avoiding equilibrium through the consumption of energy by individual constituents. How the microscopic parameters that characterize the active agents are translated to the observed mesoscopic properties of the assembly has remained an open question. These active systems are prevalent in living matter; for example, in cells, the cytoskeleton is organized into structures such as the mitotic spindle through the coordinated activity of many motor proteins walking along microtubules. Here, we investigate how the microscopic motor-microtubule interactions affect the coherent structures formed in a reconstituted motor-microtubule system. We explore key parameters experimentally and theoretically, using a variety of motors with different speeds, processivities, and directionalities. We demonstrate that aster size depends

on the motor used to create the aster, and develop a model for the distribution of motors and microtubules in steady-state asters that depends on parameters related to motor speed and processivity. Further, we show that network contraction rates scale linearly with the single-motor speed in quasi one-dimensional contraction experiments. Finally, we demonstrate that competition between motors of opposite polarity can be tuned to create various structures. In all, this theoretical and experimental work helps elucidate how microscopic motor properties are translated to the much larger scale of collective motor-microtubule assemblies.

3.2 Introduction

A signature feature of living organisms is their ability to create beautiful, complex patterns of activity, as exemplified in settings as diverse as the famed flocks of starlings in Rome or the symmetrical and dazzling microtubule arrays that separate chromosomes in dividing cells [83]. While such organization in nature has long captured the attention of artists and scientists alike, many questions remain about how the patterns and structures created by living organisms arise. In active systems such as bird flocks or microtubule-motor arrays, energy is consumed at the local level of the individual actors, and the coordinated action of many individuals creates the much larger observed patterns. How the specific microscopic activity of each individual leads to the final large-scale assembly formed remains an open question.

The motor-microtubule system is an excellent system in which to test this question, as many motor proteins with a variety of properties, such as speeds, stall and detachment forces, processivities, and directionalities exist in nature. These motors play a variety of roles in cells; some transport cargo while others localize to distinct regions of the mitotic spindle [84, 85, 86, 87, 88]. Studies have investigated how the microscopic properties of these motors makes them uniquely suited to their cellular role. For example, kinesin-1's high speed and processivity make it excellent at transporting cargo [89, 90]. However, in *in vitro* systems, kinesin-1 tetramers are able to form asters, extensile networks, and contractile networks [19, 41, 17, 23]. Ncd (Kinesin-14) and Kif11 (Kinesin-5) have similarly been shown

to form asters *in vitro*, yet it remains unclear how the properties of these motors affect the structure and dynamics of the assemblies created [19, 91].

In this work, we create motor-microtubule structures with a variety of motors and develop theoretical models to connect the motor properties to the organization and dynamics of the assemblies. Our recently developed optogenetic *in vitro* motor-microtubule system demonstrated the formation of asters and other contractile networks with kinesin-1 (K401) upon light activation [23]. We now show how this system can be made to work with kinesin-5 (Kif11) and kinesin-14 (Ncd), and form asters of varying sizes with each motor. Further, we develop a theoretical model connecting the distribution of motors and microtubules in asters; calculated distributions depend on the motor properties and fit with our experimental data. By using motors with different speeds, we find that contraction rates in quasi one-dimensional microtubule networks directly depend on the single-motor velocity. Finally, we set up competition between motors of opposite polarity, and find that the final structure formed depends on the properties and concentrations of each motor. This theoretical and experimental work sheds light on the ways that microscopic motor properties are reflected in the thousand-fold larger length scale of motor-microtubule assemblies.

3.3 Results

3.3.1 Aster Size Depends on Motor Used

We build on the foundational work that demonstrated the ability to control motor-microtubule systems with light [23] to consider a new set of motors with different fundamental properties. In brief, kinesin motors are fused to the light-dimerizable pair iLid and micro. In the absence of light, motor dimers walk along microtubules but do not organize them; upon activation with light, the motor dimers couple together to form tetramers, crosslinking the microtubules they are walking along as shown in Fig. 3.1(A). Projecting a disk pattern of light on the sample results in the formation of an aster.

In this work, we aim to determine how the properties of the motor affect the resulting structures. While experiments with this system were previously

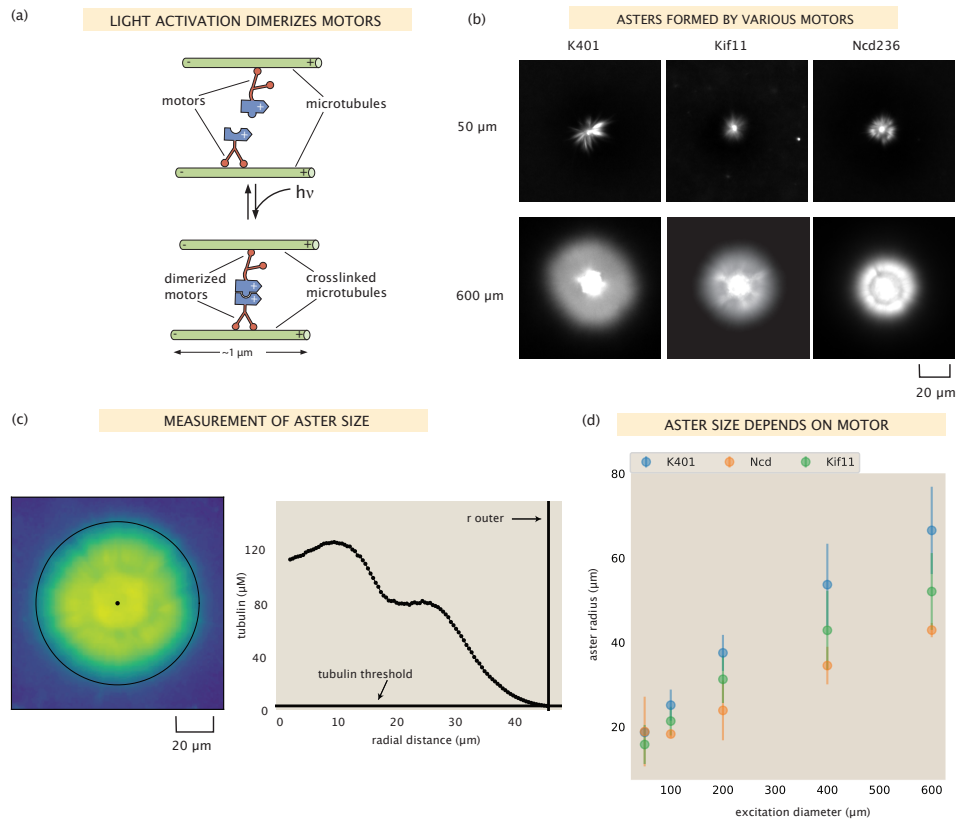


Figure 3.1: Aster size depends on motor used. (A) Motor heads are fused to optogenetic proteins such that activation with light causes the formation of motor tetramers (dimer of dimers). Motors are shown walking towards the microtubule plus-end. K401 and Kif11 walk in this direction, however Ncd is minus-end directed. (B) Images of the microtubule fluorescence for asters formed with each of the motors excited with a disk either 50 μm or 600 μm in diameter. (C) Image of the microtubule fluorescence from an aster with the measured size represented with the outer black circle. The plot on the right shows the radial microtubule concentration as revealed by fluorescence intensity; the threshold concentration used to determine the aster size is shown as a black horizontal line. (D) Mean aster size ($n \approx 5$ asters for each condition) for the three motors and different excitation diameters; the error bars represent the standard deviation.

Table 3.1: Motor proteins used and their properties

Motor	Speed	Processivity	Direction
K401	≈ 600 nm/s	≈ 100 steps [92]	plus [90]
Ncd236	≈ 115 nm/s	not processive [93, 94, 95]	minus [96]
Kif11(513)	≈ 70 nm/s	≈ 10 steps [97]	plus [97]

performed with *D. melanogaster* kinesin-1 motors (K401) [23], in the present work, we investigate if other kinesin motor species with different intrinsic properties such as speed and processivity would lead to light-induced microtubule organization. Towards this end, we use the same light-dimerizable scheme to form microtubule structures with two other motors: Ncd (*D. melanogaster* kinesin-14) and Kif11 (*H. sapiens* kinesin-5). The single-molecule properties of all three motors we use are summarized in Table 1. We measure the speed of each motor species by gliding assays (SI section 3.5.9); the processivities are based on literature values. Further, we fluorescently label the motors using mVenus or mCherry to visualize the motors and microtubules in separate imaging channels within the same assay (SI section 1).

As seen in Fig. 3.1 (B) and Fig. 3.8, each of these motors is able to form asters of varying sizes in our system. It was previously unclear whether there were limits to speed, processivity, or stall force that might prevent any of these motors from forming asters in our light-controlled system, although Ncd has previously been shown to form asters as constitutive oligomers [49, 98]. We found that all were able to form asters upon illumination by various excitation diameters ranging from 50 to 600 μm . In order to measure the size of the asters, we used the distribution of fluorescently labeled microtubules, which peaks in the center of the aster and generally decreases moving outward. We defined the outer radius of the aster as the radius at which the microtubule fluorescence is twice the background microtubule concentration, see Fig. 3.1(C) for an example aster outer radius determination. We found this method to agree well with a visual inspection of the asters (Fig. 3.8).

We find that aster size increases with excitation diameter, shown in Fig. 3.1(B) and (D), consistent with what was shown by Ross et al. for K401 [23]. Interestingly, we find that the size of the asters also depends on the motor used. For each

excitation diameter, except for the 50 μm case, K401 formed the largest asters and Ncd formed the smallest, with Kif11 producing asters of intermediate size (Fig. 3.1(D)). What is it about the different motors that confers these different structural outcomes? We found that this trend correlates with motor processivity; K401 is the most processive, followed by Kif11, and then Ncd. This is similar to the findings in [19], in which ‘intensity’ of aster formation was related to motor processivity. Other factors could also be contributing to aster size such as the ratio of microtubules to motors as was suggested by [19] but not investigated in the present work, or the motor stall force.

3.3.2 Spatial Distribution of Motors in Asters

The nonuniform distribution of filaments and motors in an aster is a key feature of its organization and has been the subject of previous studies. In these studies, continuum models were developed for motor-filament mixtures which predicted the radial profile of motors in confined two-dimensional systems [44, 20, 21, 99]. A notable example is the power-law decay prediction by Nédélec et al., who obtained it for a prescribed organization of microtubules obeying a $1/r$ decay law [44]. Measuring the motor profiles in asters formed in a quasi-two-dimensional geometry (with the z dimension of the sample being only a few microns deep) and fitting them to a power-law decay, the authors found a reasonable yet noisy match between the predicted and measured trends in the decay exponent.

In our work, we also develop and test a minimal model that predicts the motor profile from the microtubule distribution and the microscopic properties of the motor. In contrast to the earlier study [44], asters formed in our experiments are three-dimensional due to the much larger depth of the flow cells (roughly 100 μm). While the largest asters are likely partially compressed in the z -direction, we assume that this effect does not significantly alter the protein distributions in the central z -slice and hence, for modeling purposes we consider our asters to be radially symmetric, as depicted schematically in Fig. 3.2(A). Our modeling applies to locations outside the central disordered region (called aster ‘core’) with a typical radius of $\approx 15 \mu\text{m}$, beyond which microtubules have a predominantly

polar organization (see SI section 3.5.10 for the discussion of the two aster regions and an example PolScope image that demonstrates their distinction).

Similar to the treatment in earlier works [44, 21, 99], we introduce two states of the motor - an unbound state where the motor can freely diffuse with a diffusion constant D and a bound state where the motor walks towards the aster center with a speed v , illustrated in Fig. 3.2(B). In the steady-state of the system, which we assume our asters have reached at the end of the experiment, microtubules on average have no radial movement and hence, do not contribute to motor speed. We denote the rates of motor binding and unbinding by k_{on} and k_{off} , respectively. When defining the first-order rate of motor binding, namely, $k_{\text{on}}\rho_{\text{MT}}(r)$, we explicitly account for the local microtubule concentration $\rho_{\text{MT}}(r)$ extracted from fluorescence images. This is unlike the previous models which imposed specific functional forms on the microtubule distribution (e.g., a constant value [20, 21], or a power-law decay [44]), rendering them unable to capture the specific features often seen in our measured microtubule profiles, such as the presence of an inflection point, see Fig. 3.1(C) for an example.

The governing equations for the bound (m_{b}) and free (m_{f}) motor concentrations are shown in Fig. 3.2(C). They involve binding and unbinding terms, as well as a separate flux divergence term for each population. Solving them at steady state, we arrive at an equation for the total local concentration of motors defined as $m_{\text{tot}}(r) = m_{\text{b}}(r) + m_{\text{f}}(r)$. The derivation of this result can be found in SI section 3.5.15.1. As seen in the equation for $m_{\text{tot}}(r)$ in Fig. 3.2(C), knowing the microtubule distribution $\rho_{\text{MT}}(r)$ along with two effective microscopic parameters, namely, the effective dissociation constant $K_{\text{d}} = k_{\text{off}}/k_{\text{on}}$ and the length scale $\lambda_0 = D/v$, one can obtain the motor distribution up to a multiplicative constant (C in the equation). Note that in the special case where the motors do not move ($v \rightarrow 0$ or $\lambda_0 \rightarrow \infty$), the exponential term becomes 1 and an equilibrium relation between the motor and microtubule distributions dependent only on K_{d} is recovered.

To test this model, we extract the average radial distributions of microtubule and motor concentrations for each aster. Then, using the microtubule profile as

an input, we fit our model to the motor data and infer the effective parameters K_d and λ_0 (see SI sections 3.5.15.2 and 3.5.15.3 for details). A demonstration of this procedure on an example Kif11 aster is shown in Fig. 3.2(D) where a good fit to the average motor data can be observed. As a validation of our inference method, we additionally extract the radial concentration profiles inside separate wedges of the aster and show that they can be accurately captured by only choosing an appropriate multiplicative constant C for each wedge, while keeping the pair (K_d, λ_0) inferred from average profile fixed (fits to 5 out of 16 different wedge profiles are shown in Fig. 3.2(D) for clarity). The fitting error for other asters is similarly low (Fig. 3.2(E), see Fig. 3.14 for the collection of fitted profiles).

Plotting the inferred parameters K_d and λ_0 from all fits (Fig. 3.2(F)), we find that they are clustered around single values for each motor type and vary between the motors. Based on the single-molecule motor properties in Table 1 and the reported motor binding rates [100], our expectation was that the K_d values for Kif11 and K401 would have a ratio of $\approx 4.6 : 1$ (see SI section 3.5.15.4), while K_d for Ncd would be the highest due to its non-processivity. The ratio of median inferred K_d values for Kif11 and K401 is $\approx 5.6 : 1$ – close to our expectation. However, the inferred K_d values for Ncd are low and comparable to those for K401. One possible resolution of this discrepancy comes from the finding of an *in vitro* study suggesting a substantial increase in the processivity of Ncd motors that act collectively [101]. Specifically, a pair of Ncd motors coupled through a DNA scaffold was shown to have a processivity reaching $1 \mu\text{m}$ (or, ≈ 100 steps) – a value close to that reported for K401 motors. A highly processive movement was similarly observed for clusters of HSET (human kinesin-14) [102] and plant kinesin-14 motors [103]. This collective effect, likely realized for Ncd tetramers clustered on microtubules in highly concentrated aster structures, is therefore a possible cause for the low inferred values of their effective K_d . We also note that while a similar collective effect on processivity was observed for K401 motors [101], it is far less dramatic due to their already large single-motor processivity, and therefore would have a small effect on the effective K_d .

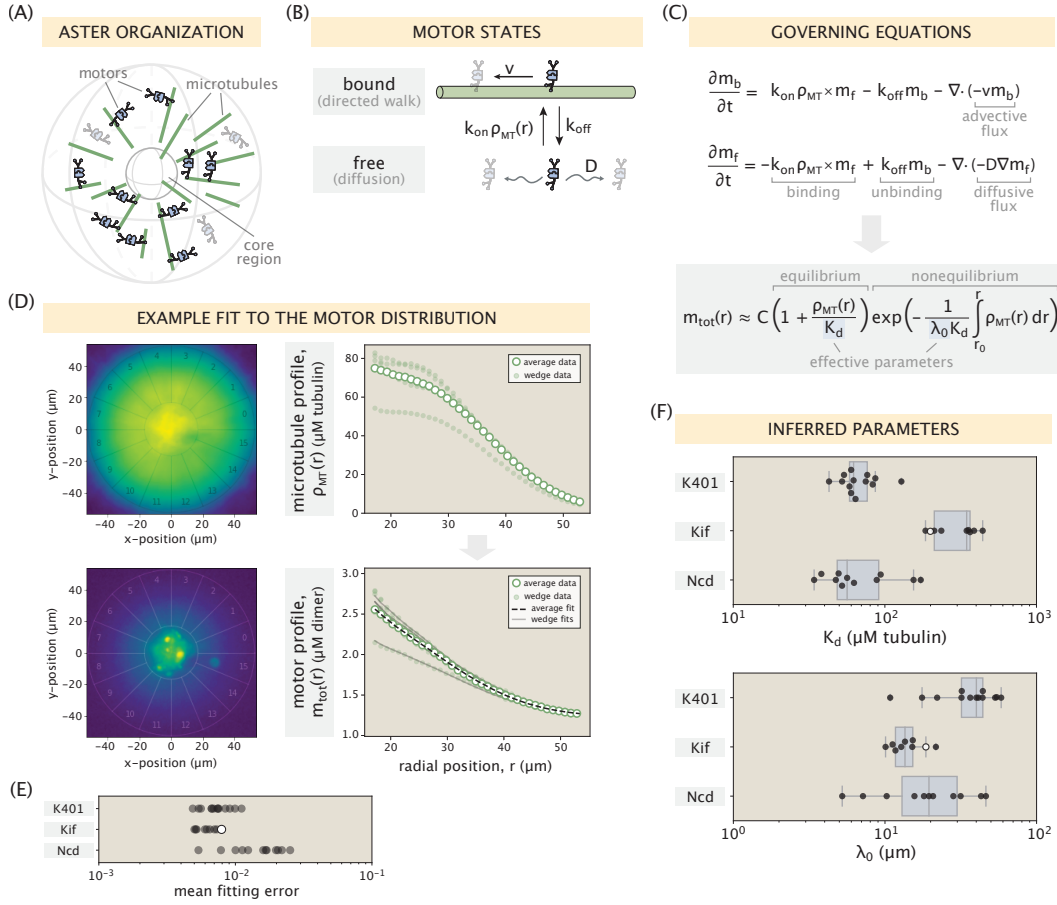


Figure 3.2: Modeling the motor distribution. (A) Schematic of the radial microtubule organization in an aster. Modeling applies to locations outside the disordered core region at the aster center. Components of the schematic are not drawn to scale. (B) Motor states and transitions between them. (C) Governing equations for the bound and free motor populations, along with our solution for the total motor distribution at steady state, expressed via effective parameters $K_d = k_{\text{off}}/k_{\text{on}}$ and $\lambda_0 = D/v$ (see SI section 3.5.15.1 for details). (D) Demonstration of the model fitting procedure on an example Kif11 aster. Fits to the average motor profile as well as to 5 out of 16 wedge profiles are shown. The outlier case with a lower concentration corresponds to wedge 13 in the fluorescence images. (E) Mean fitting errors for all asters calculated from the fits to the wedge profiles. The error is defined as the ratio of the mean residual to the concentration value at the inner boundary. (F) Inferred parameters K_d and λ_0 grouped by the kind of motor. Box plots indicate the quartiles of the inferred parameter sets. The fitting error and the inferred parameters for the Kif11 aster in panel (D) are shown as white dots in panels (E) and (F).

Next, looking at the inference results for the λ_0 parameter in Fig. 3.2(F), we can see that Kif11 and Ncd motors have an average λ_0 value of $\approx 10 - 20 \mu\text{m}$, while the average value for K401 motors is $\approx 40 \mu\text{m}$. From the measured diffusion coefficient of $D \approx 1 \mu\text{m}^2/\text{s}$ for tagged kinesin motors [89] and the single-molecule motor speeds reported in Table 1, our rough estimate for the λ_0 parameter for Kif11 and Ncd motors was $\approx 10 - 15 \mu\text{m}$, and $\approx 2 \mu\text{m}$ for K401. While the inferred values for the two slower motors are well within the order-of-magnitude of our guess, the inferred λ_0 for the faster K401 motor is much higher than what we anticipated. This suggests a significant reduction in the effective speed. One contributor to this reduction is the stalling of motors upon reaching the microtubule ends. Recall that in our model formulation, illustrated in Fig. 3.2(B), we assumed an unobstructed walk for bound motors. Since the median length of microtubules ($\approx 1.6 \mu\text{m}$) is comparable to the processivity of K401 motors ($\approx 1 \mu\text{m}$), stalling events at microtubule ends will be common, leading to a reduction of their effective speed in the bound state by a factor of ≈ 1.5 (see SI section 3.5.15.5 for details). This correction alone, however, is not sufficient to capture the factor of ≈ 25 discrepancy between our inference and the estimate of λ_0 . We hypothesize that an additional contribution may come from the jamming of K401 motors in dense aster regions. This is motivated by the experiments which showed that K401 motors would pause when encountering obstructions during their walk [104, 105]. In contrast, for motors like Ncd and Kif11 which take fewer steps before unbinding and have a larger effective K_d , jamming would have a lesser effect on their effective speed as they would unbind more readily upon encountering an obstacle. Overall, our study shows that the minimal model of motor distributions proposed in Fig. 3.2 is able to capture the distinctions in aster structure through motor-specific effective parameters, although more work needs to be done to explain the emergence of higher-order effects such as motor clustering and jamming, and their contribution to these effective parameters.

Our model also provides insights on the observation that the distribution of microtubules is generally broader than that of the motors. This feature can be observed by comparing the two example profiles in Fig. 3.2(D), and it also holds for the profiles extracted from other asters, shown in Fig. 3.14. In SI section 3.5.15.6,

we use our model to demonstrate this feature in a special analytically tractable case, and discuss its generality across asters in greater detail. We found that the relative width of the motor distribution compared to the microtubule distribution is fairly constant among asters, with their difference being the largest for Ncd motors, consistent with our model predictions. This relationship between the shapes of the distributions may be an important factor in the spatial organization of end-directed motors in the spindle where their localization to the spindle pole is of physiological importance.

3.3.3 Contraction Rate Scales with Motor Speed

Besides the steady-state structure of motor-microtubule assemblies, it is also of great interest to understand their dynamics. Ross et al. demonstrated the formation of quasi one-dimensional contractile networks by creating two asters that are initially separated by a given distance, then activating a thin rectangular region between them to form a network connecting the asters that pulls them together [23]. Example images of microtubule fluorescence during one of these experiments are shown in Fig. 3.3(A). In these experiments, an increase in maximum aster merger speed with greater initial separation was observed. We aimed to confirm this behavior with our various motors and to test the relationship between aster merger speed and single motor speed.

First, we tested the relationship between distance and speed in our experiments. Using optical flow to measure the contraction speed throughout the network, we observe a linear increase in contractile speed with distance from the center of the network, as shown in Fig. 3.3(B). This relationship suggests that the contractile network can be thought of as a series of connected contractile units. Independent contraction of each unit would generate the observed linear increase in speed because more contractile units are added with distance from the center of the network.

Next, we investigated how contractile speeds vary by motor in aster merger experiments. We repeated aster merger experiments with each motor and with various initial separations between the asters. Fig. 3.3(C) shows the results, where

each point represents the maximum aster speed measured in a single experiment by tracking the aster, and the lines are linear fits to the data for each motor (see SI section 3.5.12 for details on measuring the aster speeds). Interestingly, the ratios of the slopes of these lines match the ratios of motor speeds from Table 1. For example, the slope of the best fit line for Ncd is $\approx 0.0023 \text{ s}^{-1}$ and the best fit slope for Kif11 is $\approx 0.0013 \text{ s}^{-1}$. The ratio of these (Ncd/Kif11) is ≈ 1.8 , which closely matches the ratio of their single motor speeds (≈ 1.6). Similar calculations can be done with these two motors and K401, with the same result. Thus, we conclude that the rate of contraction in the network is set by the motor speed and the increase of network speed with distance is due to adding more connected contractile units.

3.3.4 Motor Properties Determine Outcomes of Motor Competitions

After examining how the properties of individual species of motors affect the structures they generate, we asked what would happen when two motors of opposing polarities are active in the same system. This is inspired by the situation in cells, in which many motors are present and walking towards different poles of the microtubules. Previously, ‘tug-of-war’ experiments have been performed in which specified numbers of motors are attached to microtubules and the microtubule movement is measured [106]. These experiments showed that collections of weaker motors working together can overcome stronger, faster motors [107]. In addition, experiments were performed that tested for spontaneous aster formation in the presence of both Eg5 and XCTK2 (*Xenopus laevis* kinesin-5 and -14). They found that XCTK2 always ‘won’ and formed asters, while with an engineered motor where the Eg5 motor domain was replaced with that of kinesin-1, XCTK2 and kinesin-1 sorted to form a mixture of asters with plus-end centers and minus-end centers [98]. In addition, Nédélec simulated heteromotor complexes to explore the range of steady states that could be achieved when these motors acted on already formed asters, and found that the result depended on the relative velocities of the two motors [108]. Our system offers a unique opportunity to test the effects of motor competition in aster formation by directly dimerizing opposing motors to each other.

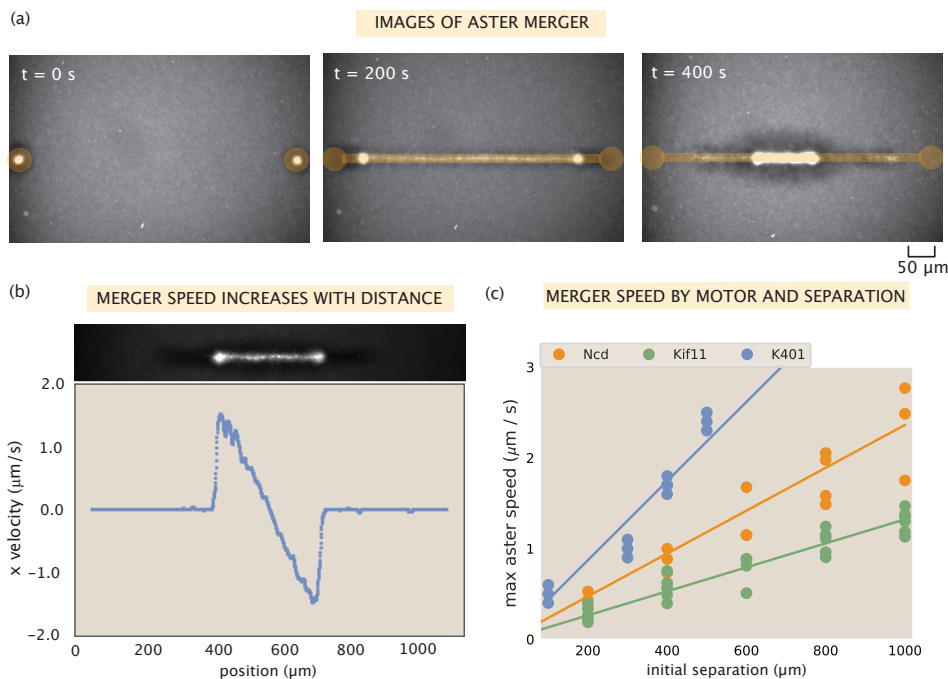


Figure 3.3: Contractile speeds in motor-microtubule networks scale with network size and motor speed. (A) Images of microtubule fluorescence during aster merger. Regions of light activation are shown in orange. (B) Example profile of speeds in an aster merger as a function of linear distance. Each dot is the mean speed measured at that x-position within the network. (C) Maximum merger speed, measured at the ends of the network for each initial separation and motor. Each dot is a single experiment and the lines are best fits to the data.

We first tested if an aster could be formed with heterodimers of K401 and Ncd (the second motor pair in Fig. 3.4(A)). Previous experiments were performed with competing K401-K401 and Ncd-Ncd dimers, and resulted in alternating plus and minus-end centered asters [15]. In our system, we form K401-Ncd heterodimers, which results in bundles of microtubules, but we did not observe overall contraction and reorganization to form an aster (Movie S1). Images of the microtubule and YFP-Ncd fluorescence are shown in the middle row of Fig. 3.4(B). Thus, neither motor dimer was able to ‘win’ to form an aster in this case.

Next, we tested if we could make an aster form by adding extra K401-K401

or Ncd-Ncd dimers. The results of these experiments are shown in Fig. 3.4(B), with the top having the most Ncd relative to K401 and moving downward, this ratio decreases. Upon the addition of even a small amount of K401-K401 dimers, (1:4 K401-K401 to K401-Ncd), K401 was able to form an aster (Movie S2). The addition of even more K401-K401 dimers increased the contraction rate of formation, indicating that the opposing force exerted by Ncd was more easily overcome (Movie S3). Interestingly, a larger amount of Ncd-Ncd dimers were required to make Ncd "win". At 2:1 Ncd-Ncd to K401-Ncd, the network still failed to contract, although a brief contractile phase was observed but then stalled (Movie S4). Ncd was finally able to form an aster with four times as many Ncd-Ncd dimers as heterodimers (Movie S5). This difference in the relative amount of motors required to make K401 and Ncd dominated asters correlate with the larger speed and processivity of K401 motors.

We performed similar experiments with YFP labeled Ncd and mCherry labeled Kif11, allowing us to visualize each motor. Similarly, with only the heterodimer, bundles of microtubules were formed but no overall contraction was observed (Fig. 3.4(C), top row, and Fig. 3.11(A)). With a ratio of only 2:1 Ncd-Ncd to Kif11-Ncd, however, an aster with Ncd at the center was formed. This can be seen by comparing the YFP and mCherry fluorescence in the bottom row of Fig. 3.4(C). While the YFP fluorescence is highly localized in the center of the aster, the mCherry fluorescence is more prevalent in the arms of the aster. The relative fluorescence intensity of each is plotted in Fig. 3.11(B), demonstrating that the ratio of Kif11 motors relative to Ncd motors increases moving radially outward from the center of the aster. This indicates that the minus ends of the microtubules are at the center, and the plus ends are pointing outwards which Kif11 walks towards. Comparing the ratio of motors required for Ncd to "win", Ncd was able to overcome the competing forces of Kif11 at a lower concentration than it did against K401. This is likely due to the lower speed and processivity of Kif11 compared to K401.

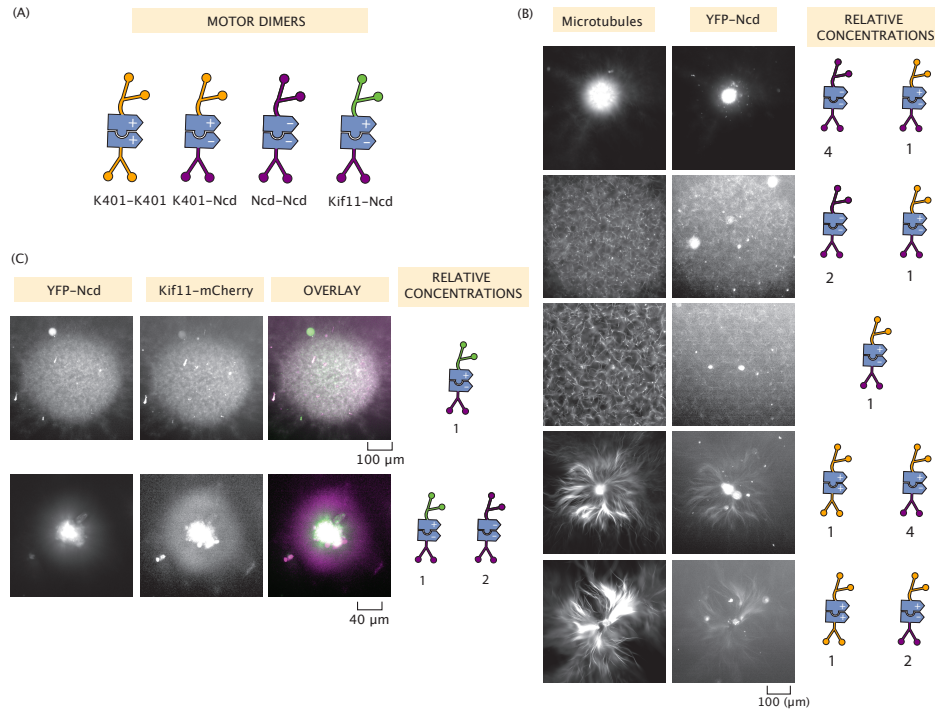


Figure 3.4: Motor competition modulates structure formation. (A) Cartoon of the various motor dimers present. Homodimers as in the previous experiments (K401-K401 and Ncd-Ncd) and heterodimers of K401-Ncd and Kif11-Ncd are formed. (B) Various combinations of K401 and Ncd motors and the resultant structures. Left – microtubule fluorescence images. Right – Ncd fluorescence images. (C) Asters formed with combinations of Ncd and Kif11 motors. In the overlay, Ncd fluorescence is in green, Kif11 is in magenta. For (B) and (C), results are shown for excitation with a 600 μm diameter disk.

3.4 Discussion

In this work, we examined how the properties of kinesin motors determine the mesoscopic properties of the structures they create. The way in which quantities such as motor speed and processivity govern the nature of the resulting motor-microtubule structures has been an open question. Previous attempts have been made to address this question, however most of these are in the context of a single motor without varying the speed or processivity and developing a model that fits the properties they measure [44, 20, 21, 99]. By varying motor speeds, processivities, and directionalities, we were able to quantify and model how these microscopic parameters connect to the properties of mesoscopic structures.

We demonstrate light-controlled aster formation with three different motors. Interestingly, the final aster size from a given illumination region varied depending upon which motor was used. Our leading hypothesis is that the key control variable is the processivity of the motors. Future work needs to be done to understand this effect and build models to explain it. Early work by Surrey et al. found that processivity affected the ‘intensity’ of aster formation in simulations, which may be related to our observations, but to the best of our knowledge no model of this effect has been developed [19]. Further, we assess the distribution of motors and microtubules in the asters we form and develop a model of the steady-state aster that predicts the motor distribution given the measured microtubule distribution, with parameters that relate to the motor speed and processivity. Interestingly, the parameters we infer differ from those we would expect from the single molecule properties of the motors, indicating that higher order effects such as increased processivity of collections of motors, are playing important roles. In addition, we measure contraction speeds in pseudo-one-dimensional networks and find that the speeds are related to the single motor speed. Finally, we show how competition between motors of opposite polarity can lead to tunable structure outcomes. The question of how individual microscopic properties are translated to the hundreds to thousands of times larger scale of the assembly has been an open question in the field of active matter. Our well controlled motor-microtubule system provides an excellent arena in which to investigate this question and our work takes a step towards a mechanistic understanding of motor-microtubule assemblies.

Table 3.2: Plasmids used

pBiex1:Kif11(1-513)-(GSG)₄-mCherry-(GSG)₂-iLid
 pBiex1:Kif11(1-513)-(GSG)₄-mCherry-(GSG)₂-micro
 pBiex1:Kif11(1-513)-(GSG)₄-mVenus-(GSG)₂-micro
 pBiex1:K401-(GSG)₄-mCherry-(GSG)₂-iLid
 pBiex1:K401-(GSG)₄-mVenus-(GSG)₂-micro
 pBiex1:mVenus-(GSG)₄-iLid-(GSG)₂-Ncd236
 pBiex1:mVenus-(GSG)₄-micro-(GSG)₂-Ncd236

3.5 Materials and Methods

3.5.1 Cloning of Motor Proteins

Human kinesin-5 (Kif11/Eg5) 5-513 was PCR amplified from mCherry-Kinesin11-N-18 plasmid (gift from Michael Davidson, Addgene # 55067). This fragment was previously shown to form functional dimers [97]. Kinesin 1 1-401 (K401) was PCR amplified from pWC2 plasmid (Addgene # 15960). Ncd 236-701 was PCR amplified from a plasmid gifted by Andrea Serra-Marques.

The optogenetic proteins, iLid and micro were PCR amplified from pQE-80L iLid (addgene # 60408, gift from Brian Kuhlman) and pQE-80L MBP-SspB Micro (addgene # 60410). mCherry was PCR amplified from mCherry-Kinesin11-N-18 and mVenus was PCR amplified from mVenus plasmid (addgene # 27793).

Constructs were assembled by Gibson assembly of the desired motor protein, optogenetic protein, and fluorophore in order to make the plasmids listed in Table 3.2.

3.5.2 Protein Expression and Purification

Protein expression and purification was done in SF9 cells. Cells were seeded at a density of 1,000,000 cells per mL in a 15 mL volume and transiently transfected with the desired plasmid using Escort IV transfection reagent, then incubated for 72 hours before purification. Cells were collected for purification by centrifugation at 500 g for 12 minutes, and the pellet was resuspended in lysis buffer (200 mM NaCl, 4 mM MgCl₂, 0.5 mM EDTA, 1.0 mM EGTA 0.5 % Igepal, 7 % Sucrose by

weight, 20 mM Imidazole pH 7.5, 10 $\mu\text{g}/\text{mL}$ Aprotinin, 10 $\mu\text{g}/\text{mL}$ Leupeptin, 2 mM ATP, 5 mM DTT, 1 mM PMSF) and incubated on ice for 30 minutes. The lysate was then clarified by centrifugation at 200,000 g for 30 minutes at 4°C. Clarified supernatant was incubated with 40 μL anti-FLAG M2 affinity gel (Sigma-Aldrich A2220) for 3 hours at 4°C. To wash out unbound protein, the resin (with bound protein) was collected by centrifugation at 2,000 g for 1 minute, the supernatant was removed and the resin was washed with wash buffer (for Ncd and Kif11: 150 mM KCl, 5 mM MgCl_2 , 1 mM EDTA, 1 mM EGTA, 20 mM Imidazole pH 7.5, 10 $\mu\text{g}/\text{mL}$ Aprotinin, 10 $\mu\text{g}/\text{mL}$ Leupeptin, 3 mM DTT, 3mM ATP; for K401: M2B with 10 $\mu\text{g}/\text{mL}$ Aprotinin, 10 $\mu\text{g}/\text{mL}$ Leupeptin, 3 mM DTT, 3mM ATP). This was repeated two more times with decreasing ATP concentration (0.3mM and 0.03mM ATP) for a total of three washes. After the third wash, about 100 μL supernatant was left and the bound protein was eluted by incubation with 10 μL FLAG peptide (Sigma-Aldrich F3290) at 4°C for 3 hours. The resin was then spun down by centrifugation at 2,000 g for 1 minute and the supernatant containing the purified protein was collected. Purified protein was then concentrated to a volume of 10-20 μL by centrifugation in mini spin filters (Millipore 50 kDa molecular weight cut-off). Protein was kept at 4°C and used the same day as purification or stored in 50% glycerol at -20°C for longer storage. Protein concentration was determined with QuBit Protein Assay Kit (Thermo Fisher Q33212).

3.5.3 Microtubule Polymerization

Microtubules were polymerized as reported previously ([23] and originally from the Mitchison lab website [62]). In brief, 75 μM unlabeled tubulin (Cytoskeleton) and 5 μM tubulin-AlexaFluor647 (Cytoskeleton) were combined with 1mM DTT and 0.6mM GMP-CPP in M2B buffer and incubated spun at 300,000 g to remove aggregates, then the supernatant was incubated at 37°C for 1 hour to form GMP-CPP stabilized microtubules.

3.5.4 Microtubule Length

The GMP-CPP stabilized microtubules were imaged with TIRF microscopy to determine their length. A flow chamber was made using a KOH cleaned slide, KOH cleaned coverslip (optionally coated with polyacrylamide) and parafilm cut into chambers. The flow cell was incubated with poly-L-lysine for 10 minutes, washed with M2B, then microtubules were flown in. The chamber was sealed with Picodent and imaged with TIRF microscopy.

Microtubules were segmented using home-written Python code and histogrammed to determine the distribution of microtubule lengths (Fig. 3.5).

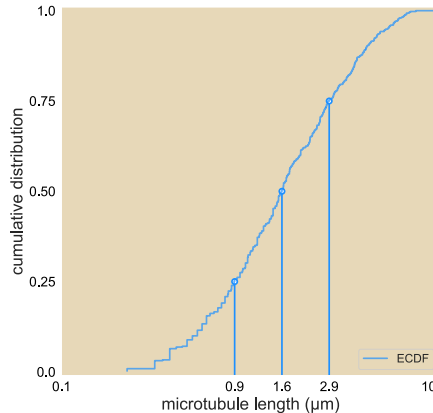


Figure 3.5: Cumulative distribution of microtubule lengths. The 25, 50, and 75% length are marked.

3.5.5 Sample Chamber Preparation

Slides and coverslips were cleaned with Helmanex, ethanol, and KOH, silanized, and coated with polyacrylamide as in [23]. Just before use, slides and coverslips were rinsed with MilliQ water and dried with compressed air. Flow chambers (3 mm wide) were cut out of Parafilm M and melted using a hotplate at 65 °C to seal the slide and coverglass together, forming chambers that are $\approx 70 - 100 \mu\text{m}$ in height and contain $\approx 7 \mu\text{L}$.

3.5.6 Reaction Mixture Preparation

The reaction mixture consisted of kinesin motors (~ 250 nM), microtubules (~ 1 μ M tubulin), and energy mix that contained ATP, an ATP recycling system, a system to reduce photobleaching, F-127 pluronic to reduce interactions with the glass surfaces, and glycerol [23]. To prevent pre-activation of the optogenetic proteins and photobleaching of the fluorophores, the motors and microtubules were always handled in a dark room where wavelengths of light below 520 nm were blocked with a filter or a red light was used to illuminate. The reaction mixture was prepared right before loading into the flow cell and then sealed with Picodent Speed.

3.5.7 Microscope Instrumentation

We performed the experiments with an automated widefield epifluorescence microscope (Nikon TE2000). We custom modified the scope to provide two additional modes of imaging: epi-illuminated pattern projection and LED gated transmitted light. We imaged light patterns from a programmable DLP chip (EKB TEchnologies DLP LightCrafterTM E4500 MKIITTM Fiber Couple) onto the sample through a user-modified epi-illumination attachment (Nikon T-FL). The DLP chip was illuminated by a fiber coupled 470 nm LED (ThorLabs M470L3). The epi-illumination attachment had two light-path entry ports, one for the projected pattern light path and the other for a standard widefield epi-fluorescence light path. The two light paths were overlapped with a dichroic mirror (Semrock BLP01-488R-25). The magnification of the epi-illuminating system was designed so that the imaging sensor of the camera (FliR BFLY-U3-23S6M-C) was fully illuminated when the entire DLP chip was on. Experiments were run with Micro-Manager [65], running custom scripts to controlled pattern projection and stage movement.

3.5.8 Activation and Imaging Protocol

For the experiments in which we make asters with excitation disks of different sizes, we use five positions within the same flow cell simultaneously in order to control for variation within flow cells and over time. Each position is illuminated

with a different sized excitation region: 50, 100, 200, 400, or 600 μm diameter cylinder. Each position was illuminated with the activation light for $\sim 50 - 200$ ms and both the microtubules (Cy5 labelled) and motors (mVenus labelled) were imaged at 10X magnification every 15 seconds. After an hour of activation, a z-stack of the microtubule and motor fluorescence throughout the depth of the flow chamber was taken at 5 μm increments in each position. Typically, one experiment was run per flow chamber. We placed the time limitations on the sample viewing to minimize effects related to cumulative photobleaching, ATP depletion, and global activity of the light-dimerizable proteins. After several hours, inactivated “dark” regions of the sample begin to show bundling of microtubules.

For the aster merger experiments, two 50 μm disks are illuminated at different distances apart. Again, five positions within the same flow cell are chosen, and the separation between the two disks varies for each position: 200, 400, 600, 800, or 1000 μm apart. For K401 and Ncd experiments, the disks are illuminated for 30 frames, for Kif11 experiments, the disks are illuminated for 60 frames (frames are every 15 seconds). Then, a bar ≈ 5 μm wide connecting the disks is illuminated to merge the asters.

3.5.9 Gliding Assay

Motor speeds were determined by gliding assay. Glass slides and coverslips were Helmanex, ethanol, and KOH cleaned. Flow cells with ≈ 10 μL volume were created with double sided sticky tape, and rinsed with M2B buffer. Then, anti-GFP antibody was applied and incubated for 10 minutes. The flow cell was then rinsed with M2B and then mVenus labeled motor proteins (at ~ 5 nM in M2B) were flowed in and incubated for 10 minutes. The flow cell was rinsed with M2B to remove unbound motors and microtubules (in M2B with 3 mM ATP and 1 mM DTT) were flowed in. Microtubules were then imaged using total internal reflection fluorescence (TIRF) microscopy at a rate of one frame per second. Individual microtubules were tracked using custom written python code to determine their speed. The mean microtubule speed (excluding those that were not moving) was determined as the motor speed. Figure 3.6 shows the histogram of speeds obtained

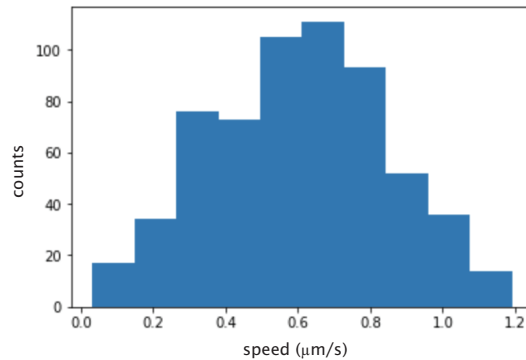


Figure 3.6: Histogram of calculated instantaneous speed of microtubules glided by K401 motors. The mean speed is ≈ 600 nm/s.

for K401 motors purified from SF9 cells.

3.5.10 Disordered Aster Core

We observe that the asters we create have centers that are very dense with motors and microtubules. By fluorescence microscopy, we do not observe organized aster arms in this region and hypothesized that the microtubules are disordered in this region. To assess the extent of microtubule organization in our asters, we imaged asters with a polarized light microscope (Pol-Scope). This microscope utilizes polarized white light to image birefringent substances. Microtubules are birefringent due to their aspect ratio; they interact differently with light polarized parallel to their long axes compared to light polarized perpendicular to their long axes. Thus, the Pol-Scope allows determination of the alignment of microtubules, but not their plus/minus end polarity [109]. When imaged with a Pol-Scope, the arms of our asters are bright, indicating high alignment, and their azimuthal angle confirms that they are radially symmetric around the center (Fig. 3.7). The center of the aster is dark, which we interpret to mean that this region is disordered. It is possible that the microtubules in the center could be aligned pointing in the z direction, which could also result in the center being dark. A disordered center may be a result from steric hindrance due to a high density of microtubules in that region, which prevents the motors from aligning the microtubules. Due to the disorder in the aster center, we exclude this region from our theoretical analysis.

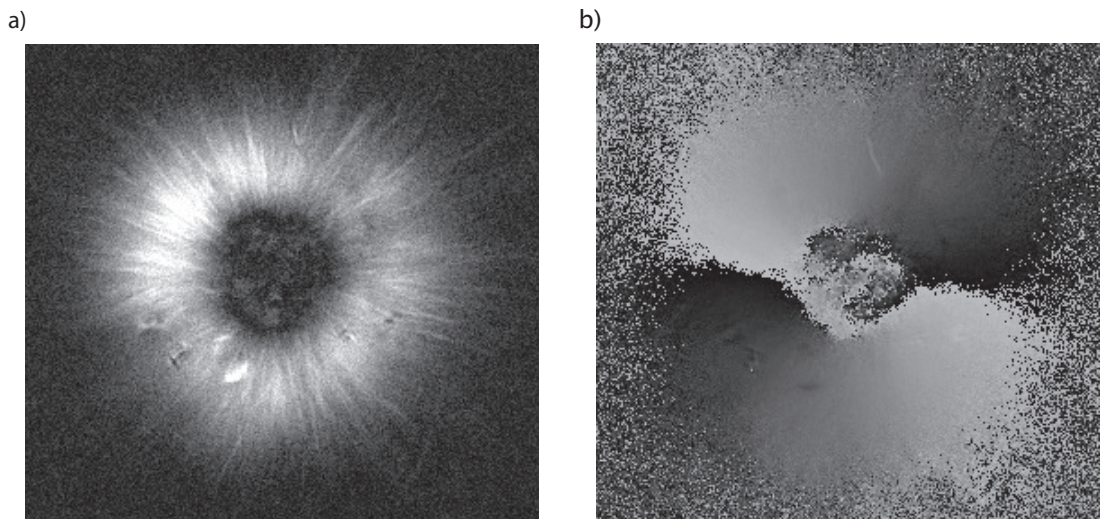


Figure 3.7: Aster centers are disordered and the arms are aligned radially. (A) Retardance image of an aster taken with a Pol-Scope. The arms of the aster result from their alignment and the magnitude of retardance is proportional to the number of microtubules in a bundle. The dark center indicates no alignment of microtubules in that region. (B) Azimuthal angle of microtubule alignment in an aster. Black is 0° and white is 180° .

3.5.11 Aster Size

We formed asters of various sizes using cylindrical illumination regions, ranging in diameter from $50 - 600 \mu\text{m}$. Fig. 3.8 shows representative images of the microtubule fluorescence of asters formed with each motor and each excitation diameter. The yellow circles are by-hand determination of the outer boundary of the aster. In order to measure the size of asters in a more systematic way, we used the measured microtubule distribution, as shown in Fig. 3.1(C). Outside of the central core region of the aster, microtubule fluorescence decreases monotonically before rising again to the background level outside of the activation region. We chose to define the outer radius of the aster as the radius at which the median microtubule fluorescence is twice the background fluorescence. This metric agrees well with a visual inspection of the asters.

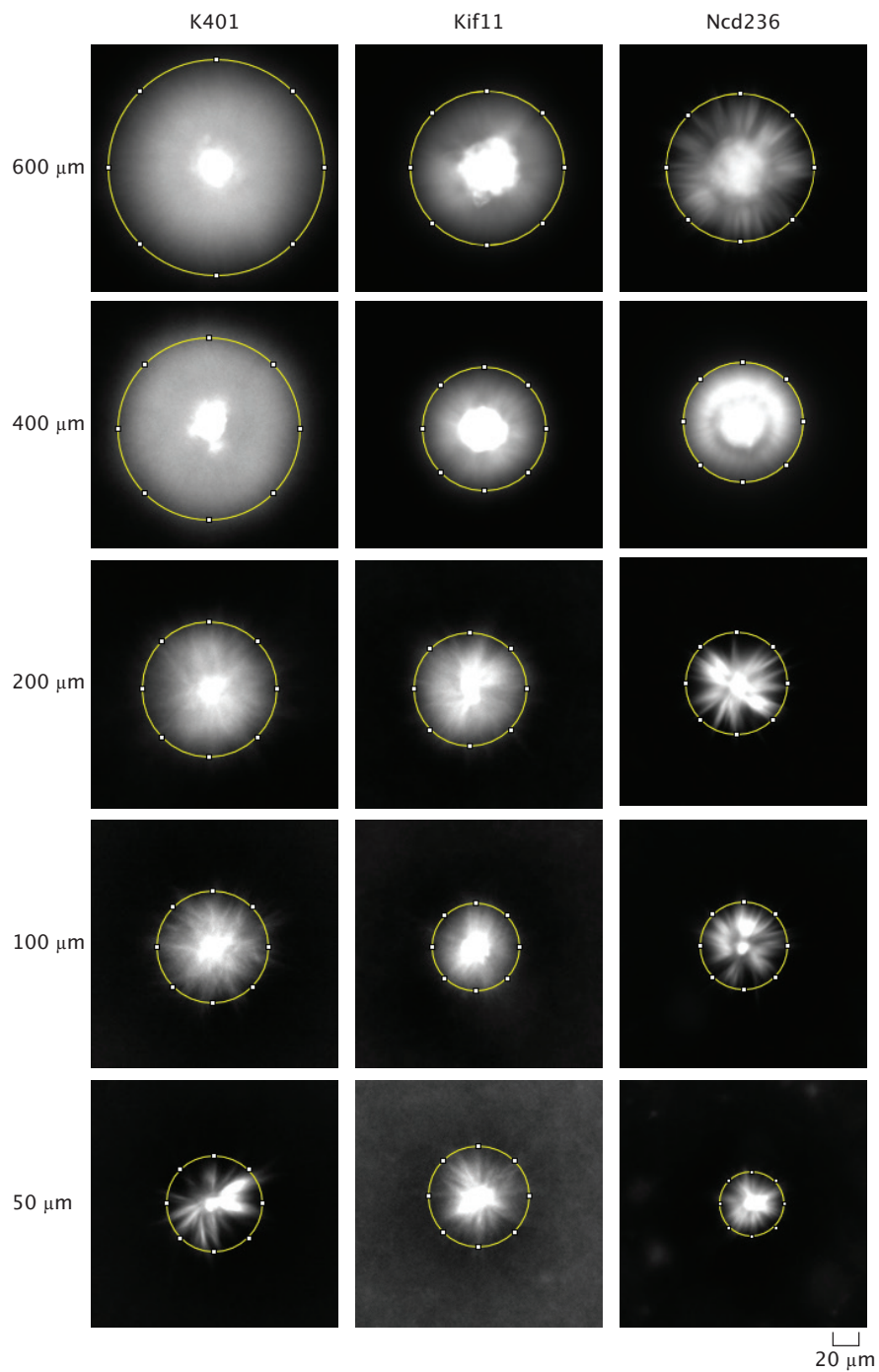


Figure 3.8: Example images of microtubule fluorescence of asters made with each motor used and each excitation diameter. The yellow circle indicates a visual measurement size of the aster.

3.5.12 Merger Analysis

To measure contractile speeds in a pseudo one-dimensional network, we performed aster merger experiments. In these experiments, two asters were formed with 50 μm diameter excitation disks, at varying initial separations (either 200 μm , 400 μm , 600 μm , 800 μm , or 1000 μm). The two disks were illuminated for 30 frames or 60 frames for experiments with Kif11 (15 seconds between frames). After 30 or 60 frames, a thin rectangle was illuminated between the two asters. This caused the formation of a network which contracted, pulling the two asters together. An image of two asters connected in this way is shown in Fig. 3.3.

We used two methods to measure contractile rates. One was using optical flow to measure speeds throughout the network, as shown in Fig. 3.3(B). The open source package, Open CV was used for the optical flow measurements. We used a dense optical flow measurement using `cv.CalcOpticalFlowFarneback`, which calculates optical flow speeds using the Gunnar Farneback algorithm.

To calculate the maximum contractile rate, we tracked the positions of the asters over time. The microtubule fluorescence in the region containing the asters and network (pixels 450 – 650 in the y direction) was summed across the y dimension to make a one-dimensional line trace of fluorescence 3.10A. This trace was smoothed using a butter filter to reduce the noise in the fluorescence intensity. An example smoothed line trace is shown in Fig. 3.10A. The asters were then identified as the large peaks in fluorescence intensity using `scipy.signal.find_peaks`. The identified peaks are shown as red dots in Fig. 3.10A. The corresponding image that the line trace is from is shown in Fig. 3.10B with a 50 x 50 pixel box around the identified aster colored in black. From the identified aster coordinates, the distance between the asters was calculated and the difference between distances in successive frames was used to calculate the merger speed. The reported speed is the speed of a single aster. Thus, the speed calculated from the change in distance between asters is divided by two, assuming the asters are moving at equal speed.

Example traces of calculated aster speeds over time after the beginning of illuminating the bar region is shown in Fig. 3.10c. The speed of the asters increases while the network connecting them forms, peaks, and then decreases as the asters

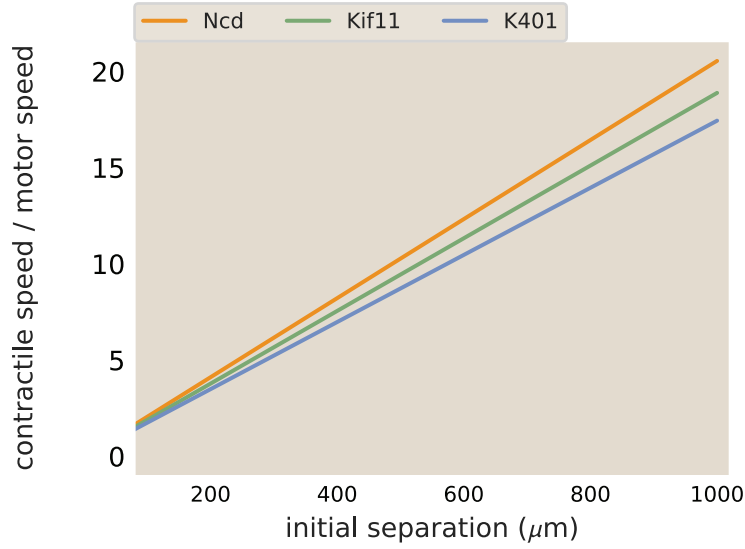


Figure 3.9: Merger speeds depend on motor stepping speed. The best fit lines from the contraction rate measured in aster mergers are plotted, normalized by the measured motor speed.

near each other. The peak speed is reported as the maximum aster speed in the main text. These maximum aster speeds as a function of initial separation were fit to a line using `scipy.optimize.curve_fit` to determine the slope of the increase in speed with separation. The best fit lines determined with this method are plotted along with the data for Ncd and Kif11 in Fig. 3.3c. The fit slope for Ncd was 0.0024 s^{-1} and 0.0013 s^{-1} for Kif11. The ratio (Ncd/Kif11) of these slopes is ≈ 1.8 . In comparison, the ratio between measured motor speeds (115 nm/s for Ncd and 70 nm/s for Kif11) is ≈ 1.6 . The slope calculated for K401 was $\approx 0.0043 \text{ s}^{-1}$. Thus, the ratio of slopes of aster speeds versus separation is in good agreement with the ratio of motor speeds, suggesting that the motor speed sets this slope. To illustrate this point, the best fit lines shown in Fig. 3.3c are plotted again, with the slopes divided by the single motor speed. In this way, the three lines now overlap, as shown in Fig. 3.9.

3.5.13 Model of Network Contraction

In our aster merger experiments, we observed a linear relationship between distance from the center of the network and aster speed. This is seen in the plot of x-velocity vs position in Fig. 3.3(B) and also in the plot of maximum aster speed versus initial separation in Fig. 3.3(C). Juniper et al. observed a similar relationship between position from the center of the network and contraction speed during aster formation in [110]. We interpret this linear trend to mean that the contractile network can be thought of as a series of independently contracting units. Going farther away from the center of the network adds more contractile units, generating the observed speed increase. One question that remained, however, is what sets the slope of the increase of the measured contractile speed. We hypothesize that the single motor speed is one factor. This results in an equation for the rate of contraction in a network to scale as

$$v_n(x) \sim \frac{v_m}{L}x \quad (3.1)$$

where v_n is the network contractile rate, v_m is the motor speed, L is a characteristic length scale, and x is the position from the center of the network. Our observed relationship that the ratio of the slope of speed increase matched the ratio of single motor speeds (as discussed in the main text) supports this hypothesis. Further, we can estimate that $L \approx 50 - 100 \mu\text{m}$, indicating that the characteristic length scale is larger than a single microtubule.

3.5.14 Motor Competition

To test the effects of competing motors, we performed experiments with two motors at a time. The reaction mixture was the same as other experiments, with the two motors added at varying concentrations. With our light activation, we were able to not only form motor homodimers (K401-K401, for example), but also heterodimers that consist of a sets of motor heads that differ (e.g. K401-Ncd). These motor heterodimers set up a direct competition between the motor heads, with one end set of motor heads walking towards microtubule plus-ends and the other set walking towards microtubule minus-ends. In order to create an aster, it

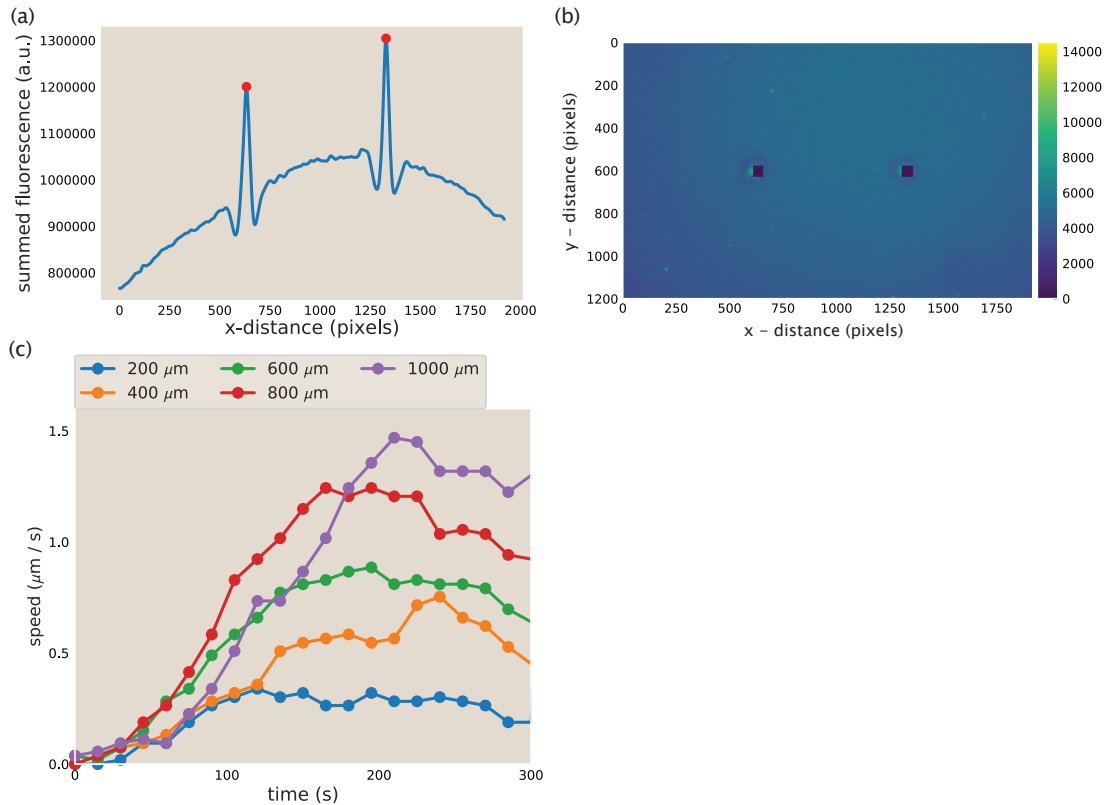


Figure 3.10: Aster identification and speed calculation in aster mergers. (A) Example y-summed microtubule fluorescence from an image during aster merger. The peaks represent the asters and the red dots are the position of the asters as identified by the code. (B) Image corresponding to the fluorescence plotted in (A). The blacked out squares are the location of the asters identified from the peaks in (A). (C) Measured aster speed versus time during aster merger. Each dot is a single measured speed and the colors represent the initial separation of the asters.

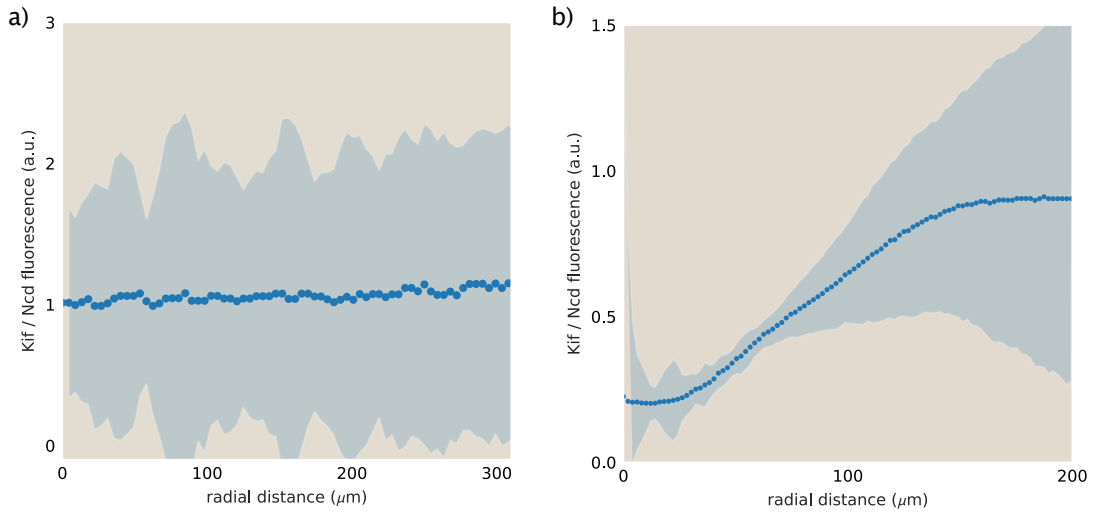


Figure 3.11: Distributions of Kif11 and Ncd in competition experiments. Ratio of Kif11:Ncd fluorescence from an experiment with Kif11-Ncd heterodimers (A) and Ncd-Ncd and Kif-Ncd (B). The shading represents one standard deviation.

is thought that one motor must ‘win’ to polarity sort the microtubules such that either the minus ends or the plus ends are brought together in the center.

We found that depending on the relative concentrations of the two motors, one of several outcomes would occur: aster with the plus-end directed motor in the center, aster with the minus-end directed motor in the center, or failed contraction. In the stalled or failed contraction case, as is the case with K401-Ncd or Kif11-Ncd, bundles of microtubules are formed that do not lead to the reorganization necessary to form an aster. In this case, we see that the fluorescence of mCherry (Kif11) and YFP (Ncd) are about equal and uniform across the activated region (Fig. 3.11(A)). In contrast, with excess Ncd dimers, an aster is formed where Ncd accumulates in the center and Kif11 is more prevalent in the aster arms. This can be seen in Fig. 3.11(B), where the ratio of Kif11:Ncd fluorescence increases moving outward from the aster center.

3.5.15 Motor distributions

3.5.15.1 Model formulation

To predict the spatial distribution of motors in aster structures, we model the dynamic steady-state of free ('f') and bound ('b') motor concentrations via

$$\frac{\partial m_b}{\partial t} = k_{\text{on}}\rho_{\text{MT}}(\mathbf{r})m_f(\mathbf{r}) - k_{\text{off}}m_b(\mathbf{r}) - \nabla \cdot \mathbf{J}_v = 0, \quad (3.2)$$

$$\frac{\partial m_f}{\partial t} = -k_{\text{on}}\rho_{\text{MT}}(\mathbf{r})m_f(\mathbf{r}) + k_{\text{off}}m_b(\mathbf{r}) - \nabla \cdot \mathbf{J}_D = 0. \quad (3.3)$$

Here, k_{on} and k_{off} are the motor binding and unbinding rates, respectively, $\rho_{\text{MT}}(\mathbf{r})$ is the spatially varying microtubule concentration at steady state (measured as μM tubulin), \mathbf{J}_v is the advective flux of bound motors, and \mathbf{J}_D is the diffusive flux of free motors. Our modeling approach is similar to that used by Nédélec *et al.* [44] with the main difference being in the handling of $\rho_{\text{MT}}(\mathbf{r})$. Namely, they imposed a particular functional form on this distribution ($\rho_{\text{MT}}(\mathbf{r}) = 1/|\mathbf{r}|^{d-1}$ with d as the spatial dimension) based on an idealized representation of microtubule organization in an aster, whereas in our treatment $\rho_{\text{MT}}(\mathbf{r})$ stands for the experimentally measured microtubule profiles which cannot be captured through an analogous idealization.

If the free motors have a diffusion coefficient D , then, in the radially symmetric setting considered in our modeling, the diffusive flux will be given by

$$\mathbf{J}_D(r) = -D\nabla m_f(\mathbf{r}) = -Dm'_f(r)\hat{\mathbf{r}}, \quad (3.4)$$

where $\hat{\mathbf{r}}$ is an outward-pointing unit radial vector. And if v is the advection speed of bound motors, then the advective motor flux on radially organized microtubules will be

$$\mathbf{J}_v(r) = -vm_b(r)\hat{\mathbf{r}}. \quad (3.5)$$

Here, we are implicitly assuming that motors constantly walk when bound, ignoring the fact that they can stall upon reaching a microtubule end. We discuss the impact of this effect later in Appendix section 3.5.15.5.

At steady state, the net flux of motors at any radial distance r should be zero ($\mathbf{J}_D(r) + \mathbf{J}_v(r) = 0$), which implies a general relation between the profiles of free and bound motors, namely,

$$0 = -Dm'_f(r)\hat{\mathbf{r}} - vm_b(r)\hat{\mathbf{r}} \Rightarrow$$

$$m_b(r) = - \underbrace{\left(\frac{D}{v}\right)}_{\lambda_0} m'_f(r). \quad (3.6)$$

Above we introduced λ_0 as a length scale parameter that can be interpreted as the distance which is traveled by free and bound motors at similar time scales, i.e., diffusion time scale (λ_0^2/D) = advection time scale (λ_0/v). Note also that the '-' sign at the right-hand side indicates that the free motor population should necessarily have a decaying radial profile ($m'_f(r) < 0$) which is intuitive since at steady state the outward diffusion needs to counteract the inward advection.

To make further analytical progress, we will assume that motor binding and unbinding events are locally equilibrated [99]. This assumption is valid if motor transport is sufficiently slow compared with binding/unbinding reactions. We will justify this quasi-equilibrium condition for the motors used in our study at the end of the section. It follows from this condition that

$$k_{\text{off}}m_b(r) \approx k_{\text{on}}\rho_{\text{MT}}(r)m_f(r) \Rightarrow$$

$$m_b(r) \approx \frac{\rho_{\text{MT}}(r)}{K_d} m_f(r), \quad (3.7)$$

where $K_d = k_{\text{off}}/k_{\text{on}}$ is the dissociation constant. Since the experimental readout reflects the *total* motor concentration ($m_{\text{tot}} = m_f + m_b$), we use our results (Eq. 3.6 and Eq. 3.7) to link $m_{\text{tot}}(r)$ with the microtubule profile $\rho_{\text{MT}}(r)$. Specifically, using Eq. 3.7 we find

$$m_{\text{tot}}(r) = m_b(r) + m_f(r)$$

$$= \left(1 + \frac{\rho_{\text{MT}}(r)}{K_d}\right) m_f(r) \Rightarrow \quad (3.8)$$

$$m_f(r) = \frac{K_d}{K_d + \rho_{\text{MT}}(r)} m_{\text{tot}}(r), \quad (3.9)$$

$$m_b(r) = \frac{\rho_{\text{MT}}(r)}{K_d + \rho_{\text{MT}}(r)} m_{\text{tot}}(r). \quad (3.10)$$

Next, substituting the above expressions for m_f and m_b into Eq. 3.6 and simplifying, we relate the motor and microtubule profiles, namely,

$$\begin{aligned} \frac{\rho_{\text{MT}}(r)}{K_d + \rho_{\text{MT}}(r)} m_{\text{tot}}(r) &= -\lambda_0 \underbrace{\left(\frac{K_d}{K_d + \rho_{\text{MT}}(r)} m'_{\text{tot}}(r) - \frac{K_d \rho'_{\text{MT}}(r)}{(K_d + \rho_{\text{MT}}(r))^2} m_{\text{tot}}(r) \right)}_{m'_f(r)} \Rightarrow \\ \frac{m'_{\text{tot}}(r)}{m_{\text{tot}}(r)} &= -\frac{\rho_{\text{MT}}(r)}{K_d \lambda_0} + \frac{\rho'_{\text{MT}}(r)}{K_d + \rho_{\text{MT}}(r)} \Rightarrow \\ (\ln m_{\text{tot}}(r))' &= -\frac{\rho_{\text{MT}}(r)}{K_d \lambda_0} + (\ln(K_d + \rho_{\text{MT}}(r)))' \Rightarrow \\ \ln m_{\text{tot}}(r) &= -\frac{R_{\text{MT}}(r)}{K_d \lambda_0} + \ln(K_d + \rho_{\text{MT}}(r)) + C_1 \Rightarrow \\ m_{\text{tot}}(r) &= C \left(1 + \frac{\rho_{\text{MT}}(r)}{K_d} \right) \exp \left(-\frac{R_{\text{MT}}(r)}{K_d \lambda_0} \right), \end{aligned} \quad (3.11)$$

where $R_{\text{MT}}(r) = \int \rho_{\text{MT}}(r) dr$ is the integrated microtubule concentration, and $C = K_d e^{C_1}$ is a positive constant. The presence of the multiplicative constant C is a consequence of the fact that the two equations used for deriving our result (Eq. 3.6 and Eq. 3.7) specify the *ratios* of motor populations. Therefore, the result in Eq. 3.11 predicts the relative level of the total motor concentration, given the two effective model parameters (K_d and λ_0), which we infer in our fitting procedure.

Note that the two variable factors on the right-hand side of Eq. 3.11 have qualitatively different structures. The first one is local and depends only on the dissociation constant (an equilibrium parameter), while the second term involves an integrated (hence, non-local) microtubule density term and $\lambda_0 = D/v$ which depends on the advection speed v (a non-equilibrium parameter). As anticipated, in the limit of vanishingly slow advection ($v \rightarrow 0$ or, $\lambda_0 \rightarrow \infty$) the second factor becomes 1 and an equilibrium result is recovered.

Connections to related works. Before proceeding further into our analysis, we briefly compare the expression for the motor distribution (Eq. 3.11) with

analogous results in the literature. Specifically, Nédélec *et al.* [44] studied quasi-two-dimensional asters and in their modeling treated microtubules as very long filaments, all converging at the aster center. This setting implied $\sim 1/r$ scaling of the microtubule concentration. With this scaling, the integrated microtubule concentration in our framework becomes $R_{\text{MT}}(r) = \int \frac{\alpha}{r} dr = \alpha \ln r$ where α is a constant. Substituting this form into the exponential term in Eq. 3.11, we find $\exp\{-(K_d\lambda_0)^{-1}R_{\text{MT}}(r)\} = \exp\{-\alpha(K_d\lambda_0)^{-1}\ln r\} \sim 1/r^\beta$, where $\beta = \alpha(K_d\lambda_0)^{-1}$. It then follows from Eq. 3.11 and the scaling $\rho_{\text{MT}}(r) \sim 1/r$ that the motor concentration is a sum of two decaying power laws (namely, $\sim 1/r^\beta$ and $\sim 1/r^{\beta+1}$) – the result obtained by Nédélec *et al.* [44]. A more detailed calculation can be done to demonstrate that the exponent β matches exactly with the result derived in the earlier work, but for the purposes of our study, we do not elaborate further on this comparison. We note that the experimentally measured microtubule profiles in asters (e.g., Fig. 2b or Fig. 3d) often have an inflection point and cannot be fitted to decaying power law functions (e.g., $1/r^2$ for 3D asters), which is why the idealized setting considered by Nédélec *et al.* [44] cannot be applied to our system.

Another set of works [20, 21] also studied motor distributions in asters, but this time under the assumption of a uniform microtubule concentration ($\rho_{\text{MT}}(r) \sim \text{constant}$). In such a setting, our framework predicts an exponentially decaying motor profile, because $R_{\text{MT}}(r) = \int \rho_{\text{MT}}(r) dr \sim \rho_{\text{MT}} r$ and thus, $m_{\text{tot}}(r) \sim e^{-\rho_{\text{MT}} r / K_d \lambda_0}$. An exponential decay was also the prediction of Lee and Kardar [20], although in their treatment all motors were assumed to be in the bound state. The two distinct motor states were considered in the work by Sankararaman *et al.* [21] who predicted an exponential decay of motor concentration modulated by a power-law tail. One can show, however, that when the decay length scale of motor concentration greatly exceeds the motor processivity (as in the case of asters which we generated), the prediction of Sankararaman *et al.* [21] also reduces into a pure exponential decay, matching the prediction of our model. But since the assumption of a uniform microtubule profile is clearly violated in our system, these predictions are not applicable for us.

Validity of the quasi-equilibrium assumption. Earlier in the section, we

assumed that motor binding and unbinding reactions were locally equilibrated, from which Eq. 3.7 followed. Looking at the governing equation of bound motor dynamics (Eq. 3.2), we can see that this assumption will hold true if $k_{\text{off}}m_b(r) \gg |\nabla \cdot \mathbf{J}_v|$. Substituting the expression of advective flux (Eq. 3.5) and recalling that in three dimensions the divergence of a radial vector $\mathbf{A} = A\hat{\mathbf{r}}$ takes the form $r^{-2}\partial_r(r^2A)$, we rewrite the quasi-equilibrium condition as

$$k_{\text{off}}m_b(r) \gg |\nabla \cdot (-vm_b(r)\hat{\mathbf{r}})| \Rightarrow \quad (3.12)$$

$$\frac{k_{\text{off}}}{v}m_b(r) \gg \left| m'_b(r) + \frac{2}{r}m_b(r) \right|. \quad (3.13)$$

Now, many of the motor profiles can be approximated reasonably well by an exponentially decaying function (see Fig. 3.14 for a collection of experimental profiles). This suggests an empirical functional form $m_b(r) \sim e^{-r/\lambda}$ for the concentration of bound motors, where λ is the decay length scale (note that the constant saturation level contributes to the *free* motor population). This functional form implies that $m'_b(r) \approx -m_b(r)/\lambda$, which, upon substituting into Eq. 3.13, leads to

$$\underbrace{\frac{k_{\text{off}}}{v} m_b(r)}_{\lambda_v^{-1}} \gg \left| -\frac{1}{\lambda}m_b(r) + \frac{2}{r}m_b(r) \right| \Rightarrow \quad (3.14)$$

$$\frac{\lambda}{\lambda_v} \gg \left| \frac{2\lambda}{r} - 1 \right|, \quad (3.15)$$

where $\lambda_v = v/k_{\text{off}}$ is introduced as the motor processivity (distance traveled before unbinding). The processivities (λ_v) of the three different kinesins used in our study, together with the observed ranges of decay length scales (λ) of corresponding motor profiles are listed in Table 3.3. As can be seen, in all cases the ratio λ/λ_v is much greater than one, verifying the intuitive expectation that the length scales of aster structures are much greater than the single run lengths of motors.

It is obvious from the presence of the r^{-1} term on the right-hand side of Eq. 3.15 that the condition can only be satisfied past a certain radius, since r^{-1} becomes very large when r approaches zero. This threshold radius (r^*) is set by $r^* \sim 2\lambda_v$, where the two sides of Eq. 3.15 become comparable. The threshold radial

distance that we choose to isolate the core is at least $5 - 10 \mu\text{m}$ for the asters of our study (see the lower x -limits in the profiles of Fig. 3.14) which exceed r^* at least a few times. This suggests that Eq. 3.15 is valid, justifying our use of the quasi-equilibrium assumption for modeling the motor distribution.

motor	processivity (steps)	processivity, λ_v (μm)	decay length scale, λ (μm)	λ/λ_v
K401	≈ 100	1	10–40	10–40
Ncd236	$\approx 1-100$	$10^{-2} - 1$	5–20	5–2000
Kif11(513)	≈ 10	10^{-1}	10–20	100–200

Table 3.3: Processivities of motors, decay length scales of motor profiles, and corresponding ratios of these two length scales. Step size of $\approx 10 \text{ nm}$ corresponding to the length of a tubulin dimer was used for estimating the motor processivities in μm units. For Ncd motors, the upper limit in processivity corresponds to that of oligomeric motor assemblies. Estimates for the decay length scales λ were made based on the motor profiles in Fig. 3.14.

3.5.15.2 Extraction of concentration profiles from raw images

In this section, we describe our approach for extracting the radial profiles of motor and microtubule concentrations from raw fluorescence images.

Fluorescence normalization and calibration

When taking images with a microscope, several sources contribute to the detected pixel intensities: the camera offset, autofluorescence from the energy mix, and fluorescence coming from the tagged proteins (tubulin or motors). In addition, due to the uneven illumination of the field of view, the same protein concentration may correspond to different intensities in the raw image.

We begin the processing of raw images by first correcting for the uneven illumination. For microtubule images, we use the first movie frames as references with a uniform tubulin concentration in order to obtain an intensity normalization matrix. Each pixel intensity of the final image frame is then rescaled by the corresponding normalization factor.

Although the motor concentration is also initially uniform, the light activation region in the first frame appears photobleached, making it unsuitable for the

construction of a normalization matrix. Instead, we obtain this matrix from the final frame, after masking out the neighboring region of the aster, outside of which the nonuniformity of the fluorescence serves as a proxy for uneven illumination. Intensity normalization factors inside the masked out circular region are obtained through a biquadratic interpolation scheme. The steps leading to a normalized motor image are depicted in Fig. 3.12a.

After fluorescence normalization, we convert intensities into units of protein concentration using calibration factors estimated from images of samples with known protein contents. For K401 and Kif11 motors, we use the conversion 1000 intensity units \rightarrow 815 nM motor dimer. For Ncd dimers, which have fluorescent tags on both iLid and Micro units, we use the 1000 intensity units \rightarrow 407 nM conversion. In all three cases, 200 ms exposure time is used in the imaging. For tubulin, we make a rough estimate that after spinning the energy mix with tubulin, around 1 μ M of tubulin remain, all of which polymerize into GMP-CPP stabilized microtubules. This leads to the calibration of 360 intensity units \rightarrow 1 μ M tubulin (100 ms exposure time).

Aster center identification

In the next step of the profile extraction pipeline, we crop out the aster region from the normalized image and identify the aster center in an automated fashion. In particular, we divide the aster into 16 equal wedges, calculate the radial profile of motors within each wedge, and define the aster center as the position that yields the minimum variability between the motor profiles extracted from the different wedges. Having identified the center, a mean radial profile for the aster is defined as the average of the 16 wedge profiles (Fig. 3.12b).

Inner and outer boundary determination

Since our modeling framework applies to regions of the aster where the microtubules are ordered, we consider the concentration profiles in a limited radial range for the model fitting procedure. As we do not have a PolScope image for every aster to precisely identify the disordered core region, we prescribe a lower threshold on the radial range by identifying the position of the fastest intensity

drop and adding to it a buffer interval (equal to 15% of the outer radius) to ensure that the region of transitioning from the disordered core into the ordered aster ‘arms’ is not included (Fig. 3.12c, top panel). As for the outer boundary, we set it as the radial position where the tubulin concentration exceeds its background value by a factor of two (Fig. 3.12c, bottom panel).

We refer the reader to Supplementary code files for more details on the profile extraction procedure.

3.5.15.3 Model fitting

Here we provide the details of fitting the expression we derived for the motor distribution (Eq. 3.11) to the profiles extracted from aster images. Since smaller asters are typically irregular and hence, do not meet the polar organization and radial symmetry assumptions of the model, we constrain the fitting procedure to larger asters formed in experiments with a minimum light illumination disk diameter of 200 μm .

The different aspects of the fitting procedure are demonstrated in Fig. 3.13. Extracting the average tubulin and motor profiles, we fit our model to the motor profile and obtain the optimal values of the effective parameters K_d and λ_0 . With the exception of a few cases, the optimal pair (K_d, λ_0) corresponds to a distinct peak in the residual landscape (Fig. 3.13c, note the logarithmic scale of the colorbar), suggesting that the parameters are well-defined. The fit to the motor data for the example aster is shown in Fig. 3.13d.

As stated in the main text, we then use the data from the separate aster wedges to assess the quality of fit for each aster. Specifically, keeping the (K_d, λ_0) pair inferred from the average profile fixed, we fit the 16 separate wedge profiles by optimizing over the scaling coefficient C for each of the wedges, and use the model residuals to assess the fit quality. In the set shown in Fig. 3.13e, with the exception of wedge 9 which contains an aggregate near the core radius, fits to all other wedge profiles are of good quality, translating into a low fitting error reported in Fig. 3.2e of the main text.

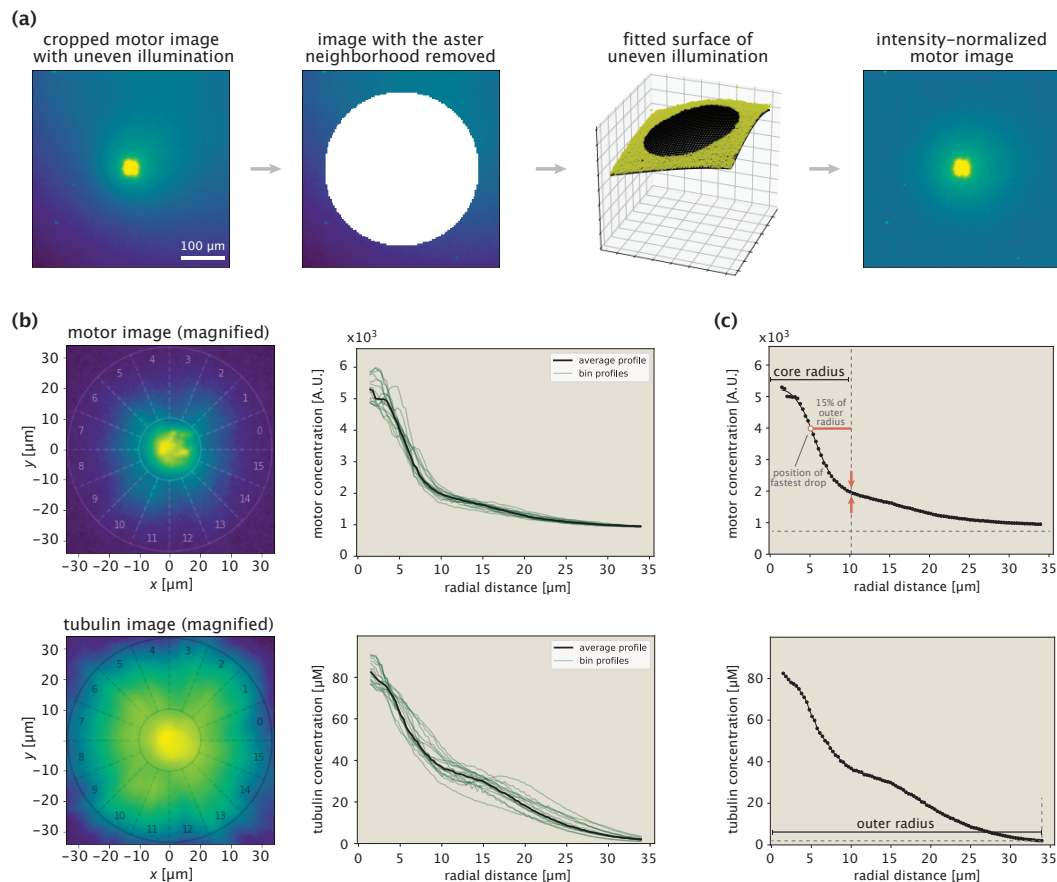


Figure 3.12: Procedure for extracting protein concentration profiles demonstrated on an example aster. (A) Steps taken in normalizing the fluorescence of motor images. The immediate aster region is shown with a saturated color to make it possible to see the nonuniform background fluorescence. (B) Aster center identification and extraction of radial concentration profiles. The numbers indicate the wedges at different angular positions. The two circles in the images indicate the inner and outer bounds. (C) Determination of inner and outer bounds based on the motor and tubulin profiles, respectively.

Repeating this procedure for all other asters, we obtain the best fits to their motor profiles and the corresponding values of the optimal (K_d, λ_0) pairs. The collection of all average profiles, along with the best model fits and inferred parameters are shown in Fig. 3.14.

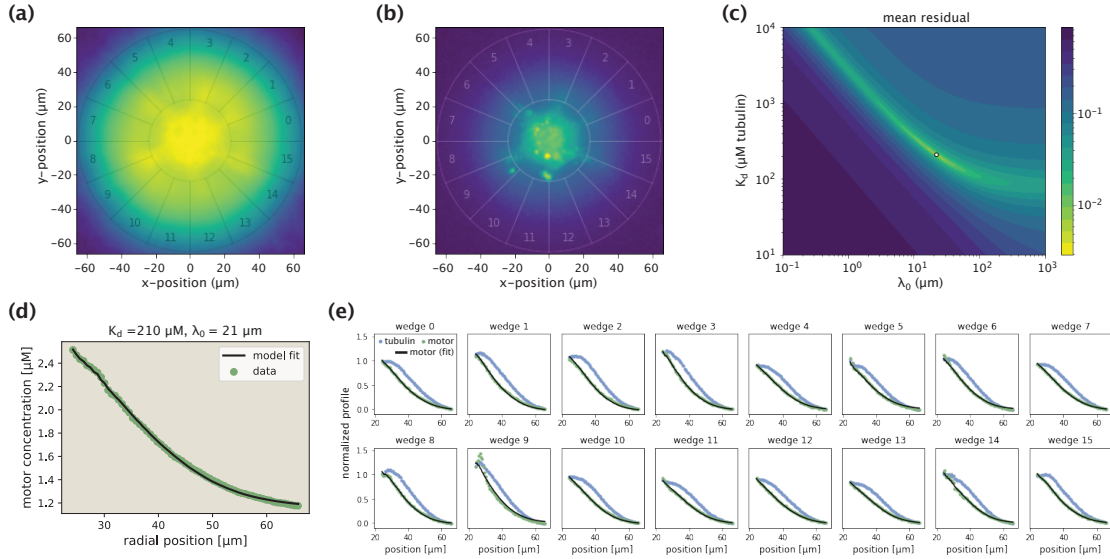


Figure 3.13: Demonstration of the model fitting procedure for average as well as separate wedge profiles. (a,b) Fluorescence images of an example Kif11 aster in tubulin (a) and motor (b) channels. 16 different wedges are separated and numbered. (c) Landscape of fit residuals when varying the two effective parameters K_d and λ_0 . For each pair, an optimal scaling coefficient C is inferred before calculating the residual. The dot at the brightest spot stands for the optimal pair (or, the arrow indicates the location of the optimal pair in the landscape). (d) Average motor profile and the model fit, along with the inferred parameters. (e) Collection of fits to separate wedge profiles using the optimal (K_d, λ_0) pair inferred from the average profile.

3.5.15.4 Expected ratio of K_d values for K401 and Kif11 motors

Here we show the steps in estimating the expected ratio of K_d values for the motors K401 and Kif11. Recall the definition $K_d = k_{\text{off}}/k_{\text{on}}$. Knowing the motor processivities and speeds from Table 1, we calculate the off-rates as $k_{\text{off}} = \text{speed}/\text{processivity}$. This yields a ratio $k_{\text{off}}^{\text{Kif11}}/k_{\text{off}}^{\text{K401}} \approx 1.15$. Then, using the reported on-rates in Valentine and Gilbert [100], we find the ratio of on-rates to be $k_{\text{on}}^{\text{Kif11}}/k_{\text{on}}^{\text{K401}} \approx 0.25$. Taken together, these two results lead to our estimate for the ratio of K_d values reported in the main text, namely, $K_d^{\text{Kif11}}/K_d^{\text{K401}} = (k_{\text{off}}^{\text{Kif11}}/k_{\text{off}}^{\text{K401}})/(k_{\text{on}}^{\text{Kif11}}/k_{\text{on}}^{\text{K401}}) \approx 4.6$.

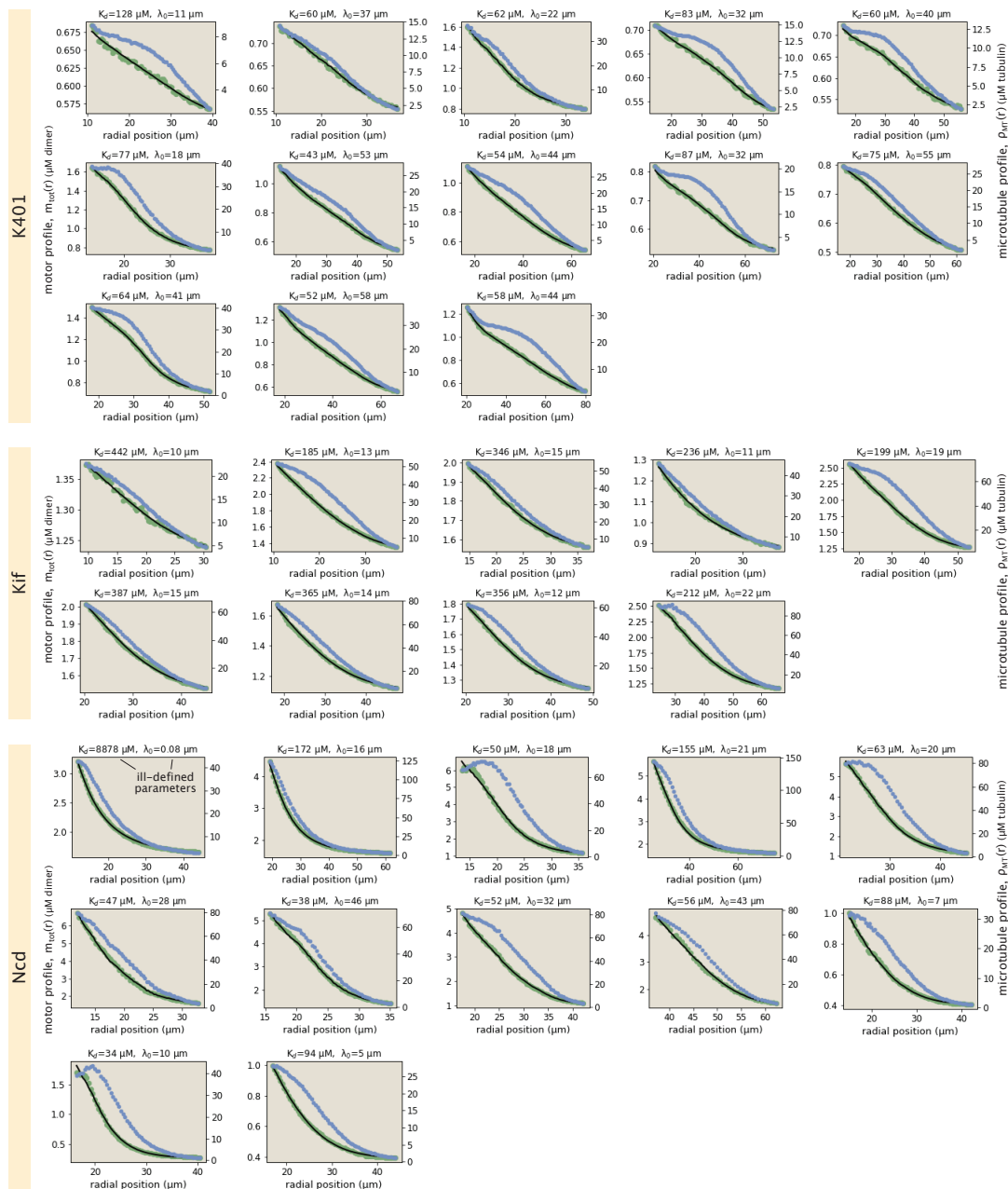


Figure 3.14: Collection of all fits to motor profiles. The green and blue dots represent the radially averaged motor and tubulin concentrations. The solid black lines represent the model fits.

3.5.15.5 Accounting for finite MT lengths

Analysis of purified microtubule images shows that the median length of microtubules is $\approx 1.6 \mu\text{m}$ (Fig. 3.5). Taking the size of a tubulin dimer to be 8 nm, this length translates into the distance traveled in ≈ 200 motor steps, which is comparable to the processivity of K401 motors reported in Table 1 of the main text. Since motors stall when reaching microtubule ends, their effective advection speed will get reduced. Here, we account for this reduction and estimate its magnitude for the different motors used in our study.

Consider the schematic in Fig. 3.15 where a motor is shown advecting on a microtubule with length L . If the distance x between the motor and microtubule end at the moment of binding is less than the motor processivity λ_v , then the motor will reach the end and stall for a time period $\tau = 1/k_{\text{off}}^{\text{end}}$ before unbinding. On the other hand, if x is greater than λ_v , the motor will not stall while bound to the microtubule and hence, its effective speed will not be reduced.

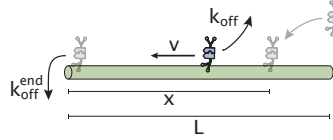


Figure 3.15: Schematic representation of initial motor binding, advection and stalling at the microtubule end.

Assuming that the location of initial binding is uniformly distributed in the $[0, L]$ interval (hence, the chances of binding between x and $x + dx$ is dx/L), we can calculate the effective speeds in the above two cases as

$$\begin{aligned}
 v_{eff}(L < \lambda_v) &= L^{-1} \int_0^L \frac{x}{x/v + \tau} dx, \\
 &= v \left(1 - \frac{v\tau}{L} \ln \left(1 + \frac{L}{v\tau} \right) \right), \\
 v_{eff}(L > \lambda_v) &= \underbrace{L^{-1} \int_0^{\lambda_v} \frac{x}{x/v + \tau} dx}_{\text{initial position } < \lambda_v} + \underbrace{L^{-1} \int_{\lambda_v}^L v dx}_{\text{initial position } > \lambda_v}
 \end{aligned} \tag{3.16}$$

$$\begin{aligned}
&= v \left(\frac{\lambda_v}{L} - \frac{v\tau}{L} \ln \left(1 + \frac{\lambda_v}{v\tau} \right) \right) + v \left(1 - \frac{\lambda_v}{L} \right) \\
&= v \left(1 - \frac{v\tau}{L} \ln \left(1 + \frac{\lambda_v}{v\tau} \right) \right). \tag{3.17}
\end{aligned}$$

As can be seen, in both cases the effective speed is lower than the walking speed v . Now, if $p(L)$ is the distribution of microtubule lengths, then the mean effective motor speed evaluated over the whole microtubule population becomes

$$\langle v_{eff} \rangle = \int_0^\infty v_{eff}(L) p(L) dL. \tag{3.18}$$

We calculate this effective speed for each motor numerically using the measured distribution $p(L)$.

The end-residence time τ was measured for rat kinesin-1 motors to be ≈ 0.5 s [111]. We take this estimate for our K401 motors (*D. melanogaster* kinesin-1) and since, to our knowledge, there is no available data on end-residence times for Ncd and Kif11 motors, we use the same estimate for them (we note that in simulation studies too the end-residence time is typically guessed [19]).

Using this τ estimate, the measured distribution $p(L)$, and the motor speed (v) and processivity (λ_v) values from Table 1, we numerically evaluate the relative decrease in the effective speeds of the motors as

$$\text{K401: } \langle v_{eff} \rangle / v \approx 0.71, \tag{3.19}$$

$$\text{Kif: } \langle v_{eff} \rangle / v \approx 0.94, \tag{3.20}$$

$$\text{Ncd: } \langle v_{eff} \rangle / v \approx 0.99 \text{ (0.71)}. \tag{3.21}$$

Here we made two estimates for Ncd, first using its single-molecule processivity (≈ 1 step) reported in Table 1, and then the 100-fold increased processivity potentially reached due to collective effects mentioned in the main text. As a consequence of this effective speed reduction, we expect factors of ≈ 1.41 , ≈ 1.06 and ≈ 1.01 (1.41) increase in the inferred λ_0 values of K401, Kif11 and Ncd motors, respectively.

3.5.15.6 Broader spread of the tubulin profile

In this section, we discuss the feature of a broader tubulin distribution in greater detail. To gain analytical insights, we first consider an idealized scenario where the motor profile can be represented as an exponential decay with a constant offset for the free motor population (Fig. 3.16a). Such a scenario is approximately met for many of our measured motor profiles. Using our modeling framework, we find that the microtubule distribution corresponding to such a motor profile has the shape of a truncated sigmoid (Fig. 3.16b, see the second part of this section for the derivation). Indeed, microtubule distributions resembling a sigmoidal shape are observed often in our asters (Fig. 3.14), two examples of which are shown in Fig. 3.16c.

One notable implication of this analytical connection between the two profiles is that microtubules necessarily have a broader distribution than motors, once the offset levels at the aster edge are subtracted off. To find whether this is a ubiquitous feature of our asters, we introduce radial distances $r_{1/2}^{(m)}$ and $r_{1/2}^{(t)}$ standing for the positions where the motor and tubulin distributions, respectively, are at their mid-concentrations (Fig. 3.16d). The ratio $(r_{1/2}^{(t)} - r_{in}) / (r_{1/2}^{(m)} - r_{in})$, if greater than 1, would then be an indicator of a wider tubulin profile. Calculating this ratio for all of our asters, we find that it is always greater than 1 for all motor types (Fig. 3.16e), suggesting the generality of the feature.

Idealized scenario with an exponentially decaying motor profile

Here, we first derive the analytical form for the tubulin distribution in the idealized scenario where the motor profile can be approximated as an exponential decay. We then demonstrate that, when normalized, this distribution is always broader than the motor distribution.

We start off by writing the motor distribution as

$$m_{tot}(r) \approx m_{\infty} + \Delta m e^{-(r-r_{in})/\lambda}, \quad (3.22)$$

where λ is the decay length scale, m_{∞} is the background motor concentration corresponding to the free motor population, and Δm is the amplitude of the

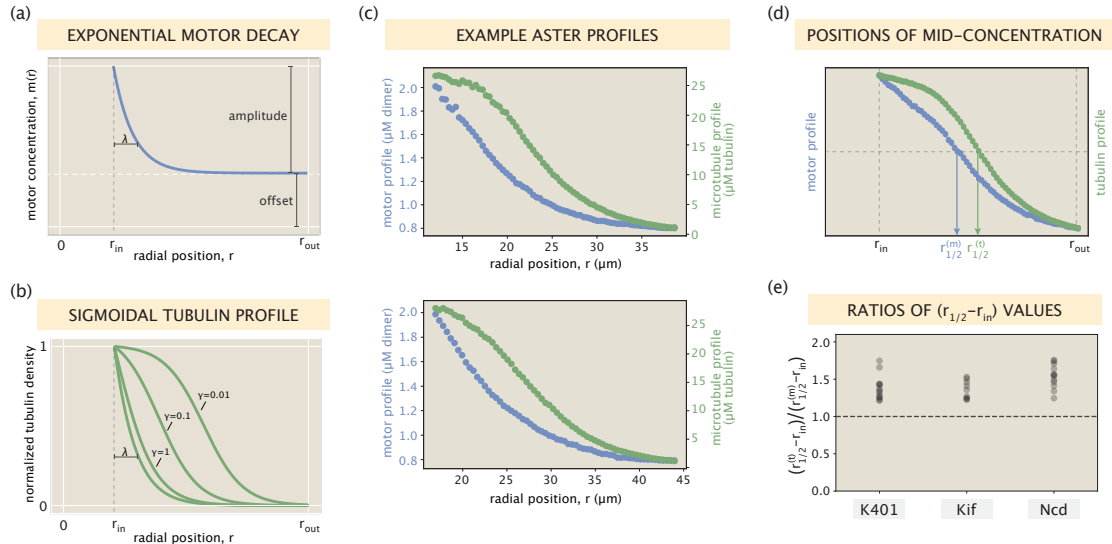


Figure 3.16: Relationship between motor and microtubule distributions. (A) An idealized exponentially decaying motor profile with a constant offset. (B) Set of sigmoidal tubulin profiles corresponding to the exponentially decaying motor profile. The precise curve depends on the shape parameters of the motor profile and the motor type via an effective constant γ (see SI section 3.5.15.6 for details). (C) Two example profiles from Ncd asters that resemble the setting in panels (A) and (B). Blue and green dots represent measured concentrations of motors and microtubules, respectively. (D) Radial positions in the $[r_{in}, r_{out}]$ interval where the motor and tubulin concentrations take their middle values. (E) The ratio $(r_{1/2}^{(t)} - r_{in}) / (r_{1/2}^{(m)} - r_{in})$ calculated for all of the aster profiles.

exponential decay. Next, using Eq. 3.6 as well as the definition $m_{tot}(r) = m_b(r) + m_f(r)$, we set out to obtain the distributions of bound and free motor populations. From Eq. 3.6, we have

$$\begin{aligned} m_b(r) &= -\lambda_0 m_f'(r) \\ &= -\lambda_0 (m_{tot}'(r) - m_b'(r)) \Rightarrow \end{aligned} \quad (3.23)$$

$$m_b'(r) - \frac{m_b(r)}{\lambda_0} = m_{tot}'(r) = -\frac{\Delta m}{\lambda} e^{-(r-r_{in})/\lambda}. \quad (3.24)$$

Solving for $m_b(r)$, we find

$$m_b(r) = C_b e^{(r-r_{in})/\lambda_0} + \Delta m \frac{\lambda_0}{\lambda + \lambda_0} e^{-(r-r_{in})/\lambda}, \quad (3.25)$$

where C_b is an integration constant. Because the approximation Eq. 3.22 applies to a finite radial interval $r \in [r_{in}, r_{out}]$, the constant C_b is generally nonzero. It

specifies the relative contributions of free and bound motor populations to the total motor distribution $m_{\text{tot}}(r)$.

The free motor population is found by simply subtracting Eq. 3.25 from Eq. 3.22, that is,

$$\begin{aligned} m_{\text{f}}(r) &= m_{\text{tot}}(r) - m_{\text{b}}(r) \\ &= m_{\infty} + \Delta m \frac{\lambda}{\lambda + \lambda_0} e^{-(r-r_{\text{in}})/\lambda} - C_{\text{b}} e^{(r-r_{\text{in}})/\lambda_0}. \end{aligned} \quad (3.26)$$

Having obtained expressions for the two motor populations (bound and free), we now recall Eq. 3.7 that relates these two populations through the local tubulin density. Using Eq. 3.7, we find the tubulin density as

$$\begin{aligned} \rho_{\text{MT}}(r) &= K_{\text{d}} \frac{m_{\text{b}}(r)}{m_{\text{f}}(r)} \\ &= K_{\text{d}} \frac{\Delta m \frac{\lambda_0}{\lambda + \lambda_0} e^{-(r-r_{\text{in}})/\lambda}}{m_{\infty} + \Delta m \frac{\lambda}{\lambda + \lambda_0} e^{-(r-r_{\text{in}})/\lambda}} \\ &= K_{\text{d}} \frac{\lambda_0}{\lambda} \frac{e^{-(r-r_{\text{in}})/\lambda}}{\frac{m_{\infty}}{\Delta m} (1 + \lambda_0/\lambda) + e^{-(r-r_{\text{in}})/\lambda}} \\ &= K_{\text{d}} \frac{\lambda_0}{\lambda} \frac{e^{-(r-r_{\text{in}})/\lambda}}{\gamma + e^{-(r-r_{\text{in}})/\lambda}}, \end{aligned} \quad (3.27)$$

where we introduced the effective parameter $\gamma \equiv (m_{\infty}/\Delta m)(1 + \lambda_0/\lambda)$. Eq. 3.27 represents a partial sigmoid, the precise shape of which in the $r > r_{\text{in}}$ region is defined through the parameter γ (Fig. 3.16b).

To formally demonstrate that the tubulin profiles predicted in Eq. 3.27 are necessarily broader than the motor profile, we first normalized them after subtracting off the concentration values at the outer boundary, namely,

$$\begin{aligned} \hat{m}_{\text{tot}}(r) &= \frac{m_{\text{tot}}(r) - m_{\text{tot}}(r_{\text{out}})}{m_{\text{tot}}(r_{\text{in}}) - m_{\text{tot}}(r_{\text{out}})} \\ &= \frac{\Delta m e^{-(r-r_{\text{in}})/\lambda} - \Delta m e^{-(r_{\text{out}}-r_{\text{in}})/\lambda}}{\Delta m - \Delta m e^{-(r_{\text{out}}-r_{\text{in}})/\lambda}} \\ &= \frac{e^{-(r-r_{\text{in}})/\lambda} - e^{-(r_{\text{out}}-r_{\text{in}})/\lambda}}{1 - e^{-(r_{\text{out}}-r_{\text{in}})/\lambda}}, \end{aligned} \quad (3.28)$$

$$\begin{aligned}
\hat{\rho}_{MT}(r) &= \frac{\rho_{MT}(r) - \rho_{MT}(r_{out})}{\rho_{MT}(r_{in}) - \rho_{MT}(r_{out})} \\
&= \frac{\frac{e^{-(r-r_{in})/\lambda}}{\gamma + e^{-(r-r_{in})/\lambda}} - \frac{e^{-(r_{out}-r_{in})/\lambda}}{\gamma + e^{-(r_{out}-r_{in})/\lambda}}}{\frac{1}{\gamma + 1} - \frac{e^{-(r_{out}-r_{in})/\lambda}}{\gamma + e^{-(r_{out}-r_{in})/\lambda}}} \\
&= \frac{(\gamma + 1)(\gamma + e^{-(r_{out}-r_{in})/\lambda})}{(\gamma + e^{-(r-r_{in})/\lambda})(\gamma + e^{-(r_{out}-r_{in})/\lambda})} \times \frac{\gamma e^{-(r-r_{in})/\lambda} - \gamma e^{-(r_{out}-r_{in})/\lambda}}{\gamma - \gamma e^{-(r_{out}-r_{in})/\lambda}} \\
&= \frac{\gamma + 1}{\gamma + e^{-(r-r_{in})/\lambda}} \times \underbrace{\frac{e^{-(r-r_{in})/\lambda} - e^{-(r_{out}-r_{in})/\lambda}}{1 - e^{-(r_{out}-r_{in})/\lambda}}}_{\hat{m}_{tot}(r)} \\
&= \frac{\gamma + 1}{\gamma + e^{-(r-r_{in})/\lambda}} \times \hat{m}_{tot}(r). \tag{3.29}
\end{aligned}$$

The local ratio of normalized tubulin and motor densities then becomes

$$\frac{\hat{\rho}_{MT}(r)}{\hat{m}_{tot}(r)} = \frac{\gamma + 1}{\gamma + e^{-(r-r_{in})/\lambda}} > 1, \tag{3.30}$$

which is always greater than 1 in the $r > r_{in}$ region. This is indicative of the ‘shoulder’ that the normalized tubulin profile often forms over normalized motor profile and demonstrates the broader spread of the tubulin distribution in this idealized setting.

Relative widths of the two distributions

We can see from Fig. 3.16e that the relative widths of the motor and microtubule distributions differ most for Ncd motors (median ratio ≈ 1.55), while for K401 and Kif, the widths are more comparable (median ratio ≈ 1.35 for both motors). Here we offer an explanation for this difference between the motors using the analytical insights developed earlier in the section.

Specifically, Eq. 3.30 and Fig. 3.16b suggest that lower values of γ correspond to broader microtubule distributions. Substituting $r = r_{in}$ in Eq. 3.27, we can write γ as

$$\gamma = \frac{1}{\rho_0} \frac{\lambda_0}{\lambda} - 1, \tag{3.31}$$

where $\tilde{\rho}_0 = \rho_{MT}(r_{in})/K_d$ is the microtubule concentration near the core in units of K_d .

From the concentration profiles in Fig. 3.14, we can estimate the motor decay length scale to be $\lambda \approx 20, 15, 10 \mu\text{m}$ for K401, Kif11, and Ncd, respectively. Then, from our model fitting procedure, we inferred $\lambda_0 \approx 40, 15, 20 \mu\text{m}$, resulting in length scale ratios $\lambda_0/\lambda \approx 2, 1, 2$ for the three motors. Lastly, again inspecting the profiles in Fig. 3.14, we find the microtubule concentrations near the core in K_d units to be $\tilde{\rho}_0 \approx 0.5, 0.2, 1.3$. Note that the microtubule concentration near the core (in K_d units) is the highest for Ncd. In the final step, we substitute these estimated values for λ_0/λ and $\tilde{\rho}_0$ into Eq. 3.31, and evaluate γ for the K401, Kif11, and Ncd, respectively, to be $\gamma \approx \{3, 4, 0.3\}$.

This matches well with the intuition of our simple analytical study (Fig. 3.16b) and the observed ratios of distribution widths (Fig. 3.16e). Namely, large values of γ for K401 and Kif11 (≈ 3 and 4 , respectively) suggest a closer correspondence between the normalized motor and microtubule profiles, while the lower values of γ for Ncd (≈ 0.3) suggests a wider microtubule distribution, as was observed in our asters.

3.6 Supplementary Movies

Movie S1 (separate file) Activation of K401-Ncd tetramers leads to microtubule bundle formation. Microtubule fluorescence following activation of K401-iLid and Ncd-micro with a $600 \mu\text{m}$ diameter excitation region is shown.

Movie S2 (separate file) Aster formation by K401-K401 and K401-Ncd tetramers. There is approximately a 1:4 ratio of K401-K401 to K401-Ncd tetramers. Microtubule fluorescence is shown.

Movie S3 (separate file) Aster formation by K401-K401 and K401-Ncd tetramers. Approximately a 1:2 ratio of K401-K401 to K401-Ncd tetramers is used. Microtubule fluorescence is shown.

Movie S4 (separate file) Failed contraction by Ncd-Ncd and K401-Ncd tetramers. There is approximately a 2:1 ratio of Ncd-Ncd to K401-Ncd tetramers. Microtubule fluorescence is shown.

Movie S5 (separate file) Aster formation by Ncd-Ncd and K401-Ncd tetramers. Approximately a 4:1 ratio of Ncd-Ncd to K401-Ncd tetramers is used. Microtubule fluorescence is shown.

Chapter 4

The Human Impacts Database: A Quantitative Repository of Humanity's Influence on the Planet

This chapter is partially derived from the preprint : Chure, G., *et al.* The Athropocene by the numbers: A Quantitative Snapshot of Humanity's Influence on the Planet. *arXiv* (2021) DOI: <https://arxiv.org/abs/2101.09620>

4.1 Abstract

The Human Impacts Database (www.anthroponumbers.org) is a curated searchable resource housing quantitative data relating to the diverse anthropogenic impacts on our planet, with topics ranging from sea level rise, to livestock populations, greenhouse gas emissions, fertilizer use, and beyond. Each entry in the database relates a quantitative value (or a time-series of values) along with clear referencing of the primary source, the method of measurement or estimation, an assessment of uncertainty, links to the underlying data, as well as a permanent identifier called an "HuID." While there are other databases that house some of these values, they are typically focused on a single topic area like energy usage or greenhouse gas emissions. The Human Impacts Database provides centralized access to quantitative information about the myriad ways in which humans impact

the Earth, giving links to more specialized databases for interested readers. Here, we outline the structure of the database and describe our curation procedures. Finally, we use this database to generate a graphical summary of human impacts on the Earth as of 2021, illustrating both their numerical values and their dense interconnections.

4.2 Introduction

One of the most important scientific developments of the last two centuries is the realization that the evolution of the Earth is deeply intertwined with the evolution of life. Perhaps the most famous example of this intimate relationship is the large-scale oxygenation of Earth’s atmosphere following the emergence of photosynthesis [24]. This dramatic change in the composition of the atmosphere is believed to have caused a massive extinction, as the organisms living at the time were not adapted to an oxygenated atmosphere [112, 113, 114]. Over the past 10,000 years, humans have likewise transformed the planet, directly affecting the rise and fall of ecosystems [115, 116, 117, 118, 119, 120, 121, 122, 123], the pH and surface temperature of the oceans [29, 30] the composition of terrestrial biological and human-made mass [124, 32], the planetary albedo and ice cover [125, 126, 127, 128, 129, 130, 131, 132, 133, 134], and the chemistry of the atmosphere [135, 136, 137, 26, 138, 25] to name just a few examples. The breadth of human impacts on the planet is so diverse that it touches on nearly every facet of the Earth system and every scientific discipline.

Technological advances in remote sensing, precision measurement, and computational power have made it possible to measure these anthropogenic impacts with unprecedented depth and resolution. However, as scientists with different training use distinct methods for measurement and analysis, report data in different units and formats, and use nomenclature differently, these studies can be very challenging to understand and relate to one another. Even seemingly simple questions such as “how much water do humans use?” can be difficult to answer when a search of the scientific literature yields an array of complicated analyses with different units, varying definitions about what constitutes water use, and distinct approaches to

quantifying flows. This problem persists beyond the primary scientific literature as governmental, intergovernmental, and industry datasets can be similarly cryptic and laborious to interpret.

Writing, as several of us are, from California, where we now have a “wildfire season” and a multi-decadal drought [139, 140], we felt compelled to overcome these difficulties and develop a deeper understanding of the ways in which human activities might have produced such dramatic and consequential changes in our local and global environment. In pursuit of basic understanding, we asked many simple questions like “how much water and land do humans use?” and “how much methane is emitted annually?” In our search for answers we often encountered the same challenges: disparate technical studies written for expert audiences must be understood, evaluated and synthesized just to answer simple questions. It seemed to us that a referenced compendium of “things we already know” akin to the CRC Handbook of Chemistry and Physics would be very useful for us and others.

In building the Human Impacts Database, we took inspiration from our previous experience building and using the BioNumbers Database [141] (bionumbers.hms.harvard.edu), a compendium of quantitative values relating to cell and organismal biology. Over the past decade, the BioNumbers Database has become a widely-accessed resource that serves not only as an index of biological numbers, but also as a means of finding relevant primary literature, learning about methods of measurement, and teaching basic concepts in cell biology [142]. We believe that a reference for quantitative data about the extent of human impacts on our planet would be similarly transformative for researchers, students and interested non-scientists. As is true for BioNumbers, reading an entry in the Human Impacts Database is not a replacement for reading the primary literature. Nonetheless, as we hope to communicate in this manuscript, building and using the Human Impacts Database has taught us a great deal about our planet, the human civilizations living on it, and their collective impacts on the Earth. We do not know which approaches to remediating these impacts are most efficient, expedient or cost-effective, but we are convinced that proposals should be evaluated in the light of a comprehensive and quantitative understanding of the Earth-human system.

Such an understanding surely begins with a concerted look at the data, data that we have begun assembling in the Human Impacts Database.

4.3 Results

4.3.1 Finding and Compiling Numbers from Scientific Literature, Governmental and Non-Governmental Reports, and Industrial Datasets

We have established the Human Impacts Database (<http://anthroponumbers.org>) as a repository for the rapid discovery of quantities describing the Earth-human system. As of this writing, the database holds > 250 unique manually-curated entries covering a breadth of data sources, including primary scientific literature, governmental and non-governmental reports, and industrial communiques. Before it is added to the database and made public, Each entry is vetted extensively by the administrators (our curation procedures are fully described in the Supplemental Information). While these 250 entries include those we consider to be essential for a quantitative understanding of human impacts on Earth, it is not an exhaustive list. This database will continue to grow and evolve as more data becomes publicly released, our understanding of the human-earth system improves, and members of the scientific community suggest values to be added.

To illustrate the structure of a database entry, let us consider the most emblematic one: the atmospheric CO₂ concentration as measured by the Mauna Loa Observatory (Fig. 4.1). At the top of the entry we find the title and category (Fig. 4.1 A-B). Primary categorization falls into one of five classes: “Land,” “Water,” “Energy,” “Flora Fauna,” and “Atmospheric & Biogeochemical Cycles.” Of course, these categories are broad and entries can be associated with several categories. For this reason, each entry is also assigned a narrower “subcategory,” such as “Agriculture,” “Urbanization,” or “Carbon Dioxide.” While this categorization is not meant to be exhaustive, and many other schemes could be implemented, we found this organization allowed us to quickly browse and identify quantities of interest.

Following the title and categorization, we find the measured atmospheric CO₂ concentration. This corresponds to the most recent reported measurement, which is,

as of writing, roughly 415 ppm in 2020 (parts per million, Fig. 4.1 C). Importantly, we report an approximate CO₂ concentration rather than a precise one to many significant digits. While the most recent entry in the linked dataset (Fig. 4.1I) gives a monthly average value of 413.95 ppm for December of 2020, this value does not account for error in the measurement, fluctuations throughout December, or seasonal oscillations in atmospheric CO₂. Therefore, we report a rounded value of 415 ppm. CO₂ measurements are quite accurate, but other measurements and inferences recorded in the Human Impacts Database are less so. We therefore strive to give an assessment of the uncertainty for all values. This can be in the form of a confidence interval, as for HuID 11827, which reports a 90% confidence interval on the extent of sea level rise due to thermal expansion since 1900, or bounds on the value, as in HuID 44641, which reports a lower bound on the number of animal extinctions since 1500 CE. In addition to error assessment, we also aim to provide legible units for all entries. Though atmospheric CO₂ is commonly reported in parts per million (ppm) units, we also report this value in other equivalent units, including the mole and mass fractions of CO₂, and the total mass of CO₂ in the atmosphere in kg CO₂ (Fig. 4.1 C). Whenever possible, entries will report values in multiple units to make quantities accessible to readers coming from diverse backgrounds.

Following the numerical value, we find the Human Impacts Database identifier (HuID, Fig. 4.1 D). The HuID is a randomly-generated five-digit number that serves as a permanent identifier of the entry. Because the HuID is permanent and static, it can be used for referencing. Rather than identifying a single value, we consider the HuID a pointer to a particular entry, so that HuID 81043 can be used to reference the atmospheric CO₂ concentration in 2020 and 1980 (Fig. 4.1 E). For example, to reference the present-day atmospheric CO₂ concentration, one could report the value as “ \approx 415 ppm (HuID 81043:2020).” Additionally, since each entry comes from a single source, we may have more than one HuID reporting similar quantities, for example HuIDs 69674 and 72086 report recent measurements of the temperature of the upper ocean.

The “Summary” field (Fig. 4.1 F) gives a succinct description of the quantity and

its relationship to “human impacts” broadly-construed, along with other pertinent information. This could include a more detailed definition of terms used in the quantity or useful historical information about the measurement. In our example of atmospheric CO₂ concentration, the summary explains that measurement is made at the Mauna Loa observatory and points out the seasonal oscillations that are observed. The following “Method” field describes the method by which the quantity was measured, inferred, or estimated (Fig. 4.1 G). This field also provides an assessment of the uncertainty in the value, which may include a description of how confidence intervals were computed or a list of critical assumptions that were made to estimate missing data.

All fields through “Method” (Fig. 4.1 A-G) depend on manual curation and interpretation by database. The following two fields, “Source” and “Dataset” (Fig. 4.1 H-I) provide direct links to the primary source reference and the relevant data. Both of these fields are direct links (shown as insets in Fig. 4.1). The “Source” can point either to the published scientific literature or the resource page of a governmental, industrial, or non-governmental organization data deposition URL. The “Dataset” field links directly to a CSV version of the datafile in our GitHub repository. As discussed in the Supplement, these data files have been converted into a “tidy-data” format [143] by database administrators, which maximizes programmatic readability.

When possible, a graphical time-series of the data is also presented as an interactive plot (Fig.4.1 J). These plots enable users to quickly apprehend time-dependent trends in the data without downloading or processing the dataset. While not available for every entry, the majority of quantities we have curated in the Human Impacts Database contain measurements over time. The last field gives the username of the administrator who generated this entry (Fig. 4.1 K). Their affiliation and contact information is available on the database’s “About” page. We invite the reader to contact the administrators directly for questions, concerns, or suggestions.

While Figure 4.1 is a representative example, each quantity in the Human Impacts Database tells a different story. Easy centralized access to different entries

allows users to learn about the magnitude of human impacts, and also study the interactions between different human activities, which, as we discuss in the next section, are deeply intertwined.

4.3.2 Global Magnitudes

In Figure 2, we provide an array of quantities that we believe to be key in developing a “feeling for the numbers” associated with human impacts on the Earth system. All of the quantities in Figure 2 are drawn from entries in the database and grouped into the same categories used in the database: land, water, flora and fauna, atmosphere and biogeochemical cycles, and energy (see color scheme at the top of Fig. 4.2 A). Though the impacts considered necessarily constitute an incomplete description of human interaction with the planet, these numbers encompass many which are critically important, such as the volume of liquid water resulting from ice melt (Fig. 4.2 B), the extent of urban and agricultural land use (Fig. 4.2 H), global power consumption (Fig. 4.2 N), and the heat uptake and subsequent warming of the ocean surface (Fig. 4.2 S). In many cases, the raw numbers are astoundingly large and can therefore be difficult to fathom. Rather than reporting only bare “scientific” units, we present each quantity (when possible) in units that are relatable to a broad audience. For example, to give context for the annual global mass of CO₂ emissions in kilograms, we note that this mass is equivalent to 2.5 two-tonne pickup trucks per person on the planet per year.

Exploring these numbers reveals several surprising quantities and relationships. For example, agriculture repeatedly appears as a major contributor to many human impacts. Agriculture dominates both global land (Fig. 4.2 H) and water use (Fig. 4.2 L), and accounts for approximately a third of global tree cover area loss (Fig. 4.2 O). In addition, an enormous amount of nitrogen is synthetically fixed through the Haber-Bosch process to produce fertilizer (Fig. 4.2 F), which is a major cause of emission of N₂O (Fig. 4.2 K), which is a potent greenhouse gas. About 30 billion livestock are raised on agricultural lands (Fig. 4.2 E), which, together with rice paddies, produce a majority of anthropogenic methane emissions (i.e. the greenhouse gas CH₄, Fig. 4.2 K). On the other hand, urban land area

(A) Atmospheric CO₂ concentration

(B) Atmospheric & Biogeochemical Cycles Carbon Dioxide
This quantity comes from a time series measurement and the most recent value (1964-2020) is reported.

(C) Value: **≈ 415 parts per million**
This is equivalent to...
 ≈ 4.15 × 10⁻⁶ mol CO₂ mol⁻¹ air
 ≈ 6.30 × 10⁻⁴ kg CO₂ kg⁻¹ air
 ≈ 3.25 × 10¹⁵ kg CO₂

(D) HuID: 81043

(E) Relevant Year(s): 1964-2020

(F) Summary: Data obtained from monthly mean carbon dioxide measured at Mauna Loa Observatory, Hawaii by the Scripps CO₂ program. The carbon dioxide data on Mauna Loa constitute the longest record of direct measurements of CO₂ in the atmosphere. They were started by C. David Keeling of the Scripps Institution of Oceanography in March of 1958 at a facility of the National Oceanic and Atmospheric Administration.

(G) Method: Uncertainty in measurements is not reported in original source, and it is assumed lower than the monthly variability. The Mauna Loa data are being obtained at an altitude of 3400 m in the northern subtropics, and may not be the same as the globally averaged CO₂ concentration at the surface. The mass of CO₂ is obtained from the concentration using the molar mass of CO₂, 44 g mol⁻¹; the molar mass of air, 29 g mol⁻¹; and the mass of the atmosphere, 5.15 × 10¹⁸ kg.

(H) Source: [Scripps CO₂ Program Primary Mauna Loa CO₂ Record \(2020\)](#)

(I) Dataset: [Monthly atmospheric CO₂ measurements from Mauna Loa Observatory \(monthly_co2_data_processed.csv\)](#)

(J) Trend

(K) Added by: ilopezgo

Station Name	Station Code	Latitude	Longitude	Elevation (m)
Mauna Loa Observatory, Hawaii	MLO	19.5° N	155.6° W	3397

Measurement	Frequency	Data File	Dates
CO ₂	monthly	monthly_hilo_co2_data	1958 - present

(I) Dataset: [Monthly atmospheric CO₂ measurements from Mauna Loa Observatory \(monthly_co2_data_processed.csv\)](#)

year	month	date (decimal)	Reported value	Concentration (ppm)
1958	3	1958.203	monthly mean	315.7
1958	4	1958.288	monthly mean	317.46
1958	5	1958.37	monthly mean	317.51
1958	6	1958.455	monthly mean	
1958	7	1958.537	monthly mean	315.86
1958	8	1958.622	monthly mean	314.93
1958	9	1958.707	monthly mean	313.21
1958	10	1958.789	monthly mean	
1958	11	1958.874	monthly mean	313.33
1958	12	1958.958	monthly mean	314.67
1959	1	1959.041	monthly mean	315.58
1959	2	1959.126	monthly mean	316.49

Figure 4.1: A representative entry in the Human Impacts Database. The entry page for HuID 81043 – “Atmospheric CO₂ concentration” is diagrammed with important features highlighted. Each entry in the Human Impacts Database has a (A) name, (B) primary and secondary categorization, (C) the numerical value with other units when appropriate, (D) a 5-digit permanent numeric identifier, (E) years for which the measurement was determined, (F) a brief summary of the quantity, (G) the method of determination, (H) a link to the source data, and (I) a link to a processed version of the data saved as a .csv file. When possible, a time series of the data is presented. (K) Finally, each entry lists the username of the administrator who curated the quantity. Their contact information is available on the anthroponumbers.org “About” page.

accounts for a very small fraction of land area use ($\approx 1\%$, Fig. 4.2 H), and the expansion of cities and suburbs accounts for only $\approx 1\%$ of global tree cover area loss (Fig. 4.2 O). This is not to say, however, that urban centers are negligible in their global impacts. As urban and suburban areas currently house more than half of the global human population (Fig. 4.2 J), many human impacts are linked to industries that directly or indirectly support urban populations' demand for food, housing, travel, electronics and other goods. For example, the pursuit of urbanization is the dominating factor in the mass of earth moved on an annual basis (Fig. 4.2 W).

Collectively, the ≈ 8 billion humans on Earth (Fig. 4.2 J) consume nearly 20 TW of power, equivalent to 23 one hundred Watt light bulbs per person (Fig. 4.2 N). Around 80% of this energy derives from the combustion of fossil fuels (Fig. 4.2 P). This results in a tremendous mass of CO_2 being emitted annually (Fig. 4.2 K) of which only \approx one-half remains in the atmosphere (HuID 70632). A sizable portion of the emissions are absorbed by the oceans (HuID 99089), leading to a steady increase in ocean acidity (Fig. 4.2 G) and posing risks to marine ecosystems [144]. Furthermore, increasing average global temperatures primarily caused by greenhouse gas emissions contribute to sea level rise not only in the form of added water from melting ice (Fig. 4.2 B and M), but also due to thermal expansion of ocean water (Fig. 4.2 M), which accounts for $\approx 30\%$ of observed sea level rise [145]. These are just a few ways in which one can traverse the impacts illustrated in Figure 4.1, revealing the remarkable extent to which these impacts are interconnected. We encourage the reader to explore this figure in a similar manner, blazing their own trail through the values.

4.3.3 Regional Distribution

While Figure 4.2 presents the magnitude of human impacts at a global scale, it is important to recognize that these impacts – both their origins and repercussions – are variable across the globe. That is, different societies vary in their preferences for food (e.g. Americans consume little fish) and modes of living (e.g. apartments vs houses), have different levels of economic development (e.g. Canada as compared

THE ANTHROPOCENE by the numbers

Griffin Chure^{1,a}, Rachel A. Banks², Avi I. Flamholz², Nicholas S. Sarai³, Mason Kam⁴, Ignacio Lopez-Gomez⁵, Yinon M. Bar-On⁶, Ron Milo⁶, Rob Phillips^{2,7,*}

California Institute of Technology, Pasadena, CA, USA, 91125:
 1. Department of Applied Physics, 2. Division of Biology and Biological Engineering, 3. Division of Chemistry and Chemical Engineering
 4. Chan-Zuckerberg BioHub, San Francisco, CA, 94158
 5. Department of Environmental Science and Engineering, 7. Department of Physics
 *. Address correspondence to rphillips@jplab.caltech.edu
 6. Weizmann Institute of Science, Rehovot 7610001, Israel; Department of Plant and Environmental Sciences
 a. Current address: Department of Biology, Stanford University, Stanford, CA, USA.

ABSTRACT

A candidate for the greatest experiment of the last 10,000 years is the presence and action of modern human beings on planet Earth, the often complex results of which are now being felt on various fronts. While there has been a deluge of careful studies exploring each facet of these "human impacts" on Earth, they are often highly focused and necessarily technical, rarely displaying their integration with other human impacts as a whole. In this snapshot, we present a diverse (yet necessarily incomplete) array of quantities that summarize the broad reach of human action across the planet.

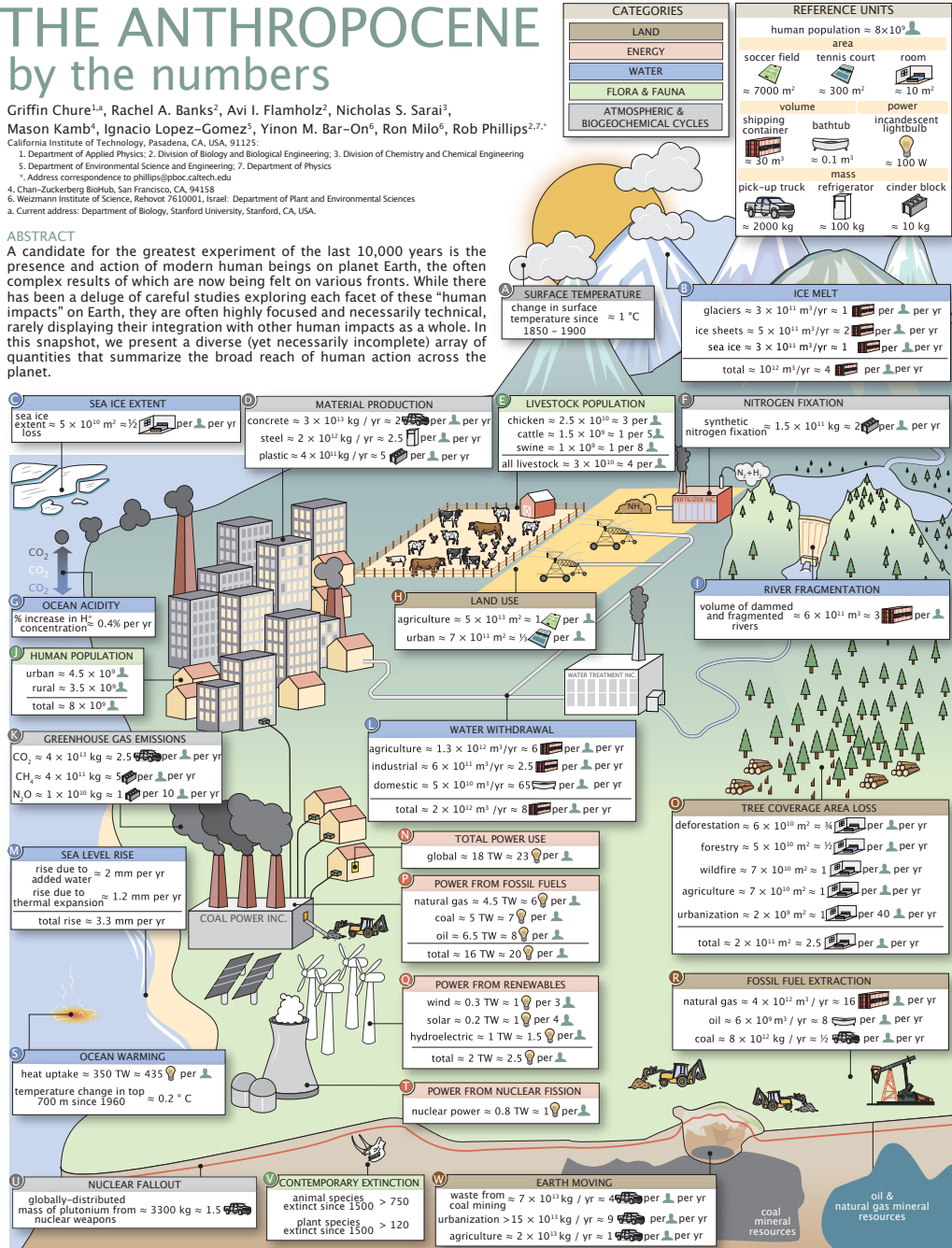


Figure 4.2: Human impacts on the planet and their relevant magnitudes. Relative units and the broad organizational categories are shown in the top-left panels. Source information and contextual comments for each subpanel are presented in the Supplemental Information.

to Malaysia), rely on different natural resources to build infrastructure (e.g. wood vs concrete) and generate power (e.g. nuclear vs coal), and promote different extractive or polluting industries (e.g. lithium mining vs palm oil farming). Some of these regional differences are evident in Figure 3, which summarizes regional breakdowns of several drivers of global human impacts, e.g. livestock populations and greenhouse gas emissions.

Just as impactful human activities like coal power generation and swine farming are more common in some regions than others, likewise the impacts of human activities affect some regions more than others [146]. Figure 4.3 displays a coarse regional breakdown of the numbers from Figure 4.2 for which regional distributions could be determined from the literature. The region definitions used in Figure 4.3 are similar to the definitions set forth by the Food and Agricultural Organization (FAO) of the United Nations, assigning the semi-continental regions of North America, South America, Africa, Europe, Asia, and Oceania. Here, we specify both the total contribution of each region and the per capita value given the population of that region as of the year(s) in which the quantity was measured.

Much as in the case of our Figure 4.2, interesting details emerge naturally from Figure 4.3. For example, Asia dominates global agricultural water withdrawal, using about 62% of the total, while North America takes the lead in industrial water withdrawal. The Organisation for Economic Co-operation and Development (OECD) defines water withdrawal as “freshwater taken from ground or surface water sources, either permanently or temporarily, and conveyed to a place of use.” Interestingly, on a per-capita basis North America withdraws the most water for all uses: agricultural, industrial, and domestic.

North America also emits more CO₂ per capita than any other region, with Oceania and Europe coming second and third, respectively. This disparity can be partially understood by considering the regional distribution of fossil fuel consumption, the dominant source of CO₂ emissions (Fig. 4.3 J). While Asia consumes more than half of total fossil fuel energy, per capita consumption is markedly lower than in North America, Europe, and Oceania (Fig. 4.3 J). Interestingly, the story is different when it comes to methane. Oceania and South America are the largest

emitters of anthropogenic CH_4 , mainly due to a standing population of cattle that rivals that of humans in those regions (Fig. 4.3 D) and produces this potent greenhouse gas through enteric fermentation [25].

Regional disparities are also apparent in the means of energy production. While consuming only 4% of total power, South America generates about 14% of renewable energy (Fig. 4.3 J). Nuclear power generation, on the other hand, is dominated by North America and Europe, while Oceania, which has a single research-grade nuclear reactor, and generates nearly zero nuclear energy (Fig. 4.3 J).

Investigating the causes of forest loss by geographic region likewise highlights interesting differences. At a global level, all drivers of forest loss are comparable in magnitude except for urbanization, which accounts for $\approx 1\%$ of total annual tree cover area loss (Fig. 4.2 O). Despite comparable magnitudes, different drivers of forest loss have different long-term consequences [137]. Forest loss due to wildfires and forestry often result in regrowth, while commodity-driven harvesting and urbanization tend to be drivers of long-lasting deforestation [147, 148]. Central and South America account for about 65% of commodity-driven deforestation (meaning, clearcutting and human-induced fires with no substantial regrowth of tree cover), whereas a majority of forest loss due to shifting agriculture occurs in Africa (where regrowth does occur) (Fig. 4.3 F). Together, wildfires in North America, Russia, China, and South Asia make up nearly 90% of losses due to fire [149]. While urbanization is the smallest driver of tree cover loss globally, it can still have strong impacts at the regional level, perturbing local ecosystems and biodiversity [150, 151].

4.3.4 Timeseries of Human Impacts

When available, the Human Impacts Database includes time series data for each quantity. Just as the regional distributions in impactful human activities help us understand differences between societies and regions, studying the history of these activities highlights recent technological and economic developments that intensify or reduce their impacts. When considering the history of human impacts on the

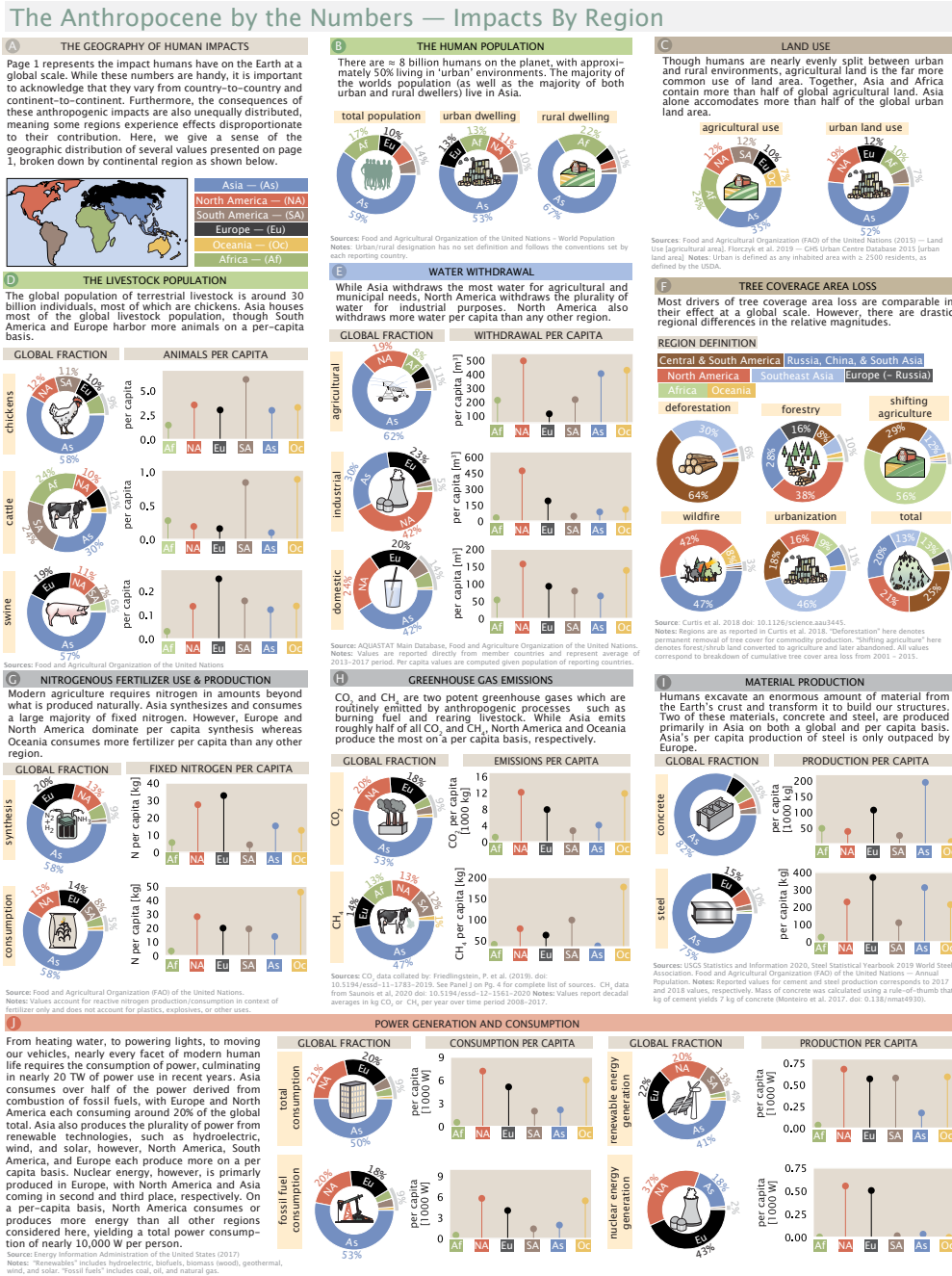


Figure 4.3: Regional distribution of anthropogenic effects. Several quantities from Figure 2 were selected and the relative magnitudes were broken down by subcontinental area (A). Donut charts show the relative contributions of each quantity by region. Ball-and-stick plots show the per capita breakdown of each quantity across geographic regions.

Earth, it is natural to start by considering the growth of the human population over time. As shown in Figure 4.4 A, the global human population grew nearly continually over the past 80 years, with the current population nearing 8 billion. Historically, most of the global human population lives in rural areas (about 70% as of 1950, HuID 93995). Recent decades have been marked by a substantial shift in how humans live globally, with around half of the human population now living in urban or suburban settings ($\approx 55\%$, HuID 93995).

Given the growth of the human population, it is reasonable to consider that human population may be the most natural scale to measure human impacts [152]. To assess this possibility, in Figure 4.4, we plot per-capita impacts over several decades. If impacts are growing in direct proportion with the human population, per capita impacts would be constant over time. Indeed, this is roughly true for per-capita water withdrawals over the last 40 years (Fig. 4.4 B). Deviations from proportionality may indicate important changes in human activities. For example, in recent decades per-capita chicken populations grew by nearly two-fold while per-capita cattle populations shrunk by roughly 25%, reflecting a modest transition away from beef and towards chicken as a source of animal meat in global diets (Fig. 4.4 C, HuIDs 40696 and 79776).

One very visible impact accompanying the shift of the human population to urban environments is the increase in production of anthropogenic mass – materials such as concrete, steel, lumber, and plastics used to build roads, buildings, machines, packaging and other useful human-made items. Since these materials are degraded very slowly, anthropogenic mass has been accumulating over time (Fig. 4.4 D). In addition, the mass of concrete, aggregates like asphalt, and bricks per capita has been increasing since the 1950s (Fig. 4.4 D). Concrete, in particular, has increased from less than 10 tons per person in the 1950s to almost 30 tons per person in the 2010s. This increase in per capita anthropogenic mass means that the increase in production of these materials is outpacing the growth of the human population.

These material production trends have been enabled, in part, by a sustained increase in power generation. As evident from Figure 4.4, total power consumption has been increasing roughly proportionally with the human population. Per capita

consumption has also increased across all generation types, including fossil fuels, hydropower, nuclear, and renewables. The growth among nuclear and renewables has been especially dramatic, and nuclear power now roughly equals hydropower production. Production of crops, aquaculture, and populations of livestock are all likewise correlated with growth in the human population (Fig. 4.4 C and E). The total number of livestock has increased with the human population, primarily due to increasing chicken populations as discussed above. The dominant means of aquatic food production has also shifted over this time: until roughly 1980, nearly all seafood was captured wild, but since then aquaculture has grown to account for roughly one-half of aquatic food production (Fig. 4.4 E, HuID 61233).

Turning our focus to greenhouse gasses, we see that annual anthropogenic CO₂ emissions have been increasing with the population (Fig. 4.4 G). Burning of fossil fuels is the dominant contributor to anthropogenic emissions, and has increased slightly on a per capita basis over the past 60 years. In contrast, as the pace of global deforestation has slowed [153, 154] emissions of CO₂ due to land-use change have decreased per capita. These two trends roughly neutralize each other, leading to little overall change in CO₂ emissions per capita since the 1960s. Akin to CO₂ emissions due to land-use change, CH₄ emissions show a sublinear trend with human population, partially due to a decline in ruminant livestock per capita (Fig. 4.4 C, H).

4.4 Discussion

Quantitative literacy is necessary for understanding in nearly all branches of science. As our collective knowledge of anthropogenic impacts expands, it has become challenging to sift through the literature to collect specific numbers useful for both calculation and communication. We have attempted to reduce this barrier to entry on several fronts. We have canvassed the scientific literature, governmental and international reports to assemble a broad, quantitative picture of how human activities have impacted Earth's atmosphere, oceans, rivers, lands, biota, chemistry, and geology. In doing so, we have created an online, searchable database housing an array of quantities and data that describe different facets of the human-Earth

The Anthropocene by the Numbers — Dynamics of Global Magnitudes

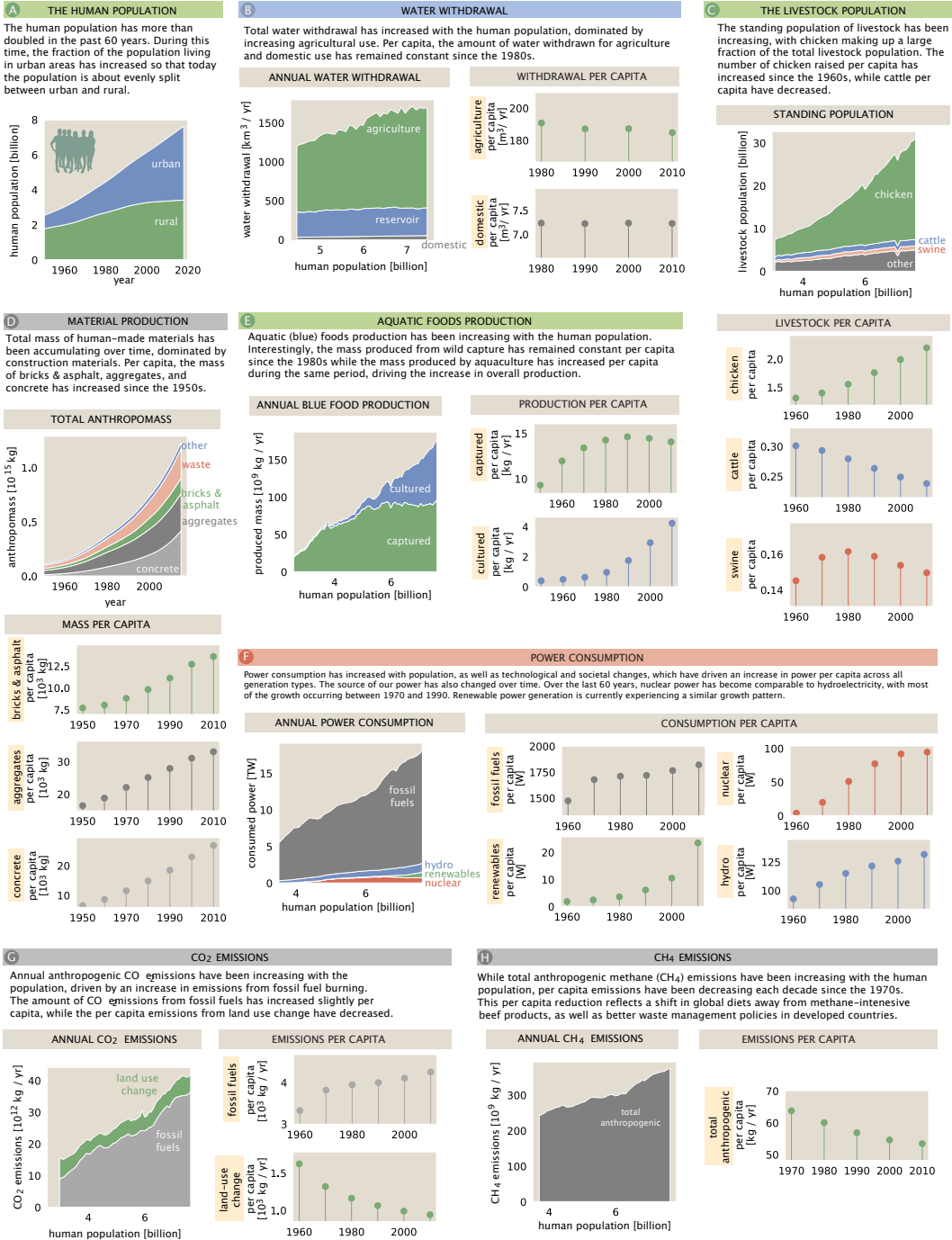


Figure 4.4: Timeseries of human impacts. Several quantities from Figure 2 that had available timeseries data are plotted as a function of either time, for cumulative impacts, or human population. Ball-and-stick plots are decadal per capita average impacts, to show how the per capita impacts have changed over time.

interface. Beyond the database, we have assembled these data into a comprehensive snapshot, with all underlying data, associated uncertainties and referencing housed in the Human Impacts Database. While necessarily incomplete, these resources provide a broad view of the ways in which human activities are impacting the Earth on multiple fronts.

One insight that emerges from considering these diverse human activities together is that they are deeply intertwined and driven by a small number of pivotal factors: the size of the human population, the composition of our diets, and our demand for materials and energy to build and power our increasingly complex and mechanized societies. Understanding the scale of human agriculture, water and power usage provides a framework for understanding most of the numerical gallery presented in Figure 4.2. Perhaps unsurprisingly, we find that feeding the growing human population is a major driver of a large swath of human impacts on Earth, dominating global land (Fig. 4.2 H, HuID: 29582) and water use (Fig. 4.2 L, HuIDs: 84545, 43593, 95345), as well as significantly contributing to tree cover loss (Fig. 4.2 O, HuID: 24388), earth moving (Fig. 4.2 W, HuID: 19415, 41496), and anthropogenic nitrogen fixation (Fig. 4.2 F, HuID: 60580, 61614), to name a few such examples.

It is common in this setting to argue that the bewildering breadth and scale of human impacts should motivate some specific remediation at the global or local scale. We, instead, take a more modest "just the facts" approach. The numbers presented here show that human activities affect our planet to a large degree in many different and incommensurate ways, but they do not provide a roadmap for the future. Rather, we contend that any plans for the future should be made in the light of a comprehensive and quantitative understanding of the interconnected ways in which human activities impact the Earth system globally (Figure 4.2), locally (Figure 4.3), and temporally (Figure 4.4). Achieving such an understanding will require synthesis of broad literature across many disciplines. While the quantities we have chosen to explore are certainly not exhaustive, they represent some of the key axes which frequently drive scientific and public discourse and shape policy across the globe.

Earth is the only habitable planet we know of, so it is crucial to understand how we got here and where we are going. That is, how (and why) have human impacts changed over time? How are they expected to change in the future? For every aspect of human entanglement with the Earth system – from water use to land use, greenhouse gas emissions, mining of precious minerals, and so on – there are excellent studies measuring impacts and predicting their future trajectories. Of particular note are the data-rich and explanatory reports from the Intergovernmental Panel on Climate Change [155, 156] and the efforts towards defining “Planetary Boundaries” [157]. We hope that the Human Impacts Database and the associated resources with this work will aim another lens on the human-Earth system, one that engages the scientific community ultimately helping humanity coexist stably with the only planet we have.

4.5 Supplementary Material

4.5.1 Curation of Human Impacts Database Entries

First of all is to determine if a quantity belongs in the Human Impacts Database. We include any number that quantifies human influence or helps one understand human impacts, such as crop yields. We also include various natural quantities that serve as a reference to compare against human impacts.

Next, we determine if the quantity is reliable. Generally, we consider data from large efforts such as the FAO or IPCC to be reliable sources of values. We are less likely to include values from industry reports, which are more susceptible to bias. Whenever possible, we include multiple sources for a single value, with energy for example, we include data from BP and the EIA. For scientific reports, we check that the methods used to determine the value are clearly stated. If the method is not stated, the value is not included. We will also compare the value to other measurements and check if the source explains any discrepancy between their measurement and others. Generally, if it is the first time a value is measured, we are less likely to include it in the Human Impacts Database as it is more likely to be error prone, and values that are measured repeatedly, such as the surface temperature, are considered more reliable.

Another consideration is if the source reports the uncertainty in their measurement. Many scientific reports will give an assessment of uncertainty, making these values favored. Many numbers come from governmental or intergovernmental bodies come from surveys and are thus self-reported by countries, adding some uncertainty to the data. In these cases, we always state this in the ‘method’ field of the entry in the database.

4.5.2 References and Explanations for Values Reported in Figure 4.2

In this section, we report our extensive and detailed referencing for each and every quantity reported in the subpanels of Figure 4.2 of the main text. As described in the Materials & Methods, each value comes from the manual curation of a piece of scientific, industrial, governmental, or non-governmental organization reports,

articles, or databases. Each value listed here contains information about the original source, the method used to obtain the value, as well as accession identification numbers for the Human Impacts Database (<https://anthroponumbers.org>), listed as HuIDs.

For each value, we attempt to provide an assessment of the uncertainty. For some values, this corresponds to the uncertainty in the measurement or inference as stated in the source material. In cases where a direct assessment of the uncertainty was not clearly presented, we sought other reported values for the same quantity from different data sources to present a range of the values. For others, this uncertainty represents the upper- and lower-bounds of the measurement or estimation.

Each value reported here is prefixed with a symbol representing our confidence in the value. Cases in which an equality (=) symbol is used represents that a measure of the uncertainty is reported in the original data source or represents a range of values from different sources that are tightly constrained (with 2 significant digits). An approximation symbol (\approx) indicates values that we are confident in to within a factor of a few. In some cases, an approximation symbol (\approx) represents a range where the values from different sources differ within three significant digits. In these cases, the ranges are presented as well. Finally, in some cases only a lower-bound for the quantity was able to be determined. These values are indicated by the use of an inequality symbol ($>$).

4.5.2.1 A. Surface Warming

Surface temperature change from the 1850-1900 average \approx 1.0 - 1.4 °C (HuID: 79598, 76539, 12147).

Data Source(s): HadCRUT.4.6 [158], GISTEMP v4 [gistemp, 159] and NOAA-GlobalTemp v5 [160] datasets.

Notes: The global mean surface temperature captures near-surface air temperature over the planet's land and ocean surface. The value reported represents the spread of three estimates and their 95% confidence intervals. Since data for the period 1850-1880 are missing in GISTEMP v4 and NOAA GlobalTemp v5, data

are centered by setting the 1880-1900 mean of all datasets to the HadCRUT.4.6 mean over the same period.

4.5.2.2 B. Annual Ice Melt

$$\text{Glaciers} = (3.0 \pm 1.2) \times 10^{11} \text{ m}^3 / \text{yr} \text{ (HuID: 32459)}$$

Data Sources: Intergovernmental Panel on Climate Change (IPCC) 2019 Special Report on the Ocean and Cryosphere in a Changing Climate. Table 2.A.1 on pp. 199-202.

Notes: Value corresponds to the trend of annual glacial ice volume loss (reported as ice mass loss) from major glacierized regions (2006-2015) based on aggregation of observation methods (original data source: [161, 162]). Ice volume loss was calculated from ice mass loss assuming a standard pure ice density of 920 kg / m³. Uncertainty represents a 95% confidence interval calculated from standard error propagation of the 95% confidence intervals reported in the original sources assuming them to be independent.

$$\text{Ice sheets} = (4.7 \pm 0.4) \times 10^{11} \text{ m}^3 / \text{yr} \text{ (HuIDs: 95798; 93137)}$$

Data Sources: [163]

Notes: Value corresponds to the trends of combined annual ice volume loss (reported as ice mass loss) from the Greenland and Antarctic Ice Sheets (2002-2020) measured by satellite gravimetry. Ice volume loss was calculated from ice mass loss assuming a standard pure ice density of 920 kg / m³. Uncertainty represents one standard deviation and considers only propagation of monthly uncertainties in measurement.

$$\text{Arctic sea ice} = (3.0 \pm 1.0) \times 10^{11} \text{ m}^3 / \text{yr} \text{ (HuID: 89520)}$$

Data Sources: PIOMAS Arctic Sea Ice Volume Reanalysis, Figure 1 of webpage as of October 31, 2020, original data source: [164].

Notes: Value reported corresponds to the trend of annual volume loss from Arctic sea ice (1979-2020). The uncertainty in the trend represents the range in trends calculated from three ice volume determination methods.

4.5.2.3 C. Sea Ice Extent

Extent of loss at yearly maximum cover (September) $\approx 8.4 \times 10^{10} \text{ m}^2 / \text{yr}$ (HuID: 33993)

Extent loss at yearly minimum cover (March) $\approx 4.0 \times 10^{10} \text{ m}^2 / \text{yr}$ (HuID: 87741)

Average annual extent loss = $5.5 \pm 0.2 \times 10^{10} \text{ m}^2 / \text{yr}$ (HuID: 70818)

Data Sources: [165, 166]

Notes: Sea ice extent refers to the area of the sea with $> 15\%$ ice coverage. Annual value corresponds to the linear trend of annually averaged Arctic sea ice extent from 1979-2015 [165] calculated from four different methods. This is in good agreement with the linear trend of annual extent loss calculated by averaging over every month in a given year ($5.5 \times 10^{10} \text{ m}^2 / \text{yr}$ HuID: 66277). The minimum cover extent loss corresponds to the linear trend of Arctic sea ice extent in September from 1979-2020 and the maximum cover extent loss corresponds to the linear trend of sea ice extent in March from 1979-2020. The Antarctic sea ice extent trend is not shown because a significant long-term trend over the satellite observation period is not observed and short-term trends are not yet identifiable.

4.5.2.4 D. Annual Material Production

Concrete production $\approx (2 - 3) \times 10^{13} \text{ kg} / \text{yr}$ (HuID: 25488; 81346; 16995)

Data Sources: [167, 168, 169, 170]

Notes: Concrete is formed when aggregate material is bonded together by hydrated cement. The USGS reports the mass of cement produced in 2019 as $4.1 \times 10^{12} \text{ kg}$. As most cement is used to form concrete, cement production can be used to estimate concrete mass using a multiplicative conversion factor of 7 ([169]). Miller et al. [168] report that the cement, aggregate and water used in concrete in 2012 sum to $2.3 \times 10^{13} \text{ kg}$. Krausmann et al. [170] report an estimated value from 2010 based on a material input, stocks, and outputs model. The value is net annual addition to concrete stocks plus annual waste and recycling to estimate gross production of concrete.

Steel production = $(1.4 - 1.9) \times 10^{12}$ kg / yr (HuID: 51453; 44894; 85981)

Data Sources: [167, 171, 170]

Notes: Crude steel includes stainless steels, carbon steels, and other alloys. The USGS [167] reports the mass of crude steel produced in 2019 as 1900 megatonnes (Mt). The World Steel Association [171] reports a production value of 1869 Mt in 2019. Krausmann et al. [170] report an estimated value from 2010 based on a material input, stocks, and outputs model. The value is net annual addition to steel stocks plus annual waste and recycling to estimate gross production of steel.

Plastic production $\approx 4 \times 10^{11}$ kg / yr (HuID: 97241; 25437)

Data Sources: [172, 170]

Notes: Value represents the approximate sum total global production of plastic fibers and plastic resin during the calendar year of 2015. Comprehensive data about global plastic production is sorely lacking. Geyer et al. [172] draw data from various industry groups to estimate total production of different polymers and additives. Some of the underlying data is not publicly available, and data from financially-interested parties is inherently suspect. Krausmann et al. [170] report an estimated value from 2010 based on a material input, stocks, and outputs model. The value is net annual addition to stocks plus annual waste and end-of-life recycling to estimate gross production of plastics.

4.5.2.5 E. Livestock Population

Chicken standing population $\approx 2.5 \times 10^{10}$ (HuID: 94934)

Cattle standing population $\approx 1.5 \times 10^9$ (HuID: 92006)

Swine standing population $\approx 1 \times 10^9$ (HuID: 21368)

All livestock standing population $\approx 3 \times 10^{10}$ (HuID: 43599)

Data Sources: [173]

Notes: Counts correspond to the estimated standing populations in 2018. Values are reported directly by countries. The FAO uses non-governmental statistical sources to address uncertainty and missing (non-reported) data. Reported values are therefore approximations.

4.5.2.6 F. Annual Synthetic Nitrogen Fixation

Annual mass of synthetically fixed nitrogen $\approx 1.5 \times 10^{11}$ kg N / yr (HuID: 60580; 61614)

Data Sources: [167, 174, 175]

Notes: Ammonia (NH₃) produced globally is compiled by the USGS and IFA from major factories that report output. The USGS estimates the approximate mass of nitrogen in ammonia produced in 2018 as 1.50×10^{11} kg N [167] and the International Fertilizer Association reports a production value of 1.50×10^{11} kg N in 2019 [174]. Nearly all of this mass is produced by the Haber-Bosch process (>96%, [175]). In the United States most of this mass is used for fertilizer, with the remainder being used to synthesize nitrogen-containing chemicals including explosives, plastics, and pharmaceuticals ($\approx 88\%$, [167]).

4.5.2.7 G. Ocean Acidity

Surface ocean [H⁺] ≈ 0.2 parts per billion (HuID: 90472)

Annual change in [H⁺] = $0.36 \pm 0.03\%$ (HuID: 19394)

Data Sources: Figures 1-2 of European Environment Agency report CLIM 043 (2020). Original data source of the report is “Global Mean Sea Water pH” from Copernicus Marine Environment Monitoring Service [176, 177].

Notes: Reported value is calculated from the global average annual change in pH over years 1985-2018. The average oceanic surface pH was ≈ 8.057 in 2018 and decreases annually by ≈ 0.002 units, giving a change in [H⁺] of roughly $10^{-8.055} - 10^{-8.057} \approx 4 \times 10^{-11}$ mol/L or about 0.4% of the global average. [H⁺] is calculated as $10^{-\text{pH}} \approx 10^{-8}$ mol/L or 0.2 parts per billion (ppb), noting that [H₂O] ≈ 55 mol/L. Uncertainty for annual change is the standard error of the mean.

4.5.2.8 H. Land Use

Agriculture $\approx 5 \times 10^{13}$ m² (HuID: 29582) **Data Sources:** [178]

Notes: Agricultural land is defined as all land that is under agricultural management including pastures, meadows, permanent crops, temporary crops, land

under fallow, and land under agricultural structures (such as barns). Reported value corresponds to 2017 estimates by the FAO.

Urban $\approx (6 - 8) \times 10^{11} \text{ m}^2$ (HuID: 41339; 39341)

Data Sources: [179] and Table 3 of [180].

Notes: Urban land area is determined from satellite imagery. An area is determined to be “urban” if the total population is greater than 5,000 and has a minimum population density of 300 people per km^2 . Reported value gives the range of recent measurements of $\approx 6.5 \times 10^{11} \text{ m}^2$ (2015) and $\approx (7.5 \pm 1.5) \times 10^{11} \text{ m}^2$ (2010) from [179] and [180], respectively.

4.5.2.9 I. River Fragmentation

Global fragmented river volume $\approx 6 \times 10^{11} \text{ m}^3$ (HuID: 61661)

Data Sources: [181]

Notes: Value corresponds to the water volume contained in rivers that fall below the connectivity threshold required to classify them as free-flowing. Value considers only rivers with upstream catchment areas greater than 10 km^2 or discharge volumes greater than 0.1 m^3 per second. The ratio of global river volume in disrupted rivers to free-flowing rivers is approximately 0.9. The exact value depends on the cutoff used to define a “free-flowing” river. We direct the reader to the source for thorough detail.

4.5.2.10 J. Human Population

Urban population $\approx 55\%$ (HuID: 93995)

Global population $\approx 7.6 \times 10^9$ people (HuID: 85255)

Data Sources: [182]

Notes: Value for total population in 2018 comes from a combination of direct population reports from country governments as well as inferences of underreported or missing data. The definition of “urban” differs between countries and the data does not distinguish between urban and suburban populations despite substantive

differences between these land uses [183]. As explained by the United Nations population division, "When the definition used in the latest census was not the same as in previous censuses, the data were adjusted whenever possible so as to maintain consistency." Rural population is computed from this fraction along with the total human population, implying that the total population is composed only of "urban" and "rural" communities.

4.5.2.11 K. Greenhouse Gas Emissions

Anthropogenic CO₂ = $(4.25 \pm 0.33) \times 10^{13}$ kg CO₂ / yr (HuID: 24789; 54608; 98043; 60670)

Data Sources: Table 6 of [135]

Notes: Value corresponds to total CO₂ emissions from fossil fuel combustion, industry (predominantly cement production), and land-use change during calendar year 2018. Emissions from land-use change are due to the burning or degradation of plant biomass. In 2018, roughly 1.88×10^{13} kg CO₂ / yr accumulated in the atmosphere, reflecting the balance of emissions and CO₂ uptake by plants and oceans [27]. Uncertainty corresponds to one standard deviation.

Anthropogenic CH₄ = $(3.4 - 3.9) \times 10^{11}$ kg CH₄ / yr (HuID: 96837; 30725)

Data Sources: Table 3 of Saunio *et al.*, 2020 [25].

Notes: Value corresponds to 2008-2017 decadal average mass of CH₄ emissions from anthropogenic sources. Includes emissions from agriculture and landfill, fossil fuels, and burning of biomass and biofuels, but other inventories of anthropogenic methane emissions are also considered. Reported range represents the minimum and maximum estimated emissions from a combination of "bottom-up" and "top-down" models.

Anthropogenic N₂O = $1.1 (+0.6, -0.5) \times 10^{10}$ kg N₂O / yr (HuID: 44575)

Data Sources: Table 1 of Tian H. *et al.*, 2020 [26].

Notes: Value corresponds to annualized N₂O emissions from anthropogenic sources in the years 2007-2016. The value reported in the source is 7.3 [4.2, 11.4] Tg N /

year. This is converted to a mass of N_2O using the fact that $\text{N} \approx 14/22$ of the mass of N_2O . Reported value is mean with the uncertainty bounds (+,-) representing the maximum and minimum values observed in the 2007-2016 time period.

4.5.2.12 L. Water Withdrawal

Agricultural = $1.3 \times 10^{12} \text{ m}^3 / \text{year}$ (HuID: 84545, 43593, 95345)

Industrial = $5.9 \times 10^{11} \text{ m}^3 / \text{year}$ (HuID: 27142)

Domestic = $5.4 \times 10^{10} \text{ m}^3 / \text{year}$ (HuID: 69424)

Total = $(1.7 - 2.2) \times 10^{12} \text{ m}^3 / \text{year}$ (HuID: 27342, 68004)

Data Sources: Figure 1 of Qin *et al.*, 2019 [38], AQUASTAT Database [184].

Notes: “Agricultural” and “total” withdrawal include one value from Qin *et al.* [38] (who reports “consumption”) and one value from the AQUASTAT database [184]. Industrial water withdrawal is from AQUASTAT and domestic withdrawal value is from Qin *et al.* Values in AQUASTAT are self-reported by countries and have missing values from some countries, probably accounting for a few percent underreporting. All values represent water withdrawals. For agricultural and domestic, water withdrawal is assumed to be the same as water consumption, which is reported in Qin *et al.*

4.5.2.13 M. Sea Level Rise

Added water = $1.97 (+0.36, -0.34) \text{ mm} / \text{yr}$ (HuID: 97108)

Thermal expansion = $1.19 (+0.25, -0.24) \text{ mm} / \text{yr}$ (HuID: 97688)

Total observed sea-level rise = $3.35 (+0.47, -0.44) \text{ mm} / \text{yr}$ (HuID: 81373)

Data Sources: Table 1 of [145]. **Notes:** Values correspond to the average global sea level rise for the years 1993 - 2018. “Added water” (barystatic) change includes effects from meltwater from glaciers and ice sheets, added mass from sea-ice discharge, and changes in the amount of terrestrial water storage. Thermal expansion accounts for the volume change of water with increasing temperature. Values for “thermal expansion” and “added water” come from direct observations of ocean temperature and gravimetry/altimetry, respectively. Total sea level rise is

the observed value using a combination of measurement methods. “Other sources” reported in Figure 1 accounts for observed residual sea level rise not attributed to a source in the model. Values in brackets correspond to the upper and lower bounds of the 90% confidence interval.

4.5.2.14 N. Total Power Use

Global power use $\approx 19 - 20$ TW (HuID: 31373; 85317)

Data Sources: [185, 186]

Notes: Value represents the sum of total primary energy consumed from oil, natural gas, coal, and nuclear energy and electricity generated by hydroelectric and other renewables. Value is calculated using annual primary energy consumption as reported in data sources assuming uniform use throughout a year, yielding $\approx 19 - 20$ TW.

4.5.2.15 O. Tree Coverage Area Loss

Commodity-driven deforestation = $(5.7 \pm 1.1) \times 10^{10} \text{ m}^2 / \text{yr}$ (HuID: 96098)

Forestry = $(5.4 \pm 0.8) \times 10^{10} \text{ m}^2 / \text{yr}$ (HuID: 38352)

Urbanization = $(2 \pm 1) \times 10^9 \text{ m}^2 / \text{yr}$ (HuID: 19429)

Shifting agriculture = $(7.5 \pm 0.9) \times 10^{10} \text{ m}^2 / \text{yr}$ (HuID: 24388)

Wildfire = $(7.2 \pm 1.3) \times 10^{10} \text{ m}^2 / \text{yr}$ (HuID: 92221)

Total tree cover area loss $\approx 2 \times 10^{11} \text{ m}^2 / \text{yr}$ (HuID: 78576)

Data Sources: Table 1 of [149] and [187].

Notes: Commodity-driven deforestation is “long-term, permanent, conversion of forest and shrubland to a non-forest land use such as agriculture, mining, or energy infrastructure.” Forestry is defined as large-scale operations occurring within managed forests and tree plantations with evidence of forest regrowth in subsequent years. Urbanization converts forest and shrubland for the expansion and intensification of existing urban centers. Disruption due to “shifting agriculture” is defined as “small- to medium-scale forest and shrubland conversion for agriculture that is later abandoned and followed by subsequent forest regrowth.” Disruption due

to wildfire is “large-scale forest loss resulting from the burning of forest vegetation with no visible human conversion or agricultural activity afterward.” Uncertainty corresponds to the reported 95% confidence interval. Uncertainty is approximate for “urbanization” as the source reports an ambiguous error of “ $\pm <1\%$.”

4.5.2.16 P. Power From Fossil Fuels

Natural gas = 4.5 - 4.8 TW (HuID: 49947; 86175)

Oil = 6.1 - 6.6 TW (HuID: 42121; 39756)

Coal = 5.0 - 5.5 TW (HuID: 10400; 60490)

Total = 16 - 17.0 TW (HuID: 29470; 29109)

Data Sources: [185, 186]

Notes: Values are self-reported by countries. All values from bp Statistical Review correspond to 2019 whereas values from the EIA correspond to 2018 estimates. Reported TW values are computed from primary energy units (e.g. kg coal) assuming uniform use throughout the year. Oil volume includes crude oil, shale oil, oil sands, condensates, and natural gas liquids separate from specific natural gas mining. Natural gas value excludes gas flared or recycled and includes natural gas produced for gas-to-liquids transformation. Coal value includes 2019 value exclusively for solid commercial fuels such as bituminous coal and anthracite, lignite and subbituminous coal, and other solid fuels. This includes coal used directly in power production as well as coal used in coal-to-liquids and coal-to-gas transformations.

4.5.2.17 Q. Power From Renewable Resources

Wind = 0.36 - 0.39 TW (HuID: 30581, 85919)

Solar = 0.18 - 0.20TW (HuID: 99885, 58303)

Hydroelectric = 1.2 - 1.3 TW (HuID: 15765, 50558)

Total = 1.9 - 2.1 TW (HuID: 74571, 20246)

Data Sources: [185, 186]

Notes: Reported values correspond to estimates for the 2019 calendar year for BP

and 2018 for EIA data, except for total renewables, which is from 2017. Renewable resources are defined as wind, geothermal, solar, biomass and waste. Hydroelectric, while presented here, is not defined as a renewable in the BP dataset. All values are reported as input-equivalent energy, meaning the input energy that would have been required if the power was produced by fossil fuels. BP reports that fossil fuel efficiency used to make this conversion was about 40% in 2017.

4.5.2.18 R. Fossil Fuel Extraction

Natural gas volume = $(3.9 - 4.0) \times 10^{12} \text{ m}^3 / \text{yr}$ (HuID: 11468; 20532)

Oil volume = $(5.5 - 5.8) \times 10^9 \text{ m}^3 / \text{yr}$ (HuID: 66789; 97719)

Coal mass = $(7.8 - 8.1) \times 10^{12} \text{ kg} / \text{yr}$ (HuID: 78435; 48928)

Data Sources: [185, 186]

Notes: Oil volume includes crude oil, shale oil, oil sands, condensates, and natural gas liquids separate from specific natural gas mining. Natural gas value excludes gas flared or recycled and includes natural gas produced for gas-to-liquids transformation. Coal value includes solid commercial fuels such as bituminous coal, anthracite, lignite, subbituminous coal, and other solid fuels. All values from bp Statistical Review correspond to 2019 whereas values from the EIA correspond to 2018 estimates.

4.5.2.19 S. Ocean Warming

Heat uptake = 346 51 TW (HuID: 94108)

Upper ocean (0 - 700m) temperature increase since to 1960 = 0.18 -0.20 C (HuID: 69674, 72086)

Data Sources: Table S1 of [30] and [188].

Notes: Heat uptake reported is the average over time period 1992-2015 with 95% confidence intervals. Range of temperatures reported captures the 95% confidence interval of temperature increase for the period 2015-2019 with respect to the 1958-1962 mean. Temperature change is considered in the upper 700 m because sea surface temperatures have high decadal variability and are a poor indicator of ocean warming; see [189].

4.5.2.20 T. Power From Nuclear Fission

Nuclear power \approx 0.79 - 0.89 TW (HuID: 48387; 71725)

Data Sources: [185, 186]

Notes: Values are self-reported by countries and correspond to estimates for the 2019 calendar year from BP and 2018 from EIA. Values are reported as ‘input-equivalent’ energy, meaning the energy that would have been needed to produce a given amount of power if the input were a fossil fuel, which is converted to TW here. This is calculated by multiplying the given power by a conversion factor representing the efficiency of power production by fossil fuels. In 2017, this factor was about 40%.

4.5.2.21 U. Nuclear Fallout

Anthropogenic ^{239}Pu and ^{240}Pu from nuclear weapons \approx $1.4 \cdot 10^{11}$ kg / yr (HuID: 42526)

Data Sources: Table 1 of [190] and [191].

Notes: The approximate mass of Plutonium isotopes ^{239}Pu and ^{240}Pu released into the atmosphere from the \approx 500 above-ground nuclear weapons tests conducted between 1945 and 1980. Naturally occurring ^{239}Pu and ^{240}Pu are rare, meaning that nearly all contemporary labile plutonium comes from human production [192]. The total mass of radionuclides released is \approx 3300 kg with a combined radioactive fallout of \approx 11 PBq. These values do not represent the entire $^{239}+^{240}\text{Pu}$ globally distributed mass as it excludes non-weapons sources.

4.5.2.22 V. Contemporary Extinction

Animal species extinct since 1500 $>$ 750 (HuID: 44641)

Plant species extinct since 1500 $>$ 120 (HuID: 86866)

Data Sources: [193]

Notes: Values correspond to absolute lower-bound count of animal extinctions over the past \approx 520 years. Of the predicted \approx 8 million animal species, the IUCN databases catalogues only \approx 900,000 with only \approx 75,000 being assigned a

conservation status. Representation of plants and fungi is even more sparse with only $\approx 40,000$ and ≈ 285 being assigned a conservation status, respectively. The number of extinct animal species is undoubtedly higher than these reported values, as signified by an inequality symbol ($>$).

4.5.2.23 W. Earth Moving

Waste and overburden from coal mining $\approx 6.5 \times 10^{13}$ kg / yr (HuID: 72899)
Earth moved from urbanization $> 1.4 \times 10^{14}$ kg / yr (HuID: 59640)

Data Sources: Supplementary table 1 of [194].

Notes: Coal mining waste and overburden mass is calculated given commodity-level stripping ratios (mass of overburden/waste per mass of coal resource mined) and reported values of global coal production by type. Urbanization mass is presented as a lower bound estimate of the mass of earth moved from global construction projects. This comes from a conservative estimate that the ratio of the mass of earth moved per mass of cement/concrete used in construction globally is 2:1. This value is highly context dependent and we encourage the reader to read the source material for a more thorough description of this estimation.

Erosion rate from agriculture $> (1.2 - 2.4) \times 10^{13}$ kg / yr (HuID: 19415; 41496)

Data Sources: Pg. 377 of [195] and pg. 21996 of [196].

Notes: Cumulative sediment mass loss over history of human agriculture due to accelerated erosion is estimated to be $\approx 30,000$ Gt. Recent years have an estimated erosion rate ranging from 12 Pg / yr [195] to ≈ 24 Pg / yr [196]. Values come from computational models conditioned on time-resolved measurements of sediment deposition in catchment basins.

Chapter 5

Numeracy for a Human-Dominated Planet: An Order of Magnitude View of the Anthropocene

5.1 Abstract

There are very few human activities that could shift the Earth's orbit around the sun. Given this intuition, it is tempting to assume that human activities are unlikely to affect our planet in lasting ways. This assumption is emphatically false, as humans are now first-order contributors to the carbon and nitrogen cycles, livestock are the dominant constituents of animal biomass on land, and roughly one-third of Earth's land area is under agricultural cultivation. Indeed, the term "Anthropocene" was coined to emphasize that we now live in a human-dominated geological epoch. In this essay, we provide a collection of carefully curated data that characterizes human impacts on the Earth system and its biological occupants. The data we consider includes measurements of impacts on air, water and land and a reflection on the scale of the already observed consequences for humans and the many other living occupants of Earth. Beyond a simple reporting of the data, we provide a collection of order-of-magnitude estimates whose ambition is to give an intuitive feeling for the scale of human impact on natural cycles, terrestrial and oceanic biota, as well as Earth's atmosphere and lithosphere. For example,

reasoning about agriculture helps us estimate the magnitude of water and fertilizer use as well as methane emissions. As we show, performing these calculations endows us with an “extra sense” for the interconnectedness of the Earth-human system and numerical scale of anthropogenic impacts, which we then leverage to trace back a myriad of impacts to the amount of power, food and water that human civilization utilizes. We further highlight the Human Impacts Database, a curated collection of data about the Anthropocene that facilitates quantitative reasoning about human impacts.

5.2 Introduction

Life forms have been shaping the Earth for billions of years, from the composition of the atmosphere to the surface of the land. A famous example is the massive increase in oxygen in the atmosphere following the evolution of organisms capable of photosynthesis, termed the Great Oxidation Event. This rapid change in atmospheric chemistry paved the way for the evolution of aerobic organisms that require oxygen in order to grow and function. More recently, humanity has become a dominant force shaping the Earth’s biosphere, atmosphere, biogeochemistry, and oceans. In recognition of the magnitude of human activity, the current geological epoch has been coined the Anthropocene, meaning that it is characterized by the influence of humanity.

The Anthropocene is estimated to have begun around 1800, the time of the Industrial Revolution. This is when human activities began to have an observable impact on atmospheric carbon dioxide concentrations, due to the large increase in fossil fuel burning to power the Industrial Revolution that transformed human society. However, there are many footprints of human activity on the human-Earth system, which have been increasingly observed and quantified, resulting in a deluge of data that can be difficult to contextualize and gain an intuition for.

Thus, to understand the ways humans are impacting the biological, chemical, and physical composition of the Earth, we take two complementary points of view. First, we find that informed thinking on the topic must be based on quantitative statements, whether about the amount of nitrogen fixed each year or the mass of

seafood produced through aquaculture. As a result, we have created a collection of carefully curated values that represent some of the ways humans are altering the Earth, which are stored in the Human Impacts Database. Second, we argue that the data alone is insufficient to develop intuition and gain an understanding of the system. Instead, order-of-magnitude estimation can be used to reveal the major factors driving human impacts to take the magnitude that they do. For example, how much water is used by humans? Accurate estimation of this requires consideration of the various ways humanity uses water. Further, we find that many of the key human impacts can be connected by a few key quantitative questions: how much land do we use? How much water do we use? And how much power do we consume? By following the threads of these questions, comparing order-of-magnitude estimates to the data, we highlight some of the main driving forces behind human influence on the Earth and the interconnection between human impacts.

5.3 Human Land Use

Humans are not uniformly distributed about the surface of the Earth. The area on which we live and can grow our food is already heavily constrained given that only $\approx 30\%$ of the Earth's surface (around 10^8 km^2) is terrestrial (Fig. 5.1). Yet, we have been able to leverage this land to sustain a tremendous population of nearly 8 billion people which is still growing, anticipated to pass 10 billion by 2050 [197]. These large numbers beg the question, how much of the Earth's land surface is utilized for human purposes? To estimate this value, we can classify human occupied lands into two categories – urban and rural/agricultural – and make some back-of-the-envelope estimates of the corresponding land usage. It should be noted that while there is a clear distinction between urban areas, whose primary uses are housing humans, their workplaces and institutions, and rural lands, whose primary use is agriculture, intermediate land uses like suburbs are harder to distinguish clearly. However, as we shall see, agricultural land use dwarfs all other human uses of land, so we will leave it to others to clearly delineate urban from suburban land uses.

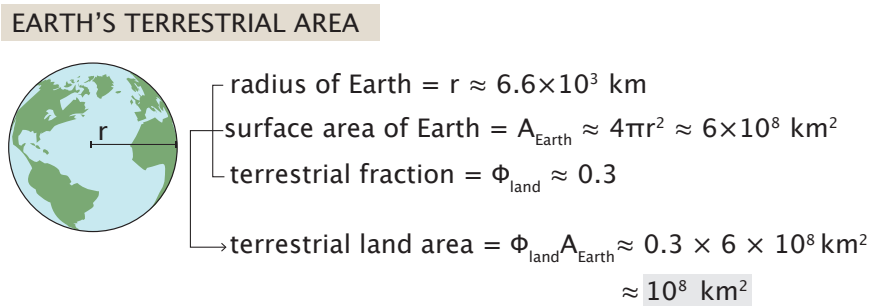


Figure 5.1: Back-of-the-envelope estimate for the Earth’s terrestrial area. We approximate the Earth as a perfect sphere with a radius of $\approx 6,600$ km. Approximately 70% of the Earth’s surface area is water with negligible direct human occupation (such as area covered by boats and other vessels, oil drilling platforms, and docks). Therefore, all human occupation can be approximated as being restrained to the 30% that is terrestrial, which is around 100 million km².

What fraction of the global population of ≈ 8 billion humans lives in urban areas? Until recently, the majority of people on the planet lived outside of cities and urban centers. The balance switched in ≈ 2007 when the rural population approximately equalled the urban population [198]. In the decade-or-so since then, the urban population has expanded, with an approximate 60-40 split in the population between urban and rural areas, respectively (HuID: 93995). While the causes of this recent change are fascinating in their own right, we will forgo that discussion here and make the approximation that the breakdown is approximately equal with a few (denoted mathematically as f) billion people living in rural and urban environments each.

5.3.1 Urban Land Use

We will define “urban” land use as the area of land that is directly occupied by large settlements (think towns, cities, and metropolises) where agriculture is not the primary use of land surface. These urban environments can be quite diverse in terms of their structure. Take, for example, the large cities of Los Angeles and Karachi, located in the United States and Pakistan, respectively. Despite having similar populations of around 14 million in the metropolitan areas, their population densities are remarkably different. Los Angeles is a sprawling metropolis with a

metropolitan area of around 6,000 km². Karachi, on the other hand, is relatively small with an area of around 800 km². These lead to population densities of around 2,000 and 20,000 people per km² for Los Angeles and Karachi, respectively. To generalize these extreme examples to a “typical” urban area, we can compute the geometric mean of these population densities, arriving at a typical density of a few thousand people per km², shown in Fig. 5.2 (A). This is concordant with the definition of an “urban center” as a contiguous area with a total population size of ≥ 5000 , as in a recent study based on satellite-imaging data [179]. The bottom panel of Fig. 5.2 (A) shows the population densities for the top 25 most populous cities in 2015 along with the average density of all urban centers being $\approx 5,000$ people / km² (HuID: 30209) shown as a transparent horizontal line, in good agreement with our estimate.

With this estimate for the urban population density in hand, and the approximation that a few billion people live in urban environments, we can make a simple estimate, diagrammed in Fig. 5.2 (B) that global urban land area is on the order of a million km². Recent measurements of global urbanization using satellite-based imaging approaches yield a similar value of $\approx 700,000$ km² (HuID: 39341, 41339), although other approaches measure a larger total land area of a few million km² [199] (HuID: 87575). While there remain difficulties and disagreement in academic circles of the precise value of urbanized terrestrial area, an estimate of ≈ 1 million km² seems appropriate as a quantitative rule of thumb and reflects that urban areas account for $\leq 1\%$ Earth’s terrestrial area.

5.3.2 Rural Land Use

For the purposes of this estimate, we will make the simplifying approximation that any region that is not urban is rural. By making this delineation, we will say that all farm land (including cropland and livestock grazing pastures) as well as suburban areas with relatively low population densities must house the half of the population that does not live in urban centers. As discussed above, this delineation is imperfect – suburbs are not urban but are not agricultural in nature. However, as we will see, estimates of agricultural land based on human population

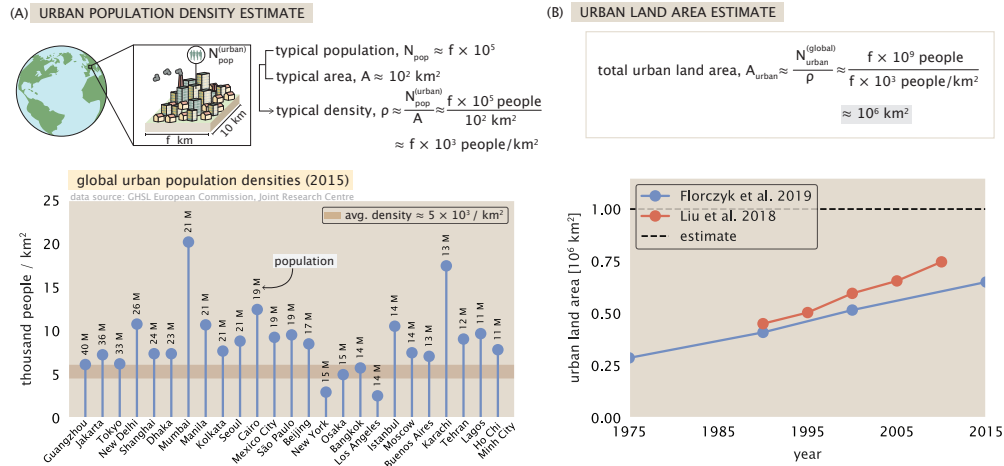


Figure 5.2: Back-of-the-envelope estimate for the extent of urban land use. (A) The density of a typical urban center. We define an urban area as a city or urban center whose primary land-use is not for agricultural processes and has a population $\geq 5,000$ people. Urban environments houses approximately half of the global population, denoted as $\approx f \times 10^9$ people. We assume a typical city size of $\approx 10^2 \text{ km}^2$ with a population of $\approx f \times 10^5$, yielding an approximate population density of $\approx f \times 10^3 \text{ people / km}^2$. The bottom panel shows the population density of the world’s top 25 most populous cities ordered by descending population from left to right (as of 2015). The average population density of all urban areas (not just those centers in the graph) of $\approx 5 \times 10^3 / \text{km}^2$ is shown as a horizontal brown line. Data comes from the European Commission Joint Research Centre [179]. (B) Estimate for the global urban land area. Assuming the estimated population density from (A), the total urban land area is estimated to be $\approx 10^6 \text{ km}^2$. Bottom panel shows the inferred global urban land area from two sources, [179] (blue) and [180] (red), both of which use satellite-based imaging techniques, but different analytical approaches, along with the estimated value shown as a dashed black line.

and agricultural output, are in very good agreement, implying that suburban areas are negligible in comparison.

To estimate the amount of land used for global agriculture, we will consider that typical agricultural areas have a low population density. It is mathematically tempting to assume some extremely low population density, such as 10 person / km^2 . This is fairly representative of the population density of industrial farming in the United States, such as the state of Kansas, which has a rural population density of $\approx 10 \text{ people / km}^2$, shown in Fig. 5.3 (A, bottom right). However, this is not representative of farming globally. Non-industrial farming operations,

such as those in the Gharbia governorate in Egypt ≈ 100 km north of Cairo, have a number of smaller farms operated by a larger local population, resulting in a higher population density while remaining a rural environment, shown in Fig. 5.3 (A, bottom left). We will assume that rural population density is on average an order of magnitude less than that of urban centers, or around 100 people / km². This density aligns with the definition of a rural area by the United States Census Bureau which reads “...rural areas consist of open countryside with population densities less than 500 people per square mile” [200], equivalent to a density of ≈ 200 people / km².

Given this population density and the approximation that there are a few billion people living in rural environments (HuID: 93995), we estimate that the total extent of rural land area, diagrammed in Fig. 5.3 (B), is on the order of a few ten-million km², an order of magnitude above the area occupied in urban areas.

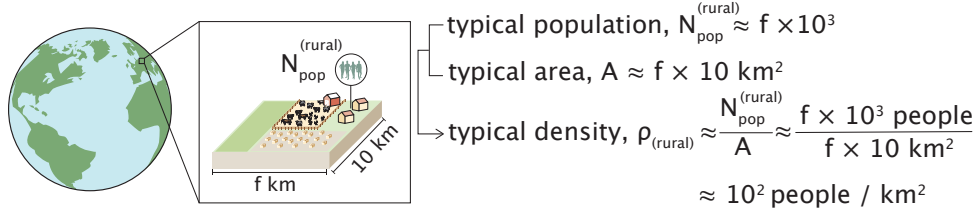
5.3.3 Agricultural Land Use

These estimates thus far suggest that “rural” land area dwarfs the land area used in urban contexts by a factor of ten or more. It is worthwhile to also approach this estimate from considering the food grown on agricultural land rather than the human population that operates it.

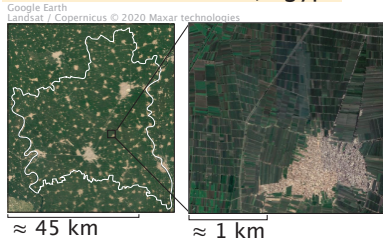
To begin, we will consider only the land area used to produce primary crops that are directly consumed by humans. While diets vary widely across the planet – due to both cultural differences and constraints on what can be easily grown and transported depending on the local environment – it is fair to estimate that a typical human will eat a few kg of plant material per week or a few hundred kg per year. In the spirit of back-of-the-envelope estimation, we will make the approximation that all of the plant-based calories of the typical diet will come from corn, though one can arrive at a similar estimate considering wheat or cereals, which all have similar yields per square meter of crop land.

How much land is needed to produce a few hundred kg of corn per year per person? To estimate this value, we will note that, despite corn fields being famously dense with tall stalks of corn, only 1-2 ears of corn are produced per stalk. The

(A) RURAL POPULATION DENSITY ESTIMATE



Gharbia Governorate, Egypt



land area, $A \approx 2 \times 10^3 \text{ km}^2$

rural population, $N_{pop}^{(rural)} \approx 3 \times 10^6$

density, $\rho \approx \frac{2000 \text{ people}}{1 \text{ km}^2}$

State of Kansas, USA



land area, $A \approx 2 \times 10^5 \text{ km}^2$

rural population, $N_{pop}^{(rural)} \approx 2 \times 10^6$

density, $\rho \approx \frac{10 \text{ people}}{1 \text{ km}^2}$

(B) RURAL LAND AREA ESTIMATE

$$\text{total rural land area, } A_{rural} \approx \frac{N_{rural}^{(global)}}{\rho} \approx \frac{f \times 10^9 \text{ people}}{10^2 \text{ people/km}^2} \approx f \times 10^7 \text{ km}^2$$

Figure 5.3: Estimate for the rural population density and global rural land area. (A) We define a rural area to be all human occupied or managed land area not constituting an urban center. This includes all villages and townships and all agricultural land including cropland and livestock pasture. Approximately half of the global population, denoted as $\approx f \times 10^9$ people, live in a rural environment by this definition. We assume a population density an order-of-magnitude less dense than a typical city center ($\approx 100 \text{ people / km}^2$). The bottom panel shows rural areas in two different locales with significantly different population densities. The Gharbia governorate in Egypt has a high density of $\approx 2,000 \text{ people / km}^2$ while remaining primarily agricultural. A satellite image of the state of Kansas, USA, is shown as an example of farmland with a low density of around 10 people / km^2 . The geometric average of these two extreme examples of rural population densities yields $\approx 100 \text{ people / km}^2$. Images come from Google Earth and data for land area and population come from various national census reports. (B) Taking this population density as being representative of rural areas globally, we arrive at an estimate of $\approx f \times 10^7 \text{ km}^2$ of rural land use.

density of planting also determines the yield of corn per m^2 . In the US, typical planting densities are ≈ 8 stalks per m^2 [201] yielding ≈ 800 g of edible mass per m^2 . However, as we have seen in Fig. 5.3 3(A), industrial farming in the US is not very representative of the world more generally, meaning we should choose a lower corn yield estimate of a few hundred grams per m^2 . The average global yield of maize over the 2010-2018 period is reported to be ≈ 0.5 kg / m^2 (HuID: 87721), in line with our estimate. Similar yields are reported for wheat, rice, and typical cereals (HuIDs: 44291; 62879; 76189).

With these approximations, we can arrive at an estimate for the per capita crop area to be ≈ 1000 m^2 per person (Fig. 5.4 (A)). However, it is important to note that this estimate corresponds only to the area used in primary crop production, meaning crops directly eaten by humans. While the meat-based portion of the typical human diet makes up only $\approx 30\%$ of yearly calories, growing and harvesting meat is a far less efficient use of land, water and fertilizer than growing crops. Recent studies [202, 203] have shown that calorie-for-calorie, the nutrient requirements for livestock rearing dwarf those for plants in most situations. For example, producing 1000 calories of beef requires ≈ 150 m^2 of land (HuID: 40176), whereas harvesting the same number of calories from wheat requires only around 2 m^2 (HuID: 10715). Notably, industrial beef production is a significant outlier in terms of land use. Chicken and pork, for example, require only around ≈ 4 m^2 and 5 m^2 per 1000 calories, respectively (HuIDs: 13245, 75575). Thus, if a typical diet is composed of ≈ 1000 calories of beef a week, land use to produce this would be about a few thousand square meters. Though some cultures consume much more beef than this, others do not eat cows at all. As such, we will assume that this value is typical – i.e. that each person is responsible for a few thousand m^2 of agricultural land per year.

As livestock rearing and primary crop production account for the vast majority of agricultural land area (with fields lying fallow, irrigation canals, and processing facilities occupying a negligible fraction of the total area), we arrive at an estimate that $\approx f \times 10^7$ km^2 of land is used for agricultural purposes. This estimate is in good agreement with the 2018 value of 50 million km^2 (HuID: 29582) reported by

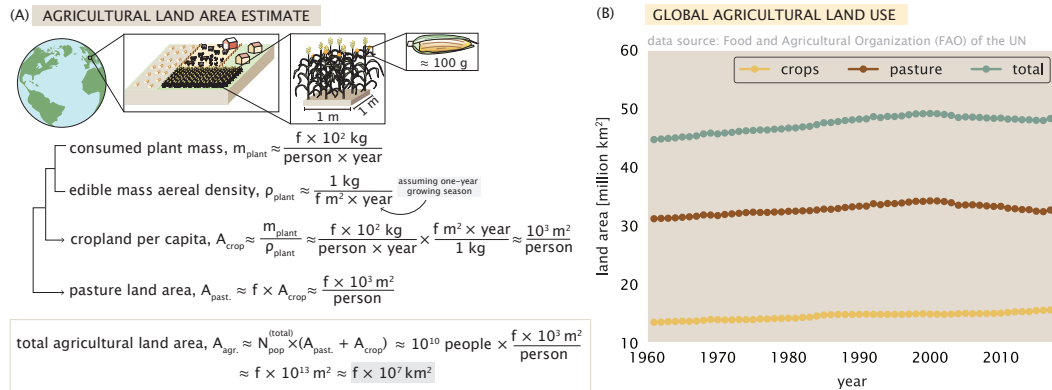


Figure 5.4: Estimating agricultural area from food consumption. (A) We approximate that the typical human consumes a few hundred kg of plant material per year, which we approximate to be coming entirely from corn. With only ≈ 8 ears of corn being generated per square meter of land, we can define an areal density for edible plant mass to be 1 kg per few m^2 of crop land. Together, these values yield an annual per capita estimate for the annual crop land to be around a few thousand m^2 per person. We further assume that the land area needed to rear livestock, while constituting a smaller portion of the typical diet, is a few times the area needed for primary crop production. Combining these estimates and approximations with an estimate of ≈ 8 billion people on Earth, we arrive at an estimate for the total agricultural land to be a few ten-million km^2 , a value in close agreement with that given in Fig. 5.3 (B) which is based solely on population density in rural areas. (B) Data for global agriculture land for crops (yellow) and pasture (brown) from 1960–2018. Data for crops, pasture, and total come from the Food and Agricultural Organization of the United Nations (FAO)[178].

the Food and Agricultural Organization of the United Nations, shown in Fig. 5.4 (B). Despite having a completely different approach, the value of this estimate is the same as that in Fig. 5.3 (B) where we estimated the rural/agricultural land area by considering population densities. Together, these estimates hammer home the fact that agriculture is the primary function of land use by humans.

5.3.4 The Terra Number

These estimates, while undoubtedly simplistic, present a telling feature of modern human society. While city centers and large urban megacities may be a potent mental image of human civilization, urbanization accounts for only a few percent of global human land usage, rural and agricultural use occupy almost all

THE TERRA NUMBER

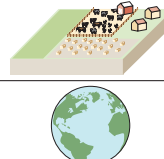
$$T_e = \frac{\text{rural/agricultural land area}}{\text{Earth's terrestrial land area}} \approx \frac{f \times 10^7 \text{ km}^2}{10^8 \text{ km}^2} \approx 0.3$$


Figure 5.5: The Terra Number. The Terra Number is defined as the sum total land use from urban and rural use relative to the total terrestrial area as estimated in Fig. 5.1.

the rest. In assessing the impacts humans have on the planet, it is worthwhile to consider how much of Earth’s surface we utilize. To this end, we define The Terra Number, illustrated in Fig. 5.5 as the fraction of the total terrestrial land surface that falls under the umbrella of urban or rural use. This number, approximately measuring the extent of land under agricultural control, comes to $\approx 30\%$ of the total land surface. It is important to note, however, that this does not reflect the total habitable land surface and includes uninhabitable surfaces such as rugged mountains, glaciers, and arid deserts. While the habitable land area may change with time, the total terrestrial land area is unlikely to drastically change on a human-relevant time scale, making the Terra Number a useful dimensionless quantity to assess the extent of human utilization of the Earth’s land.

5.4 Sizing up Human Livestock Cultivation

As we explored in the Human Land Use section, the land used to rear livestock makes up around 70% of agricultural land (HuID: 22157). This land use estimate did not even account for the number of fish, molluscs, and arthropods that are farmed through aquaculture. In this section, we explore the livestock population in quantitative terms, providing a sense for the total population and biomass of animals that we grow for food. To make these estimates we will take a “fork’s-eye” view. That is, we will start by estimating the average per-capita consumption of particular animal products (e.g. dairy from cows) and then multiply those estimates by the global population to arrive at a global value for total consumption of each animal product annually. As different cultures consume different amounts of various animal products – e.g. per-capita dairy consumption is relatively low in

East Asia – our per-person estimates are not meant to represent a “typical person” but rather an “average person.”

With an estimate for the global mass of consumed animal product in hand, we then estimate how many animals must have been used. In cases of meat, such as chicken and beef, the total number of animals is simple to estimate as the total global mass of meat product divided by the average edible mass of an animal. Animal products like dairy or eggs, which do not entail the slaughter of an animal, require us to take another approach. In these cases, we divide the globally consumed mass by an estimate of the amount of the product (eggs or dairy) that a single animal produces in a year.

Finally, we use these estimates to determine the standing population of each type of animal. Some animals have lifespans of a few years whereas others are shorter-lived with an average lifespan of only a few months until slaughter. To calculate the standing population, we will use a mathematical tool called Little’s Theorem. As an example, let us calculate the student body population of a typical university, as is diagrammed in Fig. 5.6 (A). To use Little’s Theorem, we will assume that the student body population is stationary (does not change over time) and that the rate of graduation (rate of exit) and the rate of matriculation (rate of entry) are equal. Additionally, we can reasonably assume that the typical period of study (residence time) from matriculation to graduation is about 4 years. By multiplying the rate (number of students per year) by the residence time (number of years), we arrive at an estimate of the student body population. The right-hand side of Fig. 5.6 (A) shows the same logic applied to a calculation of the standing chicken population. We will make the approximation that the rate of production (meaning, the rate of slaughter) is balanced by the rate of hatching chicks. With an estimate for the standing populations of different agricultural animals in hand, we can begin to compare population sizes and understand which species are predominantly responsible for the animal-derived portion of the global diet.

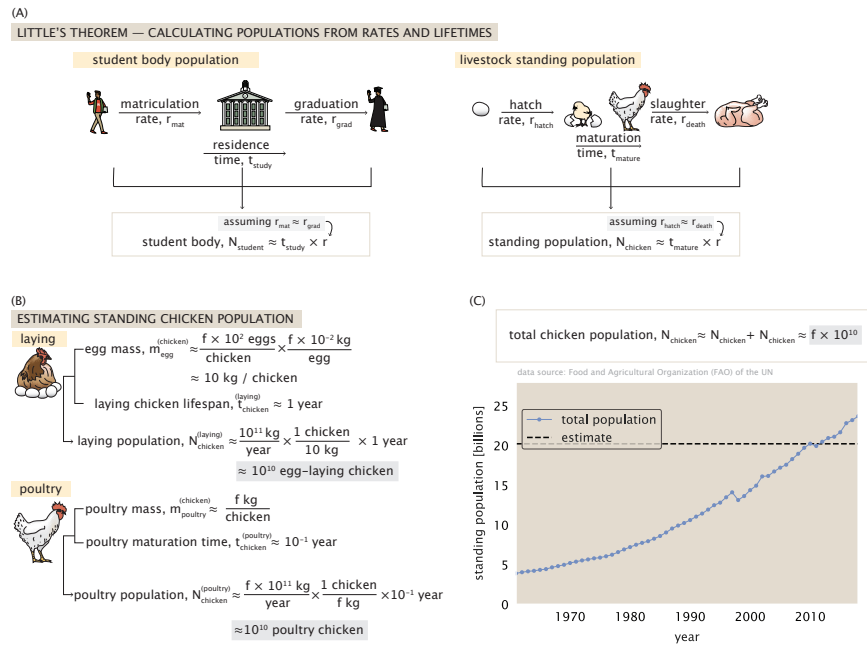


Figure 5.6: Little’s Theorem and calculating standing populations. Little’s theorem can be used to calculate standing populations under steady state, assuming a known entry/exit rate and a residence time. (A) Calculating the size of a university student body (left) given a matriculation/graduation rate and typical time period of study. (B) Calculating the standing population of poultry chicken assuming equivalent hatching/slaughter rates and a typical maturation time.

5.4.1 Chicken

We begin sizing up the global chicken population by considering the mass of poultry and eggs consumed on a yearly basis. Chicken is most-often consumed as a primary meat product, e.g. by eating chicken breasts or legs. A typical chicken breast has a mass of a few hundred grams. We will assume that our average human consumes one or two chicken breasts each week, bringing us to an annual consumed poultry mass of a few tens of kg per year per person. Generalizing to the global population of ≈ 8 billion people yields an estimated global mass of $\approx 10^{11}$ kg per year [Fig. 5.7 (A)]. The Food and Agricultural Organization (FAO) of the United Nations reported that in 2018, around 1.1×10^{11} kg of chicken poultry was produced globally (HuID: 43685) suggesting our “fork’s-eye view” approach is a reasonable one [204].

Unlike poultry, eggs are consumed both in their primary form (such as scrambled or boiled eggs) and also frequently as ingredients in other food products such as baked goods and noodles. As a consequence, it is likely that an average human eats more eggs than would be estimated based on direct consumption alone. With this in mind, we will make the assumption that the average person consumes around a carton of eggs or ≈ 10 eggs per week. While the exact mass of an egg depends on the variety, a typical egg has a mass of ≈ 30 grams or $\approx f \times 10^{-2}$ kg. Thus, the per capita average consumption is a few kg of eggs per year. If 8 billion people eat this way (on average) then we would estimate a global mass of $\approx f \times 10^{10}$ kg of eggs per year, within twofold of the 2018 produced mass of $\approx 7 \times 10^{10}$ kg (HuID: 12755) as reported by the FAO.

Taking these values together, we estimate that slightly more than 1×10^{11} kg of chicken products are consumed per year, comparable to the reported production of $\approx 1.8 \times 10^{11}$ kg chicken product in 2018 (HuID: 12395) as is shown in Fig. 5.7 (B). We note that, in our “fork’s-eye view” approach to these estimates, we are making the approximation that 100% of the food produced is consumed. This is quite certainly a detour from reality as it has recently been estimated that between 20 and 30% of food produced is lost as waste either at the consumer (plate-to-trash) or retail (shelf-to-trash) level [205]. However, as we are making a variety of approximations in considering our average human, this approach correctly estimates the right order of magnitude, often to within a factor of a few. As this is the quantity most relevant to estimating the standing population, we will continue on our journey satisfied with the estimation strategy.

How many individual chickens must humans rear to produce $\approx 2 \times 10^{11}$ kg of chicken product on a yearly basis? To estimate the population, we must first realize there is a difference between the total number of chickens processed each year and the total standing population. While it is useful to know the total number of processed animals, the standing population represents the instantaneous resource burden that rearing livestock imposes on the planet. We will use Little’s Theorem to estimate the standing population of chickens, starting with egg laying chickens. A typical egg-laying chicken can produce around 10 eggs per day for a ten-month

period, or around 300 eggs (about 10 kg) per year. The egg-laying chicken has a lifespan of about 1 year, thus, using Little's Theorem, we arrive at an estimate for the standing population of egg laying chickens of about 10 billion (Fig. 5.7 (C, top)).

What about poultry chicken? Similarly to the egg-laying variety, poultry chicken have been bred to be extremely productive. Over the span of ≈ 60 years, industrial breeding has dramatically increased the size of poultry chickens. At time of slaughter (around 2 months of age), a modern poultry chicken will carry nearly 5 kg of edible mass. In the 1950's a 2 month old chicken had a nearly fivefold smaller mass of only ≈ 1 kg [206]. Thus, with a modern mass of a few kilograms and a lifespan of around ≈ 0.1 year, we can compute the standing population of poultry chicken to be $\approx 10^{10}$ individuals [Fig. 5.7 (C, bottom)], a population comparable to that of egg-laying chicken.

Under the fair approximation that the egg and poultry producing chicken populations have negligible overlap, we can arrive at a simple estimate that the global standing population of chickens is on the order of $\approx f \times 10^{10}$ individuals [Fig. 5.7 (D, top)]. This is concordant with the value of $\approx 2.3 \times 10^{10}$ (HuID: 94934) in 2018 reported by the FAO [Fig. 5.7 (D, bottom), outnumbering the global human population by a factor of ≈ 3 .

5.4.2 Cattle

We now turn our attention to cattle – another major source of animal protein in the form of meat (beef) and dairy (milk). As we did for egg-laying and meat-producing chickens, we will consider the populations of beef and dairy cows to be distinct and non-overlapping. Like eggs, dairy can be consumed in many forms: liquid milk, cultured products such as yogurts and cheeses, or as an ingredient in a wide array of food products. Regardless of the form, all dairy products (originating from cows) begin their life as whole liquid milk. We estimate that the average person consumes a few liters of milk per week, either as liquid milk or embodied in some derived milk product. Furthermore, though milk is an emulsion of fat, sugars, and water, we will assume that it has a density around 1 kg per liter, like water.

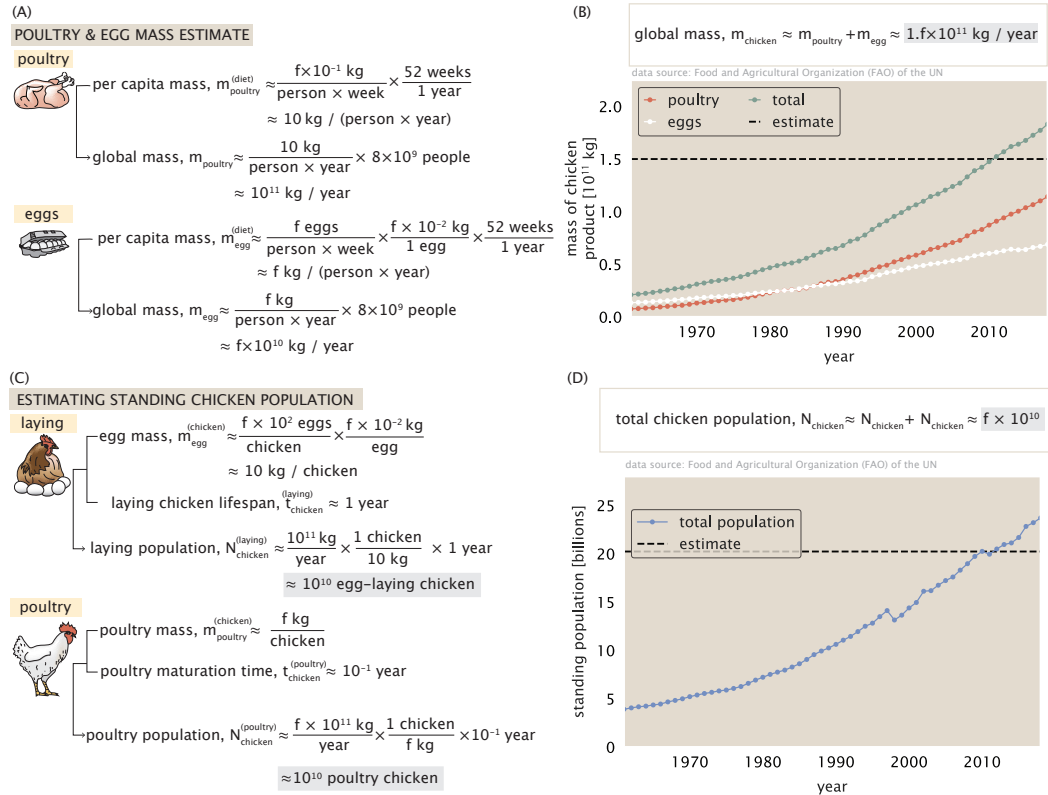


Figure 5.7: Estimates for the annual production of eggs and poultry from chicken. (A) An estimate for the total annual produced mass of chicken product, which we approximate to be composed of eggs and poultry. (B) A comparison with data from the FAO of the total produced mass of poultry (red), eggs (white), and sum total. The estimated value of $f \times 10^{11} \text{ kg} / \text{year}$ from (A) is shown as a dashed black line. (C) Estimates of the standing population of egg-laying (top) and poultry (bottom) chickens. (D) The global standing chicken population over the time period of 1961–2018 as reported by the FAO. The sum of the egg-laying and An estimate for the total annual produced mass of chicken product, which we approximate to be composed of eggs and poultry. Data from the FAO.

This assumption is almost exactly right even for whole milk (3-4% fat content). Given these values, we estimate that around 10^{12} kg of milk is consumed globally each year, shown in Fig. 5.8 (A), a value within a factor of a few from the FAO 2018 reported value of $\approx 6.8 \times 10^{11}$ kg / year [204] (HuID: 25015).

Unlike dairy products, beef is almost exclusively consumed in its primary form. In the same spirit as our estimate for poultry, we will assume that the average human consumes a 1 kg steak every few weeks, yielding yearly consumption of ≈ 10 kg per person per year. Generalizing this to the global population yields an estimate of $\approx 10^{11}$ kg of beef per year, as shown in Fig. 5.8 (A), again within a factor of a few of the FAO 2018 value of $\approx 7 \times 10^{10}$ kg [204] (HuID: 82278).

From these estimates, we can see that production of milk exceeds that of beef production by a factor of ≈ 10 , meaning that total annual mass of cow products (i.e. milk products + beef products) is approximately equal to the total milk product $\approx 10^{12}$ kg per year. This approximation is not too far from reality, as is seen in Figure 5.8 (B) where the FAO reported global production of milk, beef, and the sum total is beginning in 1961. The data for 2018 gives a total of $\approx 7.5 \times 10^{11}$ kg / year (HuID: 66626), within a factor of a few from our estimate for milk production.

How many cattle must humans rear to produce all this beef and dairy? Again, we turn to Little's Theorem to calculate the standing population. Centuries of domestication and artificial selection have yielded dairy cows capable of producing impressive volumes of milk. A high-yield dairy cow can produce around ≈ 30 L of milk or ≈ 30 kg per day. This amounts to an astounding $\approx 10^4$ kg of milk per year per cow, which is perhaps tenfold more than in non-industrial settings [207]. Given this value, we estimate that on the order of $\approx 10^8$ cows are required to satisfy the global demand for dairy. As dairy cows are typically milked for a few years, we apply Little's Theorem to estimate the standing dairy cattle population to be $\approx f \times 10^8$, as can be seen in Fig. 5.8 (C, top).

Beef cattle have been selectively bred to produce an impressive amount of meat, with typical angus beef cattle being on the order of $\approx 10^3$ kg. This "on-the-hoof" mass is not entirely edible. After slaughter, the carcass is butchered to separate

inedible parts (e.g. bones) and recover a net edible mass of $\approx f \times 10^2$ kg of beef per cow [208]. Thus, to supply a global consumption of 10^{11} kg per year, approximately 3×10^8 cattle must be slaughtered. Beef cattle are raised until they reach a mass sufficient for the economics of slaughter and sale. As for dairy cattle, beef cattle thus have lifetimes of a few years. Applying Little's Theorem once again, we arrive at an estimate of the standing beef cattle population that is on the order of one billion individuals, shown in Fig. 5.8 (C, bottom).

While a portion of beef production comes from dairy cows that can no longer produce milk, we will assume that the overlap between these populations is negligible, allowing us to estimate the total cattle standing population to be a little bit more than 1 billion [Fig. 5.8 (D, top)]. In 2018, the FAO reported a global cattle standing population of around 1.5 billion (HuID: 92006), meaning there is about 1 cow for every 5 people on Earth [173]. Once again, our simple estimate matches reality to within a factor of a few, as shown in Fig. 5.8 (D, bottom).

5.4.3 Fish

As we have seen, we collectively raise a population of chicken and cattle that is similar to the total human population (both within a factor of 10 of the ≈ 7 billion humans on Earth). Some find these numbers staggering, while others are unsurprised, having seen the extent of human agricultural activity out the window during a recent plane trip or drive. However, our extensive dependence on fishing and aquaculture for food is perhaps less visible. In this section, we explore the extent of seafood consumption. Seafood is a major source of protein globally and is becoming more prominent as a source of dietary protein over the past several decades [209]. Though 'seafood' generally has a broad definition and can include non-animal foodstuffs like seaweed, we here restrict our definition to aquatic animals. Returning to the 'average' human, let us consider a weekly diet where the majority of protein comes from seafood. A typical adult might consume 50-100 grams of seafood per day or a few hundred grams per week. As there are many forms of seafood – fish, shrimp, oysters, etc. – we will speak in broad terms and estimate that average weekly seafood consumption is on the order of a

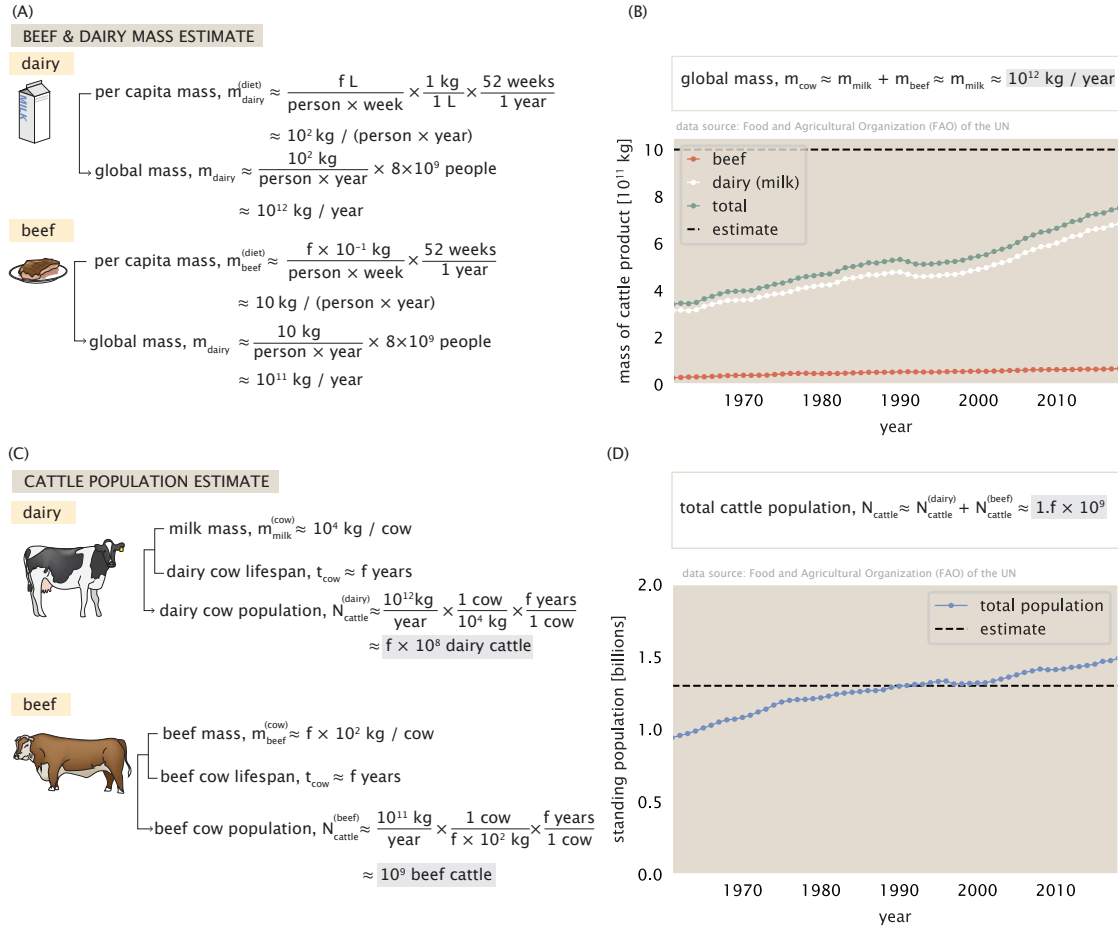


Figure 5.8: Estimates for the annual production of beef and dairy and their corresponding cattle populations. (A) An estimate for the total annual produced mass of cow product, which we approximate to be beef and dairy. (B) A comparison with data from the FAO of the total beef (red), milk (white), and sum total cow product is shown on the right-hand side. (C) An estimate for the standing population of dairy (top) and beef (bottom) cattle assuming a typical lifespan of $\approx f$ years from birth to slaughter. (D) A comparison of the estimate from (B), shown as a dashed black line, with data from the FAO of the total cattle population (blue) [173].

few hundred grams a week or a few tens of kg a year. As shown in Fig. 5.9 (A, top), with a population of 8 billion, we quickly arrive at an estimate that annual seafood consumption is on the order of $\approx f \times 10^{11}$ kg per year. Recent data from the Fisheries Division at the FAO has reported that the total mass of seafood produced in 2018 was $\approx 1.8 \times 10^{12}$ kg (HuID: 41132, 46460), within a factor of a few from our estimate.

However, unlike the vast majority of our terrestrial protein sources, only around 50% of this annual mass is aquacultural (HuID: 61233, 35535) with the remainder being caught from the wild, shown in Fig. 5.9 (A, bottom). Therefore, in order to assess the number of animals reared by humans for food, we will consider only the half of seafood mass that is raised under aquaculture. Figure 5.9 (B) illustrates the top three categories of aquatic animals that are farmed for food production – oysters/clams, carp, and shrimp – and their estimated standing populations.

The above estimates reveal that humanity is collectively responsible for raising huge populations of animals for food. As is shown in Fig. 5.10 (A), the majority of these animals ($\approx 3 \times 10^{11}$ individuals) are relatively small aquatic animals grown under aquaculture (e.g. shrimp, clams and carp). In addition to this enormous population of aquatic animals, humanity is responsible for a standing population of $\approx 3 \times 10^{10}$ terrestrial animals that are raised for food (e.g. chickens, pigs and cows raised for meat, eggs and dairy). This enormous population of aquatic livestock is primarily composed of oysters, clams, and mussels which together sum to around 250 billion, or nearly 45 times the population of humans. Shrimp and prawns hold second place in this popularity contest with around 65 billion individuals, outnumbering humans by a factor of 10. Other aquatic species, such as carp and tilapia, are also sizable but are two orders of magnitude less abundant than their shelled compatriots. Turning our gaze to terrestrial livestock, chickens dominate the population with around 25 billion individuals, around three times the human population. Interestingly, while an icon of the agricultural industry, cattle come in third place when it comes to numbers, only slightly outnumbering swine.

While it is illuminating to consider population size, it is also worth keeping in mind the total biomass of each of these categories of livestock. As seen in Fig. 5.10

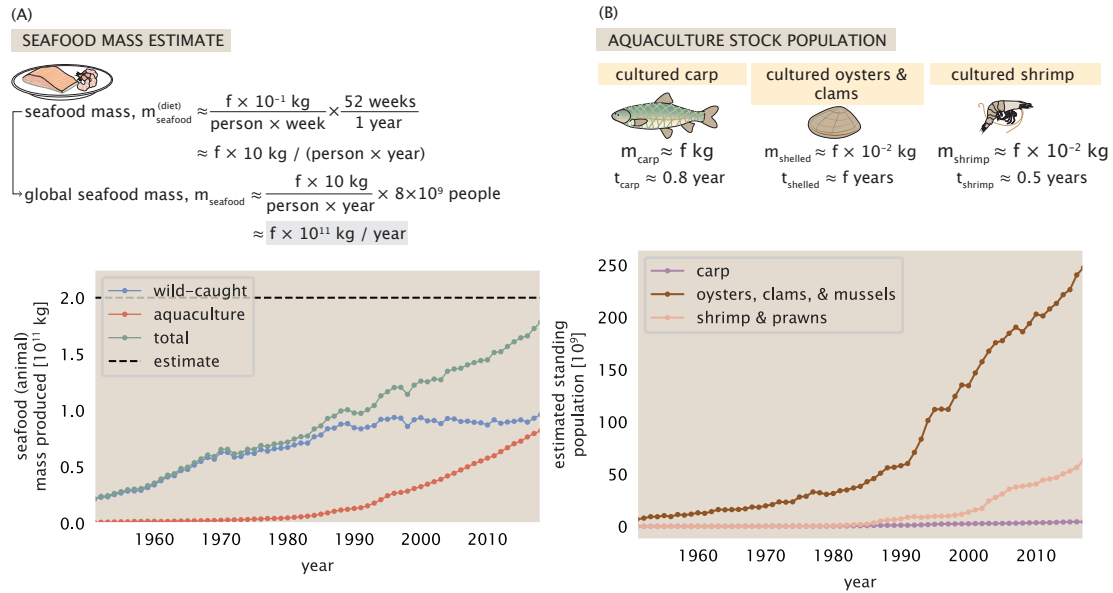


Figure 5.9: Estimate for annual seafood consumption standing populations of aquacultural animals. (A) An order of magnitude estimate for the extent of global seafood consumption assuming a few hundred grams are consumed per person per week on average. Bottom plot shows data from the FAO fishery division on the mass of seafood (animals only) produced globally (green). Blue and red curves correspond to the seafood mass produced via wild-catch and aquaculture fisheries, respectively. (B) Estimates of the standing population of carp, oysters and other shelled animals, and shrimp. Values for masses and lifespans are taken from the FAO fishery division cultured species fact sheets. Bottom plot shows the estimated standing populations of each of the aforementioned species plotted as a function of time.

(B), when we consider the total biomass of each category, the picture is reversed. In a mass-centric view, cultured aquatic organisms are negligible in comparison to the biomass of terrestrial livestock. Of the latter, cattle dominate all forms of livestock, weighing in at close to 1.52×10^{12} kg. In terms of mass, aquatic organisms rank 5th after swine, chicken, and horses/camels, with no single aquatic species commanding a majority of biomass. This result makes intuitive sense, of course, because cows, pigs and chickens are larger and more massive than the most abundant aquatic livestock (clams, oysters, shrimp and muscles) by a factor of 10-100 at least.

How do we put these findings in context? While there are many compelling

reasons to critically assess standing populations of various kinds of livestock, we believe that examining the distribution of livestock biomass most fairly represents how resources – namely, carbon, water, fertilizer and land area – are apportioned to the animals that we rear for food. It has recently been estimated that there is around $\approx 17 \times 10^{12}$ kg animal biomass (both terrestrial and aquatic) on the planet, including sizable populations of arthropods and worms [124]. Our estimate of $\approx 2 \times 10^{12}$ kg of animal biomass in livestock means that just over 10% of all animal biomass on the planet, both terrestrial and aquatic, is under the direct care of humans for the production of food. As can be seen in Figure 5.10 (B, left), nearly the entirety of this livestock biomass is terrestrial. Thus, it is reasonable to compare the total biomass of animal livestock ($\approx 2 \times 10^{12}$ kg) to their wild counterparts. The most recent estimate places the total biomass of wild mammals and birds to be approximately 7×10^{10} kg.

5.4.4 The Barnyard and Fishery Numbers

Thus, we define The Barnyard Number as the ratio of the total biomass of terrestrial livestock (which is nearly exclusively mammals and birds) to the total biomass of wild mammals and birds on Earth. Shown schematically in Fig. 5.11 (A), The Barnyard Number is approximately ≈ 30 . This remarkable ratio captures the extent to which human livestock has become the dominant proportion of mammals and birds (in terms of mass) on the planet, effectively replacing the biomass void left after the Quaternary Megafauna Extinction event.

We can compute an analogous quantity for aquatic organisms, defining The Fishery Number, diagrammed in Figure 5.11 (B) as the total mass of aquatic animal livestock to the wild animal biomass in marine environments. As illustrated in Fig. 5.11 (B), the total aquatic livestock sums to a modest $\approx 9 \times 10^{10}$ kg whereas recent work has estimated the animal biomass in the oceans to be nearly 13×10^{12} kg [124, 31], bringing The Fishery Number to a factor of ≈ 0.007 . This small factor reveals that only a tiny fraction of the marine world is farmed with the vast majority of the animal biomass remaining “wild.” While human seafood production is a very small share of the total aquatic biomass, human fishing practices have

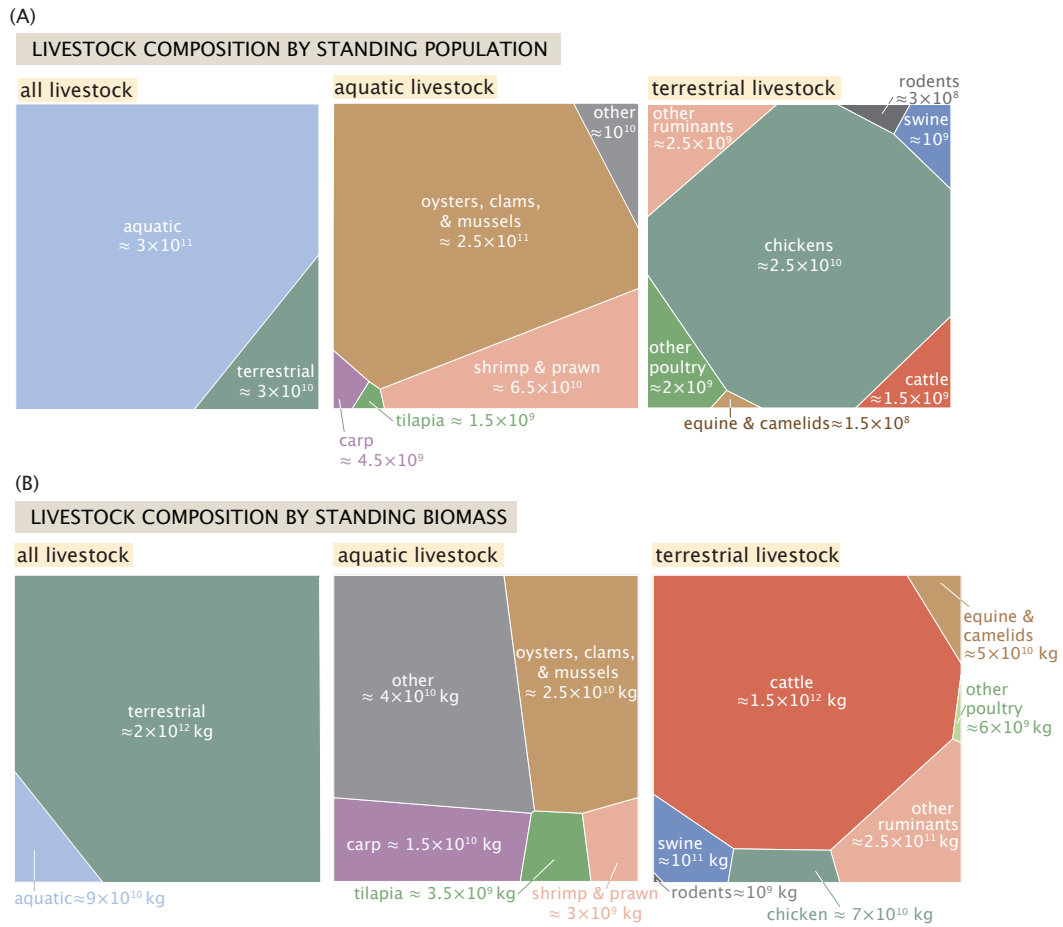
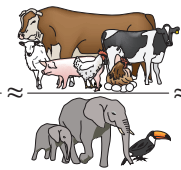


Figure 5.10: Livestock population and biomass distribution by type and animal category. (A) Standing population breakdown (left) aquaculture versus terrestrial livestock, (middle) aquatic livestock, and (right) terrestrial livestock. (B) Estimate for the standing biomass of each livestock category and corresponding animal. Plots were created using the voronoi proteomap tool via bionic-vis.

(A)

THE BARNYARD NUMBER

$$B_y = \frac{\text{terrestrial livestock animal biomass}}{\text{terrestrial wild animal biomass}} \approx \frac{2 \times 10^{12} \text{ kg}}{7 \times 10^{10} \text{ kg}} \approx 30$$


(B)

THE FISHERY NUMBER

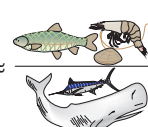
$$F_y = \frac{\text{aquatic livestock animal biomass}}{\text{aquatic wild animal biomass}} \approx \frac{9 \times 10^{10} \text{ kg}}{13 \times 10^{12} \text{ kg}} \approx 0.007$$


Figure 5.11: The Barnyard and Fishery Numbers. (A) The Barnyard Numbers defined as the total terrestrial animal biomass in livestock relative to the total animal biomass in wild mammal and bird populations. The value of the numerator comes from this work whereas the denominator comes from other recent work [124]. (B) The Fishery Numbers is defined as the fraction of aquatic animal biomass under the umbrella of human aquaculture relative to the wild-animal biomass in marine environments. The value of the numerator comes from this work whereas the denominator is estimated by others [124, 31]. Note the denominator in this quantity also includes marine animal livestock, however the wild-animal fraction dwarfs the cultured biomass, making the latter a negligible component of the denominator. Denominators in both (A) and (B) were calculated from recent estimates of the total carbon mass and was converted to wet mass assuming that carbon represents $\approx 50\%$ of the dry mass and the total dry mass is $\approx 30\%$ of the total wetmass of the organism.

had substantial effects on aquatic ecosystems. For example, about one-third of fisheries are overfished, meaning that the biomass is being depleted year by year.

5.5 Human Water Use

Water is a requirement for life on planet Earth as we know it. Hence, it is not hyperbole to state that water is one of (if not the) most valuable resources on the planet, and the pursuit of reliable water sources has sculpted modern human society. The hunt for reliable sources of potable water has dictated our experience on the planet throughout history, with river systems such as the Nile, the Amazon, the Mississippi, and the Tigris and Euphrates playing central roles in defining

where we live and how we utilize land. As time marches forward, our population, our power consumption, and the land area that we occupy and utilize significantly increases. As a result of this multi-dimensional growth, the amount of water required to sustain our growth has increased, posing new challenges as well as opportunities for how water can be withdrawn, used, and ultimately recycled.

How much water do humans use in modern society? Beyond the direct consumption of water through drinking and eating, modern agriculture and industrial operations require quantities of water that rival the flow rate of some of our planet's greatest rivers. Agriculture, for example, requires water for the growth of crops as well as for rearing livestock. Industrial processes use water in a variety of ways, such as in the generation of electricity, cooling of high-temperature machinery, and as a general purpose solvent. In this section, we will quantitatively explore the various ways in which humans use water, ultimately changing its phase, purity, or temperature before it is returned back into the natural water cycle. In considering these effects, however, we must define what it means to “use” water.

The scientific literature discusses two kinds of water “use” – water withdrawal and water consumption. Water withdrawal considers the volume of water that is temporarily removed from Earth's water cycle, used for a particular purpose, and is ultimately returned to the water cycle in a liquid state. For example, water used for thermoelectric power production is withdrawn from a natural source, converted to steam to spin turbines and generate electricity, is condensed back to a liquid state, and is returned directly back to its source. This is not to say, however, that such processes have no effect on the water cycle: water is often returned at a higher temperature or with a level of purity different than when it was originally withdrawn, both of which can have important ecological consequences [210]. Conversely, water consumption refers to water that is removed from the source and is ultimately dispersed through evaporation, transpiration by plants, or direct consumption by humans and livestock. For example, water consumed by humans is ultimately released back into the water cycle primarily as a sewage, which requires further treatment and purification before it is restored to the “natural” water cycle.

In the coming sections, we will estimate water withdrawal and consumption for a variety of processes. Estimated volumes are often unimaginably large, measured in trillions of liters.

5.5.1 Domestic Water Use

While humanity uses water in many ways, the most obvious and relatable is direct consumption. While it obviously depends on your level of activity and your local temperature, it is recommended to drink around 8 glasses of water a day. With a typical glass being a few hundred mL, this brings us to a conservative estimate that the average person consumes a few liters of water a day directly. Assuming you're usually not dehydrated, this rule-of-thumb tells us that the $\approx 8 \times 10^9$ people on the planet collectively drink $\approx f \times 10^{13}$ L of water every year, shown in Fig. 5.12 (A).

Of course, drinking is not the only way people use water domestically. In addition to cooking, the average person uses a considerable volume of water for sanitation. This includes water used in flushing toilets, bathing, and doing laundry. While the amount of water used for each of these categories is highly variable from region to region, we can make some fast-and-loose approximations for the volume used on average for each category.

For example, let us assume that toilets use only a few L of water for every flush [211]. If a typical person flushes a few times a day, then each person will flush about 10 L of water a day. We can take a similar approach to estimating the water volume used for bathing. For bathing, we will make the bold approximation that the typical person takes a 10 minute shower a few times a week. Typical high-pressure showerheads will deliver water at a rate of a few L per minute, thus we arrive at a simple estimate that the typical person uses on average ≈ 10 L of water per day for bathing.

Finally, we consider the water volume used for laundry services. While washing machines are becoming significantly more efficient in both their power and water use [212], a typical washing machine uses ≈ 102 L of water per load of laundry. Assuming that the typical person washes their clothes once every two weeks,

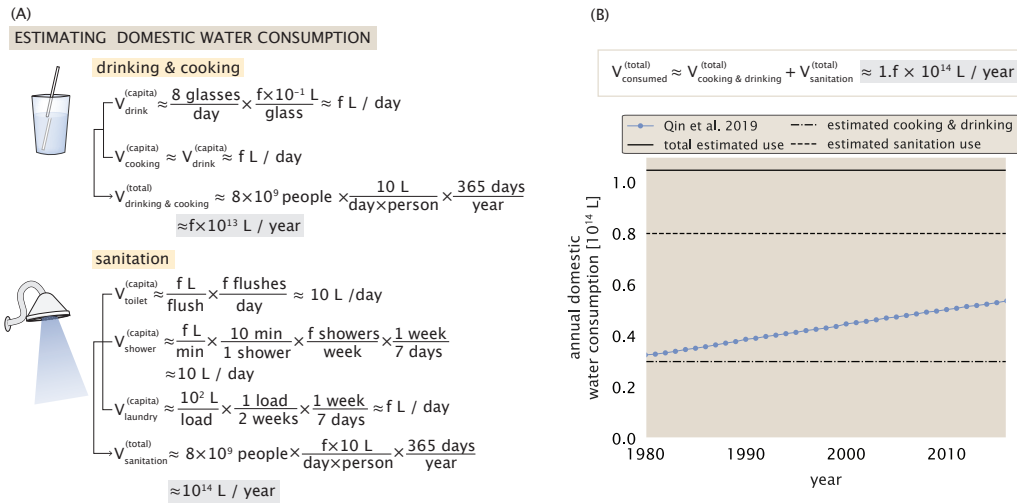


Figure 5.12: Estimating domestic water use. We define domestic water use to include all water used for drinking, cooking, and for sanitation. (A) An estimate for the volume of water used for drinking and cooking (top) assuming that approximately equal volumes are used for each activity. The bottom estimate is for sanitation, which we define as including lavatory water, water used for bathing, and water used for cleaning clothes. (B) An estimate for the total volume of water used for domestic purposes [sum of estimates in (A)] and a comparison with data ([38], blue glyphs). Solid (—), dashed (---), and dash-dot (-.-) horizontal lines correspond to the total estimated water use, estimated water use for sanitation, and water used for cooking and drinking, respectively.

washing will use a relatively modest daily volume of a few L, comparable to our estimate for drinking water.

Pulling these estimates together, we arrive at a total daily water budget of $\approx f \times 10$ L of water used for sanitation each day. Scaling this value up to the global population of $\approx 8 \times 10^9$ people yields an estimate of around 10^{14} L of water a year. This volume is a few times larger than the volume estimated to be directly consumed via drinking and cooking. Thus, as schematized in Fig. 5.12 (B), we arrive at our final estimate for the volume of water used for domestic purposes to be on the order of $\approx 1.1 \times 10^{14}$ L per year. This is reasonably close to the reported value of consumed water which in 2016 was found to be $\approx 6 \times 10^{13}$ L per year [38] (HuID: 69424).

5.5.2 Agricultural Water Use

While the volume of water used for domestic purposes is sizable, it is remarkably small compared to the water used in one of humanity's most significant activities – agriculture. As discussed in the Barnyard Number section, humans collectively manage a standing population of ≈ 30 billion terrestrial animals as livestock, together weighing in at a whopping $\approx 2 \times 10^{12}$ kg. This is about an order of magnitude more mass than the 8×10^9 humans (HuID: 85255), which weigh in at about $f \times 10^{11}$ kg. Like humans, these livestock drink a sizable volume of water every day to stay hydrated and grow. To estimate the volume of water directly consumed by livestock, we will make the assumption that humans and livestock require approximately the same amount of water per kilogram of body weight. In Fig. 5.12 (A), we relied on the old adage that each human must consume ≈ 8 glasses or a few L of water every day. Assuming that the typical person is ≈ 50 kg, we arrive at a rule of thumb that around 0.1 L of water is needed per kg per day. Combined with our estimate of $\approx 2 \times 10^{12}$ kg of livestock on the planet, we quickly arrive at an estimate that livestock collectively consume $\approx 10^{14}$ L of water every year, as seen in Fig. 5.13 (A, top), a few times more than the volume of water drunk by humans.

What volume is needed to grow the crops that feed both us and our livestock? A portion of water for crops comes from natural sources, such as rainfall. Thus, to account for human water use, we consider only the water purposefully brought in and used for irrigation. What fraction of the global agricultural land is actually irrigated? As discussed in the Terra Number section, humans manage $\approx f \times 10^{13}$ m² of the Earth's surface solely for agriculture (HuID: 29582). According to the Food and Agriculture Organization (FAO) of the United Nations, only about 10% of this land, or $\approx f \times 10^{12}$ m², is irrigated while the rest is rainfed. Then, we need to know how much water is used per m². We assume that around one L of water is used for each m² of irrigated land each day, which is enough to cover the land area by one mm of water. Thus, as shown in Fig. 5.13 (A, bottom), the total amount of water used for irrigation is $\approx 10^{15}$ L per year.

How do these relatively naive estimates stack up to the actual data? Together,

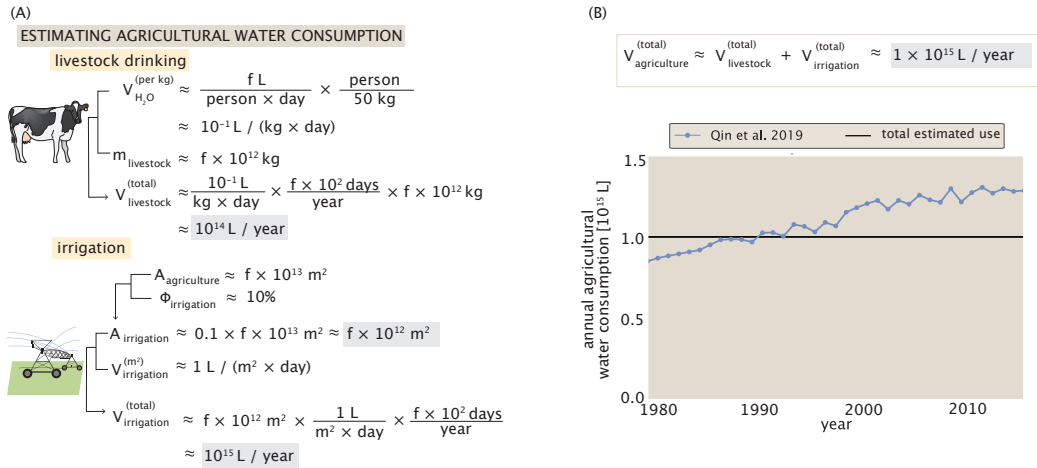


Figure 5.13: Estimating agricultural water use. We define agricultural water use as the water consumed by livestock drinking and water used for irrigating crops. (A) An estimate of the amount of water consumed by livestock (top). We assume that livestock and humans drink the same amount of water per kg. The bottom estimate is the amount of water used for irrigating crops. (B) A sum of estimated agricultural water use, which is dominated by the water used for irrigation and a comparison with the data from [38].

our estimates suggest that $\approx 10^{15}$ L of water are used each year for agriculture, with irrigation dominating the sum. This is in agreement with the amount of agricultural water use reported by [38] of about $0.9 - 1.3 \times 10^{15}$ L, shown in Fig. 5.13 (B) (HuID: 43593). This data also shows that the amount of agricultural water consumption has risen since 1980. While land used for agriculture has not changed much, historically rainfed land has been transitioned to irrigation, thus increasing the water use for irrigation [210]. Like the water used for domestic purposes, agricultural water use we estimate here qualifies as consumed, leaving its natural source and entering the bodies of livestock, plants, and the ground.

5.5.3 Thermoelectric Water Use

Thus far, our exploration of the extent of human water use has focused on direct use cases. However, water is used for purposes other than keeping ourselves clean or growing our food (plant and animal alike). A major and perhaps non-obvious use of water is in the production of around 10^{20} J of electricity per year from thermoelectric power plants. In the late 18th century, James Watt patented

a design for a rudimentary steam engine – a machine that coupled the formation of steam to the generation of mechanical work. Roughly 200 years of technological advances in how steam-driven engines operate (and the type of work they can perform) have culminated in electricity generators like the one schematized in Fig.5.14 (A). While these technological advancements have greatly improved the efficiency and power output of steam engines, their basic operating principle is the same as that proposed by James Watt over two centuries ago. Steam is formed by boiling water using the heat from combusting fuel sources such as gas or coal and the resulting pressure performs mechanical work. In modern generators, this mechanical work is in the form of rotation of a steam turbine, shown in the middle of Fig. 5.14 (A), which produces electricity via electromagnetic induction. Once the steam has spun the turbine, it is then condensed back to a liquid state, ready to be boiled again. This cycle of vaporization and condensation is repeated over and over again with a negligible volume of water lost through leaks, suggesting that the volume of water that is boiled is not the dominant avenue of water use.

To condense the steam back to liquid water, it must be cooled by transferring heat to a colder medium. This is mediated by a condenser system, which one can think of as an array of tubing filled with near room temperature water that is surrounded by the steam that has already been used to spin the turbine. There are many ways in which these condenser systems can be built and operated, but the basic principle is the same—cool water comes in, absorbs heat from the steam which then condenses back to liquid water, and leaves the cooling water in the condenser at a higher temperature. A common mechanism of this condensing process is a “once-through” cooling system where water from an external source (such as a lake, river, sea, or ocean) is passed into the condenser, absorbs heat, and is returned to the source. Many of the early industrial power stations were constructed using a once-through system, though the past few decades have seen the development of other water-based cooling systems that avoid returning higher temperature water to the environment entirely.

What volume of water is needed to cool the steam? To begin, we will make the approximation that all thermal power plants use a once-through cooling system,

generating $\approx 10^{20}$ J of electricity each year. Of course, as these turbines are not 100% efficient, the energy input to the system must be larger than the electric output. The turbines in industrial power plants are objective marvels of modern engineering, and boast efficiencies as high as 70%, meaning that for every joule of energy input to the turbine system, around 0.7 joules are produced as electricity. Assuming that 70% efficiency [213] is characteristic of modern thermoelectric power plants, we can compute the input energy of the steam to be $\approx 1.4 \times 10^{20}$ J, as is shown in Fig. 5.14 (B). The difference between the input and output is the remaining energy of the steam that needs to be dissipated by the cooling system, coming out to $\approx 0.3 \times 10^{20}$ J.

The volume of water needed to dissipate this energy is dependent on one key parameter – the maximum tolerable temperature difference between the water flowing in and the water flowing out and back into the environment. This temperature difference is strongly regulated by various environmental protection agencies and places an inherent limit on the minimal volume of water that must be used to perform the cooling. While the temperature difference is dependent on the location of the power generator, the time of year, and even the time of day, we make the estimate that the acceptable difference is a few degrees Celsius, a value in range with guidelines from the US and European Union [214].

The volume of water needed to absorb this energy and maintain a few degree temperature difference is defined by a chemical property termed the specific heat. The specific heat, shown as C_{H_2O} in Fig.5.14 (B), defines the amount of energy needed to raise 1 kg of water by 1 °C, at a standard temperature and pressure, which we will approximate to be $\approx 4.2 \times 10^3$ J per kilogram per °C. With ≈ 3 °C as our self-imposed temperature difference, a total energy of $\approx 1.3 \times 10^{19}$ J of residual heat to be dissipated, and assuming a density of 1 kg per L of water, Fig. 5.14 (B) shows that an estimated volume of around 3×10^{15} L of water would be passed through the condenser system.

Another way of representing this volume is to compute the withdrawal coefficient for once-through cooling, the volume of water needed for cooling for every 10^6 joules of electricity produced. With our estimate of 3×10^{15} L needed for

cooling for the generation of 10^{20} J, we estimate a withdrawal coefficient of $f \times 10$ L / 10^6 J as diagrammed in Fig. 5.14 (C, top). This value is in line with the reported withdrawal coefficients for a variety of once-through cooling technologies for different fuel sources, as shown in Fig. 5.14 (C, bottom) and tabulated by Macknick et al. [215], suggesting that our relatively simplistic estimate captures the major features of once-through cooling systems.

Due to the ecological impact of once-through cooling and the need to build power plants in environments where water can be scarce, more recent power plants employ different types of cooling systems that permit recycling of the cooling water through the condenser or systems which require little to no water at all. Such recirculation systems (such as cooling towers, cooling ponds, and dry-cooling systems which use air as the exchange medium) withdraw significantly less water than once-through systems. This can be seen in Fig. 5.15 (A, left) where the withdrawal coefficient (L of water per 10^6 J of produced electricity) is shown for different fuel sources and cooling systems. From these data, we can approximate the withdrawal coefficients to be $\approx f \times 10$ L / 10^6 J for once-through cooling and ≈ 10 L / 10^6 J for the various types of recirculation cooling technologies. In the United States, around 30% of the annual thermoelectric power production comes from once-through cooling [216]. More recent work for the global thermoelectric power estimates that $\approx 50\%$ is cooled using once-through systems [38] with the remaining using recirculation systems. In Fig. 5.15 (B), we use the reported withdrawal coefficients from Fig. 5.15 (A) to estimate the total volume of water withdrawn for both varieties of cooling technologies, totalling to $\approx 2 \times 10^{15}$ L of water per year, a volume 20 times than the discharge volume of Niagara Falls.

Our estimates, however, consider only the volume of water withdrawn to cool steam. Water consumed in the generation of thermoelectricity refers to the water evaporated and thus is distinct from the water withdrawn. The consumption coefficients, defined as the volume of water evaporated per 10^6 J, are shown in Fig. 5.15 (C) for different cooling systems and fuel sources. Unlike the withdrawal coefficient, once-through systems ultimately consume less water (≈ 0.2 L / 10^6 J) compared to that of recirculation systems (≈ 0.5 L / 10^6 J), though it is highly

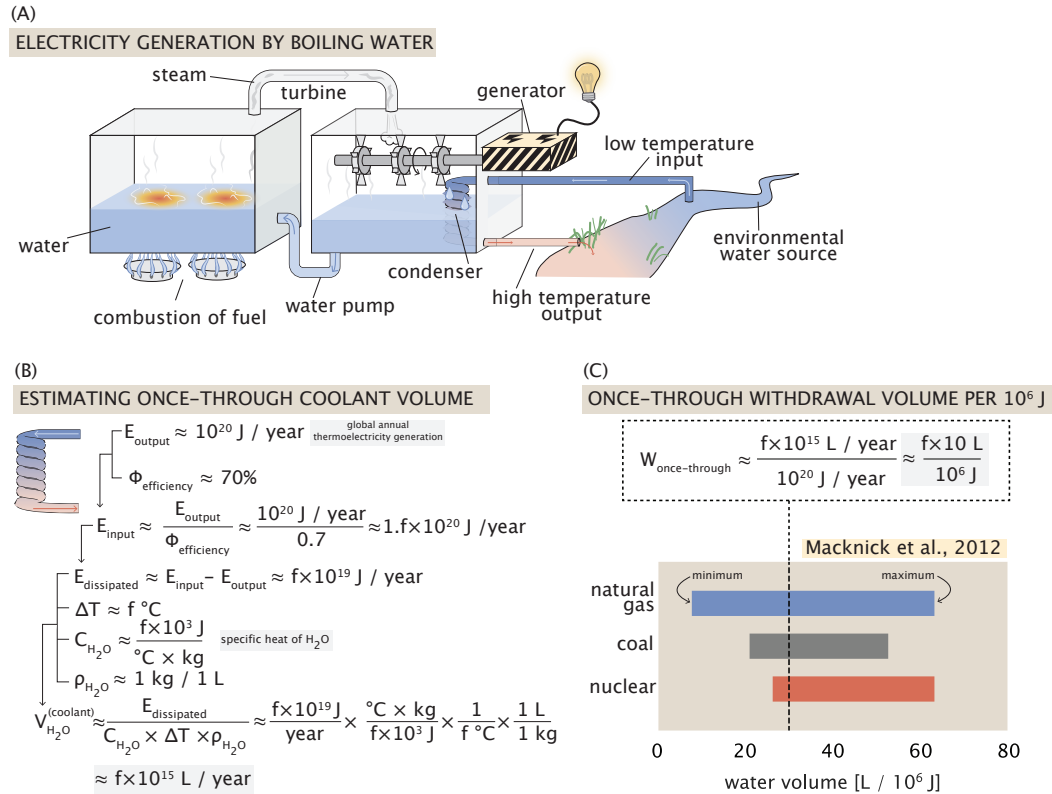


Figure 5.14: Thermoelectric power generation and estimates for required water volumes. (A) A schematic of a simplified thermoelectric power plant. A volume of water is boiled by the combustion of fuel (such as coal or natural gas) or by nuclear fission. This generates steam which is pressurized and directed to a turbine system where a pressure differential causes the turbine to rotate, generating electricity. After spinning the turbine, the steam is condensed and returned to the boiler. The steam condensation is mediated by a cooling system where external water (such as from a lake, sea, or river) is passed through a condenser system in the steam filled turbine chamber. As steam condenses, the latent heat is transferred to the coolant water. In a “once-through” cooling system, the water in the condenser is returned back to the natural source, though at a higher temperature. (B) An estimate for the volume of water needed to condense the steam if all thermoelectric power was cooled via a once-through system. (C) The amount of water needed to be withdrawn for cooling per 10^6 J of energy produced (termed the withdrawal coefficient) is estimated to be a $f \times 10 \text{ L} / 10^6 \text{ J}$. This estimate is inline with reported values as reported by [215]. Width of the bars represent the minimum and maximum withdrawal coefficients reported for a variety of different once-through cooling systems for different fuel sources. Vertical dashed line shows the estimated coefficient.

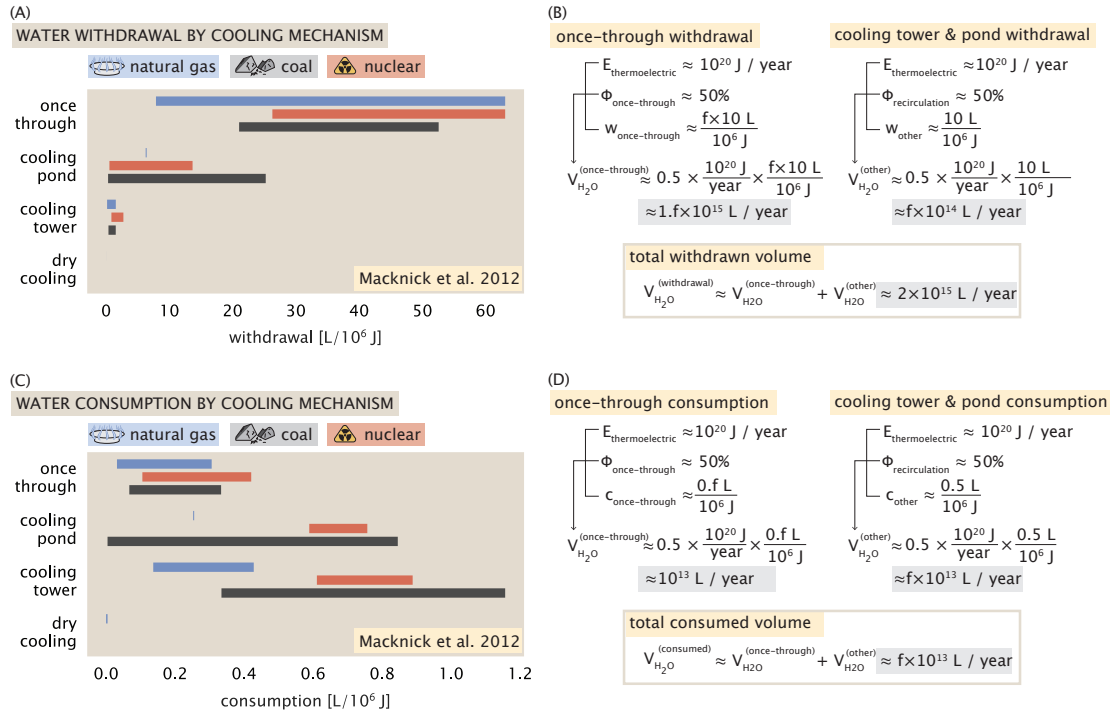


Figure 5.15: Water withdrawal and consumption volumes for different thermoelectric cooling systems. (A) Data on the range of withdrawal coefficients for different fuel sources and cooling mechanisms. (B) Estimates on the illustrate the volume of water withdrawal for once-through and recirculated cooling mechanisms. (C) Data on the range of consumption coefficients for different fuel sources and cooling mechanisms. (D) Estimates correspond to the total volume of water consumed for once-through and recirculating cooling mechanisms. The width of each bar of (A) and (B) corresponds to the maximum and minimum reported values for withdrawal and consumption of each fuel source and cooling technology, respectively. Data corresponds to that reported by [215].

dependent on the type of fuel used. Again using an approximation that $\approx 50\%$ of the global electricity production uses once-through cooling systems, we can estimate the total volume of water consumed. This estimate is shown in Fig. 5.15 (D), culminating in a consumed volume of $\approx f \times 10^{13} \text{ L a year}$.

These estimates reveal that the global drive for electricity requires a sizable volume of water to be withdrawn, even if it is ultimately returned. Of course, thermoelectric generators are not the only way to generate electricity from water.

5.5.4 Hydroelectric Water Use

Hydroelectricity has been growing as a source of power globally. It is enticing because it does not consume materials like fuels, coal, and gas. But rather, hydroelectricity works by harnessing the energy of falling water. As anyone who has ever dropped something knows, gravity brings that object to crash on the ground or whatever surface is underneath it. The potential energy of a falling object can simply be calculated from its mass, the acceleration due to gravity, and the height that it will fall. An object farther from the ground or with a larger mass has a larger potential energy than one closer to the ground or with a smaller mass. In the case of hydroelectricity, water falls down a dam, gaining kinetic energy, hits a turbine, and causes it to turn and generate electricity, diagrammed in Fig. 5.16 (A).

At first glance, it may seem as though water used to generate hydroelectric power does not have much of an impact on the hydrological cycle since the water continues flowing down the river after turning the turbine. However, it is important to consider this water usage for a couple reasons. Firstly, although much of the water is returned to the river, building a dam has a huge impact on the ecosystem by inhibiting fish migration or causing flooding of the nearby region, for example. Further, these dams have very large reservoirs of water, which increase the surface area of the water, thereby increasing evaporation. This water is considered consumed by hydroelectric plants since the evaporated water is not immediately returned to the river system.

How much water is used for hydroelectricity? We can estimate this number from the total amount of energy produced by hydroelectricity, about 10^{19} J per year (HUID: 50558) [185]. For reference, this is about 15% of the total energy produced globally in a year. As noted above, this energy comes from falling water, which we can easily calculate if we know the height that the water falls. The average height of all dams used for hydroelectricity is $\approx f \times 10$ m [217]. However, we make the simplifying assumption that most hydroelectric power is produced by a handful of large, tall dams, with an average height $\approx 10^2$ m. Along with the density of water at standard temperature and pressure ≈ 1 kg / L, we estimate

that one of these dams produces 10^3 J / L. As diagrammed in Fig. 5.16 (A). In order to generate 10^{19} J, a whopping volume of $\approx 10^{16}$ L of water is required.

Determining how much water actually flows through these dams in a year is a bit tricky, however. We can calculate the total reservoir capacity of all dams used to generate hydroelectricity, which amounts to $\approx 5 \times 10^{15}$ L (HUID: 94828), a factor of a few below the amount of water we estimated is required. This implies that each year, hydroelectric dams pass more water than they store, by a factor of a few fold! For this to be true, the flow of water into the reservoirs must be large enough to offset the discharge rate. To determine if this seems reasonable, let us consider a real-life example. The Three Gorges dam in China is the largest producer of hydroelectric power in the world. Reaching completion in 2012, it reliably generates $\approx f \times 10^{17}$ J of electricity each year, is about 200 m tall, and has a reservoir capacity of 4×10^{13} L [218]. Using the same strategy as above, we can estimate the amount of water required to flow through the Three Gorges to generate $f \times 10^{17}$ J. As shown in Figure 5.16 (B), we estimate that the Three Gorges dam requires 10^{14} L per year, a volume a few times larger than the reported reservoir capacity. Repeating this exercise for the Kolnbrein Dam in Austria and the Grand Coulee dam in Washington State yields the same conclusion, that they must flow a few times their reservoir volume through the dam each year.

As a final sanity check, we can estimate the amount of water that would flow over the Three Gorges dam if the river were allowed to flow at its natural rate. The flow rate of the Yangtze river at the dam is $\approx 30,000$ m³/s [219]. At this rate, about 3×10^{14} L of water would flow over the dam in a year. However, the dam only operates about half the year, so about 10^{14} L is a better estimate, a quantity similar to our estimated flow volume through the dam. This suggests that, while relatively simplistic, our estimate of water used based on the energy of falling water is a fair approximation of the amount of water harnessed to generate hydroelectric power.

While the amount of water harnessed for hydroelectricity is massive, this water is neither withdrawn nor consumed since it continues on its way down the river after spinning the turbine. However, as noted above, hydroelectric dams consume

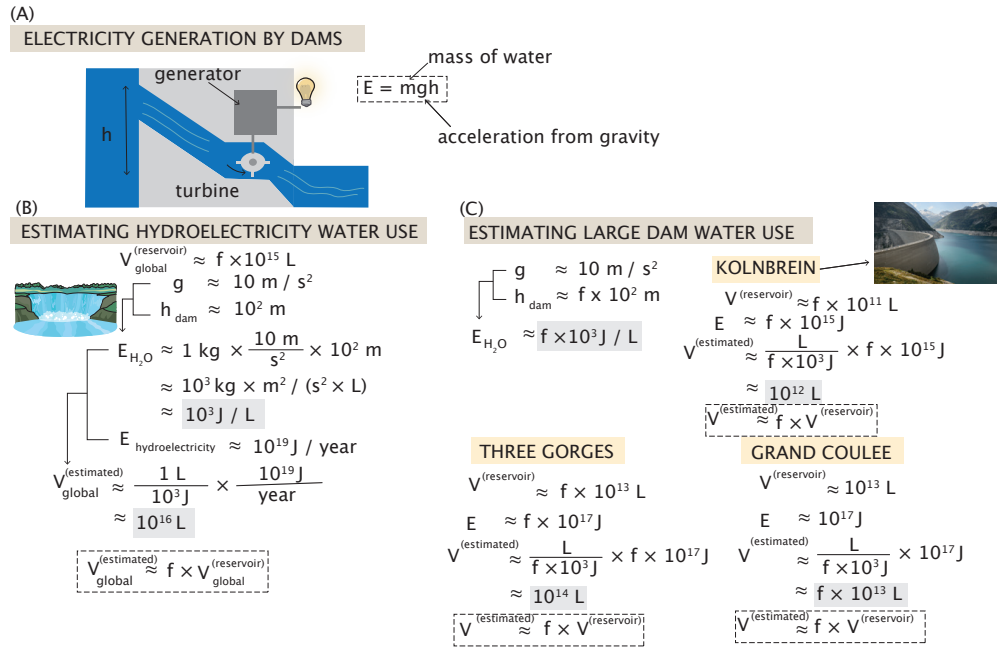


Figure 5.16: Estimating hydroelectric water use. (A) Schematic diagram of how falling water is harnessed to produce electricity. Water falls down a dam and hits a turbine, causing it to turn. The spinning turbine powers a generator, creating electricity. (B) Estimate of the global required water to generate hydroelectricity. 10^{19} J of electricity are produced each year by hydroelectricity, we estimate that 10^3 J are produced per liter of water. This means that about 10^{16} L are required each year. (C) Comparison of reservoir capacity for a few very large dams to the volume of water we estimate is required to generate the amount of electricity that they produce. For each of these examples, we find that our estimated water required is a few times larger than the reservoir capacity.

water through evaporation from their reservoirs. This increase in evaporation can cause considerable water consumption from hydroelectricity production. We can estimate the amount of water evaporated by hydroelectric dams in a couple ways. First, as noted above, the total global reservoir capacity of hydroelectric dams is $\approx 6 \times 10^{15}$ L, this corresponds to about $4 \times 10^{11} \text{ m}^2$ of surface area [217] (HUID: 23616). A survey of US hydroelectric dams found that the average height of water evaporated in the reservoir of each dam was about 1 m over the course of a year. Using this as a global average, we estimate that $\approx 10^{14}$ L of water is evaporated from hydroelectric dam reservoirs, shown in Fig. 5.17 (A) [220]. Another approach to this estimate is to use a reported value of water consumption

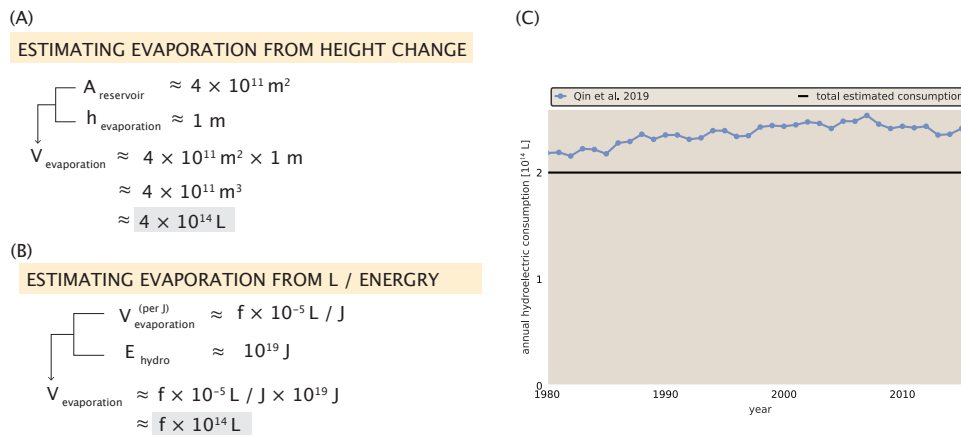


Figure 5.17: Estimating water consumption by hydroelectricity. (A) An estimate of the water evaporated from hydroelectric reservoirs if the global average evaporative loss was 1 m / year. (B) An estimate of water evaporated from hydroelectric reservoirs using a reported value of 2×10^{-5} L / J consumed. This estimate is a few times larger than the estimate based on an average height change. (C) Comparison of our estimated water consumption by hydroelectric dams with the value reported in [38].

per energy generated. Water consumption by hydroelectric dams varies widely by dam, however the global average has been estimated to be $39 \text{ m}^3 / \text{GJ}$, which we will call $f \times 10^{-5}$ L / J [221]. Using this value, along with the total energy produced from hydroelectricity, 10^{19} J, we estimate that $f \times 10^{14}$ L of water is consumed to generate hydroelectric power, as shown in Fig. 5.17 (B). These values agree well with the value reported by Qin *et al.* [38] of 3×10^{14} L, plotted in Fig. 5.17 (C)). Thus, $\approx f \times 10^{14}$ L of water is consumed each year through evaporation from hydroelectric reservoirs.

5.5.5 The Water Number

In this vignette, we have investigated and quantified the various ways in which humans use water. Water has shaped the way human civilization has developed and continues to play a major role in our everyday lives. Some of our uses are obvious, like drinking and showering, and others are more complex, like the water used to generate electricity via steam turbines and hydroelectric dams. In investigating these various uses of water, it is important to recognize that all uses are not equal

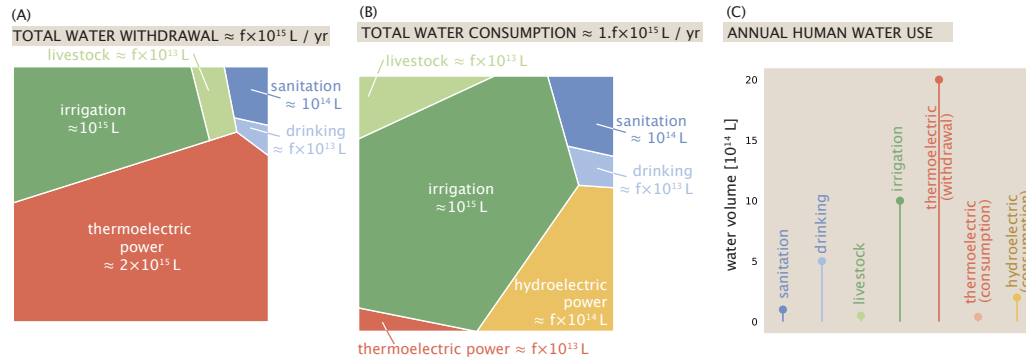


Figure 5.18: The distribution of human water use. (A) Breakdown of human water withdrawal shows that thermoelectric power dominates. (B) Breakdown of human water consumption shows that irrigation dominates. (C) Comparison of all major human uses of water by volume.

on how much they disrupt the natural water cycle.

Some interesting observations arise when we compare the share of water withdrawn, consumed, and used by the various ways we quantified in this vignette. When it comes to water withdrawal, shown in Fig. 5.18 (A), the cooling of thermoelectric power plants dominates the total volume with $\approx 2 \times 10^{15}$ L withdrawn per year. This is approximately twice the volume withdrawn for irrigation of cropland and pasture, followed by direct consumption by livestock and humans. Conspicuously missing in the breakdown of Fig. 5.18 (A) is hydroelectric power production which by definition withdraws no water, though it is certainly harnessed by humanity.

When considering water consumption, shown in Fig. 5.18 (C), irrigation becomes the dominating factor. The water consumed for the production of hydroelectric power is the next largest factor, which, despite producing an order of magnitude less power than thermoelectricity, consumes a few times more water each year. Fig. 5.18 (C) compares the magnitude of all major human uses of water, showing the dominance of irrigation and thermoelectric power production.

In total, humans withdraw and consume over 3×10^{15} L each year. Does this represent a significant impact on the hydrological cycle? To answer this question, we need to know a bit about the hydrological cycle and how much water

THE WATER NUMBER

$$W_a = \frac{\text{annual water volume used by humans}}{\text{global annual river discharge volume}} \approx \frac{\text{[Illustration of water withdrawal]} }{\text{[Illustration of river discharge]}} \approx 0.05$$

Figure 5.19: The Water Number. We define the water number as the total volume of water withdrawn by humans divided by total annual river discharge volume, ≈ 0.05

is available for human use. While a portion of the freshwater used by humans comes from groundwater stores, the recharge rate of this water is very slow and can be considered exhausted once it has been used [222]. To understand the magnitude of human water use, we should consider the amount of renewable freshwater resources. In theory, the maximum renewable freshwater available for human use is the difference between terrestrial precipitation and evapotranspiration [222]. A majority of this water becomes river discharge, $\approx 4.5 \times 10^{17}$ L each year, which we will consider the amount of available renewable freshwater resources [222]. Thus, we define the Water Number as the volume of water used by humans divided by the flux of water through rivers each year, ≈ 0.05 5.19.

Globally, humans use less than 10% of the available renewable freshwater, however the picture is different at the regional level. Unlike other resources, such as food, transporting water over large distances is very difficult. Thus, the human impact on the hydrological cycle varies widely across the globe. For this reason, the water scarcity index has been defined, which compares water use to available water at the local level, and a score above 0.4 is considered highly water stressed [222]. Using this index, it has been estimated that about one third of the human population lives in highly water stressed regions [222].

5.6 Human Impacts on the Nitrogen Cycle

Water is only one supply that humans provide for our crops. Equally as important is the addition of nitrogen. Nitrogen is the fourth most abundant element in biomass after carbon, hydrogen, and oxygen, making up a few percent of the mass of typical cells and organisms. N is required to build molecules critical for life, namely: DNA, proteins, and other essential biological molecules like heme and chlorophyll. Dinitrogen (N_2) is the most abundant constituent of Earth's atmosphere, but contains a very stable triple bond. Therefore, to satisfy biological needs for N, dinitrogen must be "fixed" into a form that is biologically accessible. Today this occurs by means of several important mechanisms, outlined in Figure 5.20, including fixation catalyzed by bacteria and plants (both agricultural and wild), fixation catalyzed by energy-intensive human technologies (the "Haber-Bosch" process), and fixation catalyzed by lightning. These processes mostly produce biologically accessible nitrogen in the form of ammonia (NH_3) which can be directly assimilated by plants and bacteria.

Once nitrogen enters biomass in the form of nitrogenous compounds like nucleotides and amino acids, it can be retained in a biologically-accessible form for extended periods. For example, when plants die, fungi and bacteria can take up and utilize the nitrogenous compounds released during their decomposition (e.g. amino acids in proteins). In addition to the recycling of biological N, nitrogen compounds can be used as energy sources by bacteria and fungi, which produces a "microbial N cycle" in soils, oceans and rivers, diagrammed in Fig. 5.20. These energy-conserving pathways interconvert several nitrogen oxides (NO_2^- , NO_3^- , N_2O) that play different roles due to their distinct chemistries. For example, N_2O is volatile and the third most significant greenhouse gas by some measures, while NO_2^- and NO_3^- are significant contributors to nitrogen run-off due to their negative charge. In contrast with N_2 , which is intrinsically stable and not very reactive on its own, NH_3 , NO_3^- , NO_2^- , N_2O , NO_2 (nitrogen dioxide), and NO (nitric oxide) are together called "reactive N" because they are used by living organisms and participate in rapid chemical reactions.

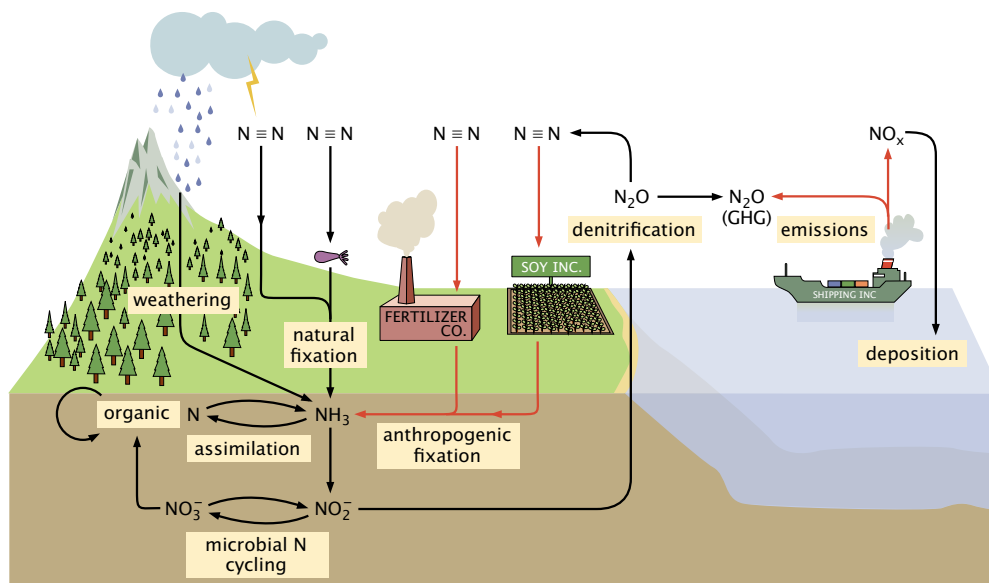


Figure 5.20: A simplified view of the global nitrogen cycle. Nitrogen is naturally provisioned to Earth’s surface via fixation of N_2 to NH_3 (ammonia) by N-fixing organisms and lightning. Additionally, NH_3 that is stored in rocks can be released by way of weathering. The NH_3 provided by these processes is then used by living organisms to synthesize important N-bearing molecules like DNA and proteins which we collectively term “organic N.” When plants decompose, organic N may be converted back to NH_3 . Beyond biological assimilation of NH_3 , a complex web of N transformations is mediated by microorganisms (“microbial N cycling”). NH_3 can be oxidized to nitrite (NO_2^-) which can then be further oxidized to nitrate (NO_3^-). NO_3^- and NO_2^- can be assimilated into plant biomass when microbes convert them into NH_3 . NO_3^- can also be “denitrified” to N_2 by way of N_2O . A small portion of this N_2O is released to the atmosphere, where it acts as a potent greenhouse gas. Humans have become major contributors to the nitrogen cycle via fixation of N_2 to NH_3 by the Haber Bosch process, fixation by leguminous crops, and emissions of NO_x and N_2O from fossil fuel burning. Some of these emissions return to the surface as reactive nitrogen species (“deposition”). For simplicity, this diagram depicts parts of the nitrogen cycle as occurring entirely on land or the ocean when in reality the nitrogen cycle is important in both. Also, this simplified diagram of the N cycle omits many intermediate N compounds as well as many interconversions between these species.

In this vignette we will attempt to estimate the magnitudes of the major fluxes in the nitrogen cycle diagrammed in Figure 5.20. As we will see, due to the large number of important N species and the complex processes that interconvert them, this is no simple task. To get started, however, it is important to estimate the size of nitrogen reservoirs on Earth. The majority of nitrogen that is accessible to biological and technological N fixation is stored in the atmosphere, which is $\approx 80\%$ N_2 , with most of the remainder being O_2 [223]. Reminding ourselves that N and O are adjacent on the periodic table (14 and 16 mass units, respectively) we can make the approximation that N_2 and O_2 have roughly the same mass. Given this assumption, we can calculate the mass of atmospheric N from the mass of the atmosphere as $\approx 4 \times 10^{18}$ kg N_2 [224] [Fig. 5.21 (A)].

The atmosphere is not actually the largest N reservoir, it is only the largest reservoir of accessible N. A very large quantity of mostly-inaccessible N is bound up in minerals in the Earth's crust. A small fraction of this is released annually via a process called "weathering" where reactions of water with minerals liberate N mostly in the form of NH_3 . How much N is bound up in minerals? Based on the terrestrial surface area of Earth, the average density of rocks, and their average N content, we estimate that the top meter of the Earth contains $\approx 10^{14}$ kg N (Fig. 5.21 (B)). This very closely approximates a value estimated from a more detailed set of parameters including sampled rock N concentrations, maps of surface rock types and abundances, and detailed rock density data [225]. Notably, we ignored deeper N-bearing minerals based on the assumption that they are less accessible to weathering in the short term. Including deeper rocks increases the size of the geological N reservoir by several orders-of-magnitude to $\approx 10^{19}$ kg N [224].

Prior to the industrial revolution, most reactive nitrogen was derived from biological nitrogen fixation – the transformation of atmospheric N_2 to NH_3 by bacteria – with small amounts generated by lightning discharge in the atmosphere and by rock weathering. Since the availability of reactive nitrogen is a primary factor that limits agricultural yields [226], a growing human population drove research into technologies to generate reactive N to support agriculture. One such

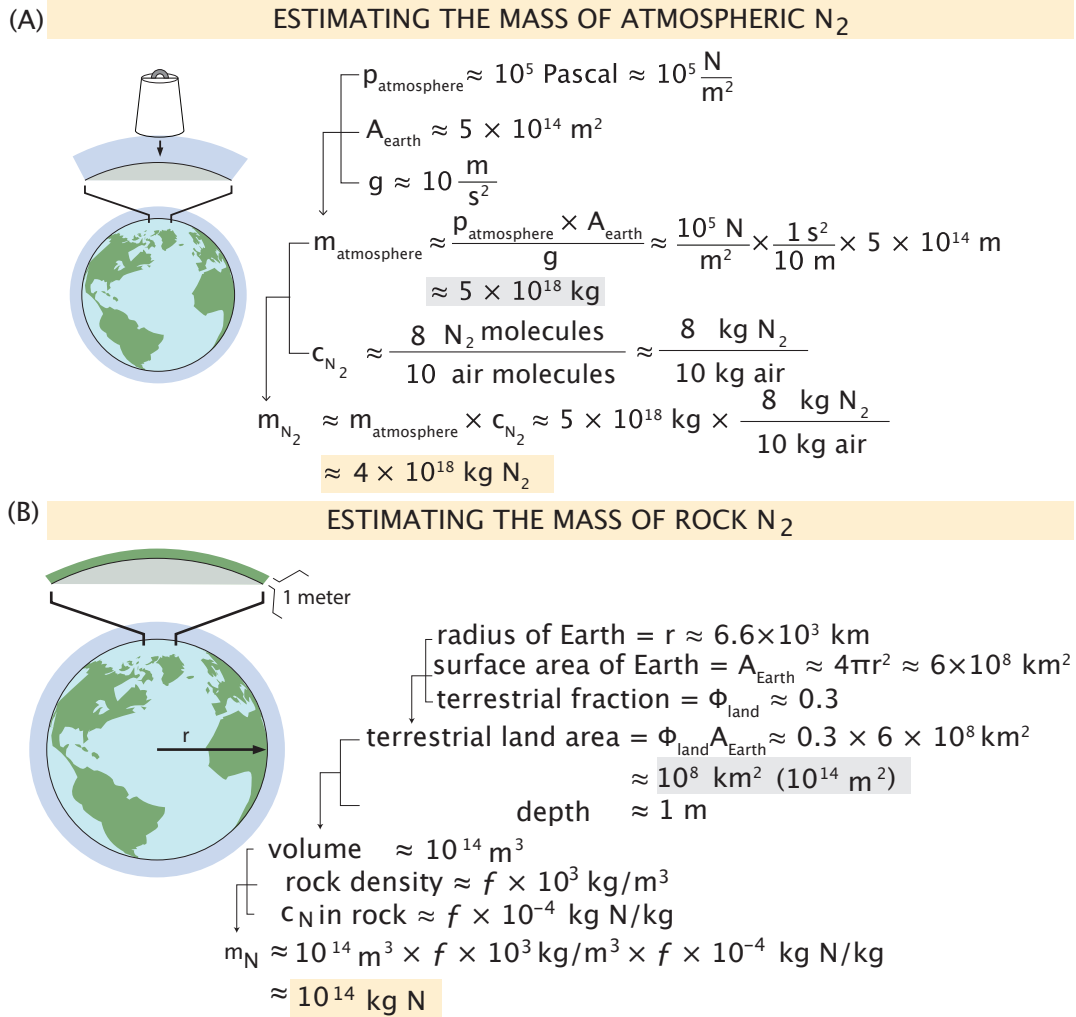


Figure 5.21: Estimate of the mass of nitrogen in the atmosphere and land. (A) Estimate of the mass of N_2 in the atmosphere. We first estimate the mass of the atmosphere and then use the fact that $\approx 80\%$ of the atmosphere is composed of N_2 to estimate that $\approx 10^{18}$ kg of N_2 is in the atmosphere. (B) Estimate of the mass of nitrogen stored in rocks in the top meter of the Earth's crust. We estimate the terrestrial surface area of the Earth, combined with the average and concentration of nitrogen in rocks in order to estimate that $\approx 10^{14}$ kg of nitrogen is stored in this reservoir.

process, called the Haber-Bosch process after its inventors, was developed during World War 1 and produces NH_3 from N_2 at high-pressure and temperature. This process is so crucial to global agricultural production that today, Haber-Bosch NH_3 production is responsible for a few percent of global power consumption. To build intuition for the scale of human impacts on the nitrogen cycle, we will estimate the natural and anthropogenic sources of reactive nitrogen. As we will see, anthropogenic sources of reactive N now rival natural sources, highlighting our impressive impact on the planet and its biosphere.

5.6.1 Natural Sources of Reactive Nitrogen

Most organisms cannot utilize N_2 in its atmospheric form, but a small minority of organisms, called diazotrophs, are capable of N fixation. Despite the incredible strength of the N_2 triple bond, evolution found a way to cleave it. The majority of natural nitrogen fixation is catalyzed by diazotrophic bacteria that use an enzyme called “nitrogenase” to convert N_2 into NH_3 . Nitrogenase is a fascinating and complex enzyme worthy of detailed study [226] but we will not dwell on its structure, mechanism, or evolution here. One salient detail bears mention, however: nitrogenase is exceptionally sensitive to O_2 and so biological N_2 fixation typically occurs in environments with low O_2 .

Plants can also fix nitrogen, but indirectly through association with diazotrophic bacteria. In this symbiosis, plant roots provide carbon compounds (food) and a low O_2 environment supporting N fixation by partner bacteria, which, in turn, provide reactive N to the plant as NH_3 . This symbiosis is best studied and most common in leguminous plants like soybeans, which grow special structures called root nodules that house diazotrophic bacteria. Nodules maintain a low O_2 concentration because they slow O_2 entry, express proteins that sequester O_2 (“leghemoglobin”) and are also composed of plant and bacterial tissue that consume O_2 to produce metabolic energy. In contrast, this tissue does allow atmospheric N_2 to enter the nodule and become fixed by resident bacteria.

Roughly 80% of terrestrial biological nitrogen fixation occurs in evergreen broadleaf forests and savannas, where diazotrophs living in symbiosis with plant

hosts provide fixed N in return for reduced carbon compounds [227]. The remaining 20% of terrestrial biological nitrogen fixation is conducted by diazotrophs living independently from their plant hosts. The estimated flux of terrestrial biological nitrogen fixation has been estimated to fall within the range of $\approx 6 - 20 \times 10^{10}$ kg / year [228], which we approximate to 1×10^{11} kg / year (HuID: 15205).

Although biological nitrogen fixation predominates, it is not the only natural route to furnish fixed nitrogen to Earth's ecosystems and is augmented by other minor sources. When lightning strikes, the spatiotemporally intense energy release is strong enough to break the bonds in O_2 , forming O radicals. These radical species react with N_2 under the high-temperature conditions to form nitrogen oxides ($NO_x = NO$ and NO_2). NO is the predominant form that results from the lightning strike and its subsequent reaction with ozone (O_3) can form NO_2 . NO_x are key molecular constituents of the atmosphere that govern ozone concentrations, regulate bulk atmospheric chemistry, and cause a variety of changes that impact the strength of the greenhouse effect [229]. Lightning bolts vary in strength and frequency but a reasonable estimate for their average energy discharge is $\approx 1 \times 10^9$ J / strike [230] and they have been measured to occur on average 44 times per second globally [231]. Based on the energy required to dissociate molecular nitrogen and molecular oxygen, the energy released by lightning strikes, and the global frequency of lightning strikes, we estimate that lightning produces $\leq f \times 10^{10}$ kg NO_x per year which translates to $\approx 1 \times 10^{10}$ kg N per year, shown in Fig. 5.22 (A). The simplifying assumptions made here mean that this estimate is an upper bound on the flux of N_r produced by lightning. More detailed estimates suggest that the nitrogen fixation by lightning is about one order of magnitude smaller and that lightning strikes produce $\approx 5 \times 10^9$ kg of N_r per year [228, 229]. As precipitation gathers this N_r , it deposits it on Earth's surface.

It was long assumed that lightning was the only natural source of N_r besides biological nitrogen fixation. Over geological timescales, the fixed nitrogen that has accumulated in sediment as a result of ancient forms of biological nitrogen fixation becomes trapped in rocks and sediments [225]. Although this nitrogen is reactive, it is sequestered from the biosphere and can only be made available

via chemical and physical weathering processes. Although this supply does not represent new fixation of nitrogen, it does represent delivery of what is newly available N_r to ecosystems. Based on global averages of weathering rates and average rock nitrogen concentrations, we estimate that weathering releases $\approx 5 \times 10^9$ kg of N_r per year, shown in Fig. 5.22 (B). More detailed approaches have estimated that the weathering of rocks on Earth's terrestrial surface releases $\approx 1.5 \times 10^{10}$ kg of N_r per year to the biosphere [225].

To summarize the natural sources of reactive nitrogen, biological nitrogen fixation enabled by diazotrophs expressing nitrogenase dominates the natural terrestrial flux of N_r and provides $\approx 1 \times 10^{11}$ kg of ammonia per year. Roughly 5×10^9 kg of NO_x is generated by lightning strikes in the upper atmosphere. Finally, weathering of rocks and sediments frees $\approx 1.5 \times 10^{10}$ kg of ammonia per year. In sum, $\approx 2 \times 10^{11}$ kg of reactive nitrogen is delivered to the terrestrial biosphere on an annual basis via natural processes. Although we do not rigorously discuss marine nitrogen fixation in this vignette, it is estimated that biological nitrogen fixation contributes $\approx 1 - 2 \times 10^{11}$ kg of N_r per year to marine ecosystems. In preindustrial times, these sources represented the total amount of reactive nitrogen that was input to the biosphere each year. Understanding the magnitude of these sources will help us understand the scale of anthropogenic nitrogen fixation that we develop below.

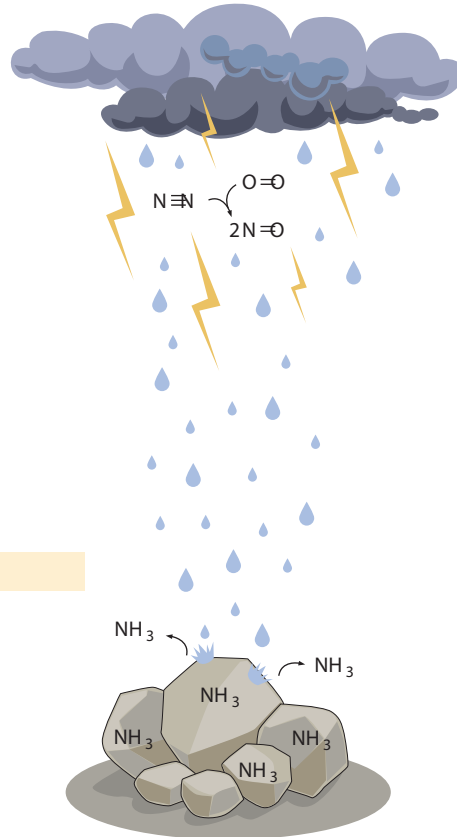
5.6.2 Anthropogenic Nitrogen Fixation

Since the advent of industrial agriculture and extensive fossil fuel use, humans have become formidable rivals to natural formation of N_r . Intentional production of fixed nitrogen arose to meet needs for fertilizer, explosives, and other nitrogen-containing molecules. The natural sources of fixed nitrogen described above were insufficient to meet the nitrogen needs of a rapidly growing human population and incited research into alternative means of provisioning nitrogen to agricultural soils and obtaining nitrogenous feedstocks.

Fertilizer is the predominant use of synthetically fixed nitrogen. Therefore, we will estimate the amount of nitrogen required for agricultural production,

(A) ESTIMATING NITROGEN FIXATION BY LIGHTNING

$$\begin{aligned}
 & \begin{cases} E(\text{per mol NO}) & \approx 10^6 \text{ J}/(\text{mol NO}) \\ E_{\text{lightning}} & \approx 10^9 \text{ J}/\text{strike} \\ m_{\text{N}}^{(\text{mol NO})} & \approx 10^{-2} \text{ kg}/\text{mol NO} \end{cases} \\
 & \downarrow \\
 m_{\text{N}}^{(\text{per strike})} & \approx \frac{1 \text{ mol NO}}{10^6 \text{ J}} \times \frac{10^9 \text{ J}}{\text{strike}} \times \frac{10^{-2} \text{ kg}}{1 \text{ mol NO}} \\
 & \approx 10 \text{ kg N}/\text{strike} \\
 & \begin{cases} r_{\text{strike}} & \approx f \times 10 \text{ strikes}/\text{s} \\ t_{\text{year}} & \approx f \times 10^7 \text{ s}/\text{year} \end{cases} \\
 & \downarrow \\
 m_{\text{N}}^{(\text{per year})} & \approx \frac{10 \text{ kg}}{\text{strike}} \times \frac{f \times 10 \text{ strikes}}{\text{s}} \times \frac{f \times 10^7 \text{ s}}{\text{year}} \\
 & \approx 10^{10} \text{ kg N}
 \end{aligned}$$



(B) ESTIMATING FIXED NITROGEN RELEASED BY WEATHERING

$$\begin{aligned}
 & \begin{cases} m_{\text{N}}^{(\text{per kg rock})} & \approx 5 \times 10^{-4} \text{ kg N}/(\text{kg rock}) \\ r_{\text{weathering}} & \approx 10^{-1} \text{ kg rock}/(\text{m}^2 \times \text{year}) \\ A_{\text{land}} & \approx 10^{14} \text{ m}^2 \end{cases} \\
 & \downarrow \\
 m_{\text{N}} & \approx \frac{5 \times 10^{-4} \text{ kg N}}{\text{kg rock}} \times \frac{10^{-1} \text{ kg rock}}{\text{m}^2 \times \text{year}} \times 10^{14} \text{ m}^2 \\
 & \approx 5 \times 10^9 \text{ kg N}/\text{year}
 \end{aligned}$$

Figure 5.22: Back-of-the-envelope estimates for NO_x produced by lightning and N released by weathering. (A) When lightning strikes, it cleaves the bonds in N_2 and O_2 which is a prerequisite for NO_x to be formed. To break the bonds in one mole of O_2 requires 498 kJ and cleavage of the even stronger N_2 bond requires 945 kJ / mol. Assuming perfect conversion, 2 moles of NO would then be formed, but the formation of this molecule requires additional input of energy. Thus, the formation of one mole of NO requires at least 1×10^6 J of energy input. (B) As rocks are weathered by chemical, physical, and biological processes, they release their constituent components to the environment. By multiplying the average concentration of nitrogen in geological material, the average global rate of rock weathering and the area of Earth's terrestrial surface, we estimate that 5×10^9 kg N / year are released from the rock nitrogen reservoir.

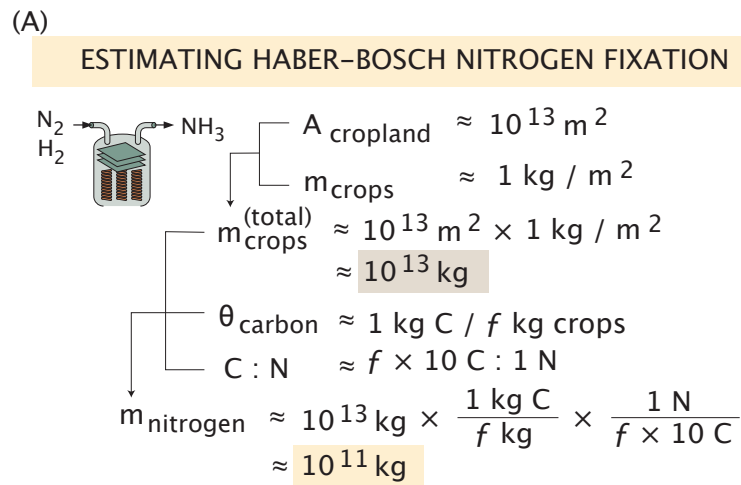


Figure 5.23: Estimate of nitrogen fixation for fertilizer. Most of the nitrogen synthetically fixed by the Haber-Bosch process goes to nitrogenous fertilizer. We estimate the amount of nitrogen contained in crops to estimate Haber-Bosch fixed nitrogen.

shown in Figure 5.23. In the "Extent of human land use" section, we showed that agricultural land used for crops is $\approx 10^{13} \text{ m}^2$ and that the edible mass density is $\approx 1 \text{ kg} / f \text{ m}^2$. We will assume that edible mass is about half of total plant mass. Thus, the total mass of plants raised as crops is $\approx 10^{13} \text{ kg}$. We will estimate the amount of nitrogen in this plant mass. About one half of plant mass is dry weight, of which about half is carbon. The ratio of nitrogen to carbon in plants is well known, about $1 : f \times 10$, i.e. there is one nitrogen atom for every $f \times 10$ carbon atoms. Since nitrogen and carbon have approximately the same mass, we can use this information to estimate that the mass of nitrogen in crops is $\approx 10^{11} \text{ kg}$. Since these crops are grown yearly, if all of this nitrogen is derived from the application of synthetic nitrogenous fertilizer, then $\approx 10^{11} \text{ kg}$ of nitrogen would need to be produced through the Haber-Bosch process each year. In fact, the USGS reports that $\approx 1.5 \times 10^{11} \text{ kg}$ of NH_3 was produced in 2019, in good agreement with this estimate [167].

The Haber-Bosch process is not the only way that humans are impacting the nitrogen cycle, however. Cultivation of legumes, plants that have root nodules that house nitrogen fixing bacteria such as soy beans and lentils, also increases the

amount of biologically fixed nitrogen. This adds $\approx 6 \times 10^{10}$ kg of fixed ammonia annually [228]. Furthermore, burning fossil fuels results in the production of NO_x through the reduction of fuel nitrogen molecules and the production of radicals through the high heat which react with N_2 in the atmosphere. In this way, burning of fossil fuels results in an additional $\approx 4 \times 10^{10}$ kg of fixed ammonia yearly [228].

These additional inputs of reactive nitrogen arising from human activities have enabled numerous the growth of the human population and societal developments while simultaneously stimulating ill effects on the environment. Excessive addition of nitrogen results in runoff to the ecosystem which can have many negative effects, including acidification and eutrophication of terrestrial and aquatic systems [232]. Eutrophication is an effect from an excess nutrient supply which can lead to altered growth such as toxic algal blooms that degrade the natural ecosystem. In addition, reactive nitrogen in the atmosphere degrades stratospheric ozone, contributing to global climate change [233].

5.6.3 The Nitrogen Number

Humans demand fixed nitrogen to sustain high agricultural yields which support the human population and our livestock. Fixed nitrogen is also generated to meet demand for other nitrogen-containing compounds, such as explosives and many pharmaceuticals. Our reliance on the nutrient is as deeply rooted in chemistry and biology and is rivaled in importance perhaps only by water. This undeniable importance has led humans to invest great time and effort in devising numerous ways to acquire fixed nitrogen.

Currently, humans are impacting the nitrogen cycle in a number of ways. Figure 5.24 summarizes the anthropogenic and natural sources of reactive nitrogen. From this, it is clear that many anthropogenic activities rival the magnitude of natural nitrogen fixation pathways. We summarize this anthropogenic effect on the nitrogen cycle through the Nitrogen Number, which we define as the mass of synthetically fixed nitrogen to the mass that is naturally fixed 5.25. This ratio is now approximately one, indicating that humans now rival the magnitude of natural processes in the nitrogen cycle. This perturbation to the natural cycle is

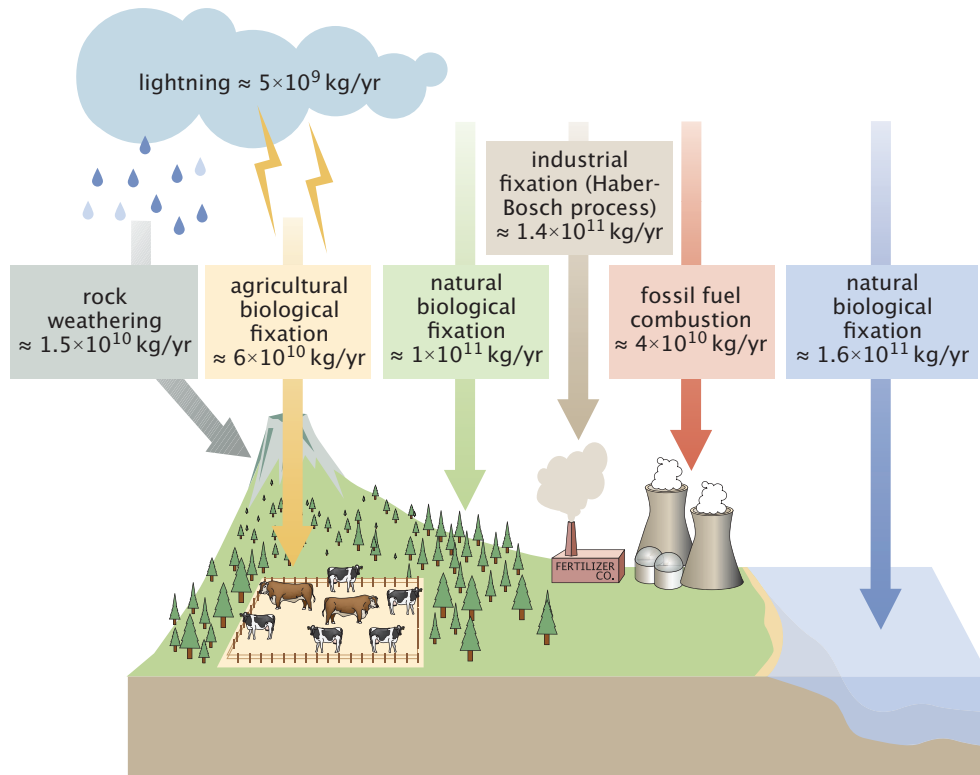


Figure 5.24: The global natural and anthropogenic sources of reactive nitrogen. Reactive nitrogen is created from the atmospheric N_2 reservoir from a limited set of pathways. The advent of industrial nitrogen fixation, legume cultivation, and fossil fuel combustion have led to high levels of anthropogenic reactive nitrogen creation. These now rival the natural sources of reactive nitrogen: marine & terrestrial biological nitrogen fixation, rock weathering, and lightning.

having many consequences for ecosystem health, biodiversity, and climate change [232].

5.7 Human Impacts on Methane Emissions

What are side effects and byproducts from our agricultural practices? One is the production of methane gas. In discussions of climate change, the most important gases are the "greenhouse gases." Sunlight heats the surface of the Earth, which in turn emits energy in the form of thermal radiation. Similar to the roof of a greenhouse, the 'greenhouse gases' trap and absorb this thermal radiation,

THE NITROGEN NUMBER

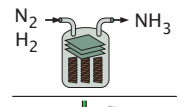

$$N_2 = \frac{\text{mass of N}_2 \text{ fixed via the Haber-Bosch Process}}{\text{mass of N}_2 \text{ fixed biologically}} \approx \frac{\text{Haber-Bosch Process Diagram}}{\text{Biological Fixation Diagram}} \approx 1$$



Figure 5.25: The nitrogen number. We define the nitrogen number as the mass of synthetically fixed nitrogen through the Haber-Bosch process to the mass of naturally fixed nitrogen, which is about one.

heating the lower atmosphere and increasing the surface temperature around the globe. Anthropogenic emissions of greenhouse gases contribute to climate change through this absorption process, with the major culprits being carbon dioxide (CO_2), methane (CH_4), tropospheric ozone (O_3) and nitrous oxide (N_2O).

Greenhouse gases absorb radiation at different wavelengths, and how much they absorb depends on their concentration. Water vapor (H_2O), for example, is so abundant that it already absorbs most radiation at its absorption wavelengths. Due to this, small increases in H_2O concentration lead to limited warming — there is not much radiation left to absorb. Methane is much less prevalent in the atmosphere, so similar concentration changes lead to much stronger absorption and global warming. Despite having a concentration ≈ 200 times smaller than carbon dioxide, it is estimated that methane has about 25 times the warming potential of an equal mass of CO_2 over 100 years. Thus, methane is an extremely potent greenhouse gas that is critical to include in conversations of climate change. While produced in a variety of ways, agriculture and human land use are responsible for a large fraction of anthropogenic methane emissions, as we will dive into in this section. We will quantitatively explore the various sources and sinks of CH_4 — both natural and anthropogenic — to understand how human activities are impacting the fluxes of this critical greenhouse gas.

5.7.1 Assessing Atmospheric Methane Concentrations

The past 200 years (and especially since 1970) have seen a rapid acceleration in the production of greenhouse gases including CH_4 . The measured concentration of CH_4 in the atmosphere since about 1970 is shown in Figure 5.26 (A). Before we can consider the scale of anthropogenic CH_4 emissions, we will try to understand the total amount of CH_4 that has been added to the atmosphere over this timeframe. This will give us an important sense for how large of an impact humanity has had on the concentration of this important greenhouse gas.

Recent estimates place the CH_4 composition of the atmosphere at ≈ 1900 parts per billion (ppb) [28](HuID: 44953), meaning that ≈ 2 out of every million molecules in the atmosphere are CH_4 . To estimate the mass of CH_4 that this corresponds to, we first estimate the mass of the atmosphere ($\approx 5 \times 10^{18}$ kg). Using the molecular mass of CH_4 compared to the average molecular mass of the atmosphere, we estimate that $\approx 5 \times 10^{12}$ kg CH_4 is currently in the atmosphere, as shown in Figure 5.26 (B).

At the end of the pre-industrial period in ≈ 1800 C.E., the CH_4 concentration is estimated to have been ≈ 800 parts per billion, or around 1 in every ≈ 2 million atmospheric molecules being CH_4 [234]. This value is measured by examining the CH_4 concentration in air bubbles trapped in long-lived ice that can be dated to that time period [235]. Using the conversion between methane concentration and total mass presented in Figure 5.26 (B), we arrive at a preindustrial methane mass of $\approx 2.5 \times 10^{12}$ kg, about a factor of two smaller than in the present day. While 200 years may seem long on the timescale of a human lifespan, it is a fraction of an instant compared to the geological history of Earth and represents the rapidity with which we can sculpt the composition of the atmosphere.

The mass of atmospheric methane is changing rapidly. In recent decades, methane concentrations have been increasing at a rate of ≈ 10 parts per billion per year, meaning that $\approx 2 \times 10^{10}$ kg of methane are deposited into the atmosphere every year [25]. How do human activities contribute to increasing methane concentrations? What fraction comes from ruminant animals versus extraction of fossil fuels? How long does CH_4 persist in the atmosphere, and where does it go when it

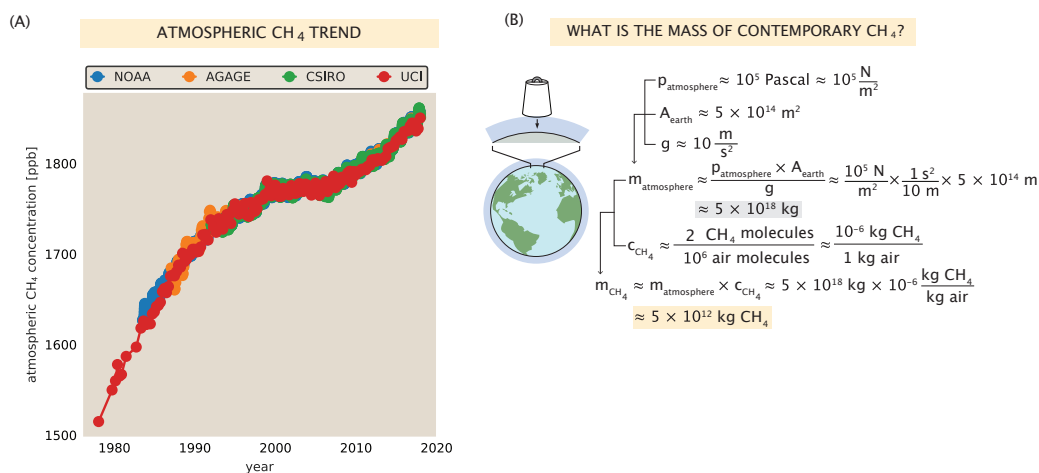


Figure 5.26: Sizing up the mass of methane in the atmosphere. (A) Trend in atmospheric methane concentration over time, measured by four different sources. (B) Estimate of the mass of methane in the atmosphere. First, we estimate the mass of the atmosphere based on pressure, then use the concentration of methane in the atmosphere to estimate the mass of methane in the atmosphere.

decays? We will dive into these questions in the rest of this vignette.

5.7.2 Methane Emissions Due to Agriculture

Agricultural methane emissions are driven primarily by the metabolic processes of microorganisms consuming organic matter. Microorganisms take the large, complex organic molecules that biology produces, and use chemical reactions to break them down into simpler forms, which releases energy that can be used to power the functions of life. Most microorganisms produce CO₂ as a byproduct of metabolism. In order to produce CO₂, however, a source of oxygen is required. In environments where oxygen is scarce, microorganisms break down carbon-containing molecules and produce CH₄ as a byproduct instead. This process is the dominant way that methane is produced in many agricultural and natural settings.

The majority of agricultural methane emissions come from ruminant livestock, including cattle, asses, and goats. These methane emissions are driven by a process known as enteric fermentation, which occurs when microorganisms living in the oxygen-deprived animal gut digest food and produce CH₄ as a byproduct. Being

the most populous of the ruminants, cattle produce the majority of methane emissions due to enteric fermentation. Microbial metabolism of animal manure also produces methane emissions, but is a much smaller contributor.

How much methane does a cow produce? We can perform an estimate based on our understanding of cow metabolism. Let us assume that an adult cow eats ≈ 10 kg of dry feed each day. As we estimated in sizing up human livestock cultivation, to sustain our consumption of dairy and beef, a population of ≈ 1 billion cattle is needed. Feeding this population thus requires $\approx f \times 10^{12}$ kg of feed per year. Since the feed is plant matter (e.g. corn or soy in the USA), it is roughly 50% carbon by weight. If we assume that a few percent of cattle feed by mass is converted into CH_4 , then the total amount of methane produced from cattle is $\approx 10^{11}$ kg CH_4 per year, as shown in Figure 5.27 (A).

Reports of methane produced by cattle range from $\approx 50 - 120$ kg per cow annually, which is in agreement with our estimate of a few percent conversion efficiency of feed to methane [236]. Indeed, our estimate of $\approx 10^{11}$ kg CH_4 annually is in good agreement with data from the FAO, which reports $\approx 10^{11}$ kg CH_4 per year from enteric fermentation, as shown in Figure 5.27 (B). This value includes emissions from other animals such as asses, mules, and sheep, but cattle are the dominant source of emissions from enteric fermentation, accounting for greater than 75% of emissions.

Another important contributor to agricultural methane emissions is rice cultivation. Rice is cultivated in flooded paddy fields. Because the water in these fields stays stagnant for long periods of time, the soils they cover are often oxygen-poor and become hospitable to anaerobic microbes. Similarly to the cattle gut, the microbes in the soil of rice paddies consume carbon containing compounds produced by the plants, producing methane as a byproduct.

To estimate methane production from rice paddies, we must first know the area of rice paddies globally. We can estimate this through the amount of rice consumed by each person. If we assume that the average diet consists of a few kg of rice per week, then global rice consumption is $\approx 5 \times 10^{11}$ kg annually, as shown in Figure 5.28 (A). Further, the average crop yield is 1 kg per few m^2 per year, as

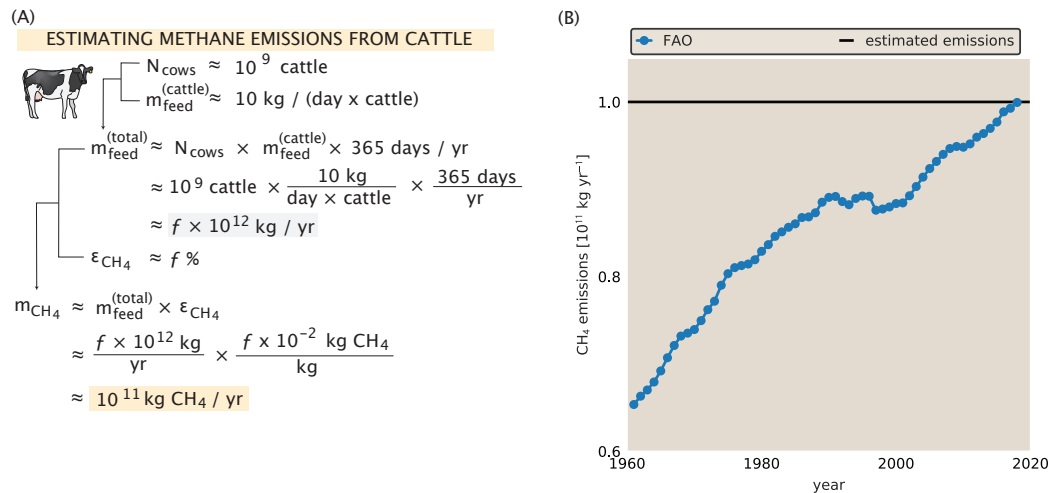


Figure 5.27: Estimate of methane emissions from cattle. (A) Estimate of methane emissions due to fermentation in the digestive tract of cattle based on the amount of feed they consume and the fraction that is converted to methane. (B) Comparison of our estimate with data from the FAO [237].

was also used in our “extent of human land use” section. With our estimate for consumption, then, we estimate that the area of rice paddies globally is $\approx 10^{12} \text{ m}^2$. This is in agreement with data from the FAO, which reports $\approx 1.6 \times 10^{12} \text{ m}^2$ of rice paddies [237], shown in Fig. 5.28 (B).

Next, we wish to know how much methane is produced across this area of rice paddies. The amount of methane produced per m^2 of rice paddy depends on the amount of microbes living in the soil, including not only ones that produce methane but others that consume methane, thereby reducing the amount of methane that reaches the atmosphere. To avoid these complications, we will use the measured value of $\approx 10 \text{ mg CH}_4$ per m^2 per hour. We can simply multiply this value by our estimate for rice paddy area and the number of hours in a year to arrive at an estimate of $\approx 10^{10} \text{ kg CH}_4$ per year emitted by rice paddies, as shown in Fig. 5.28 (C). This estimate is within a factor of a few of the value reported by the FAO, $\approx 2.5 \times 10^{10} \text{ kg CH}_4 / \text{yr}$, plotted in Fig. 5.28 (D) [237] (HuID: 14300).

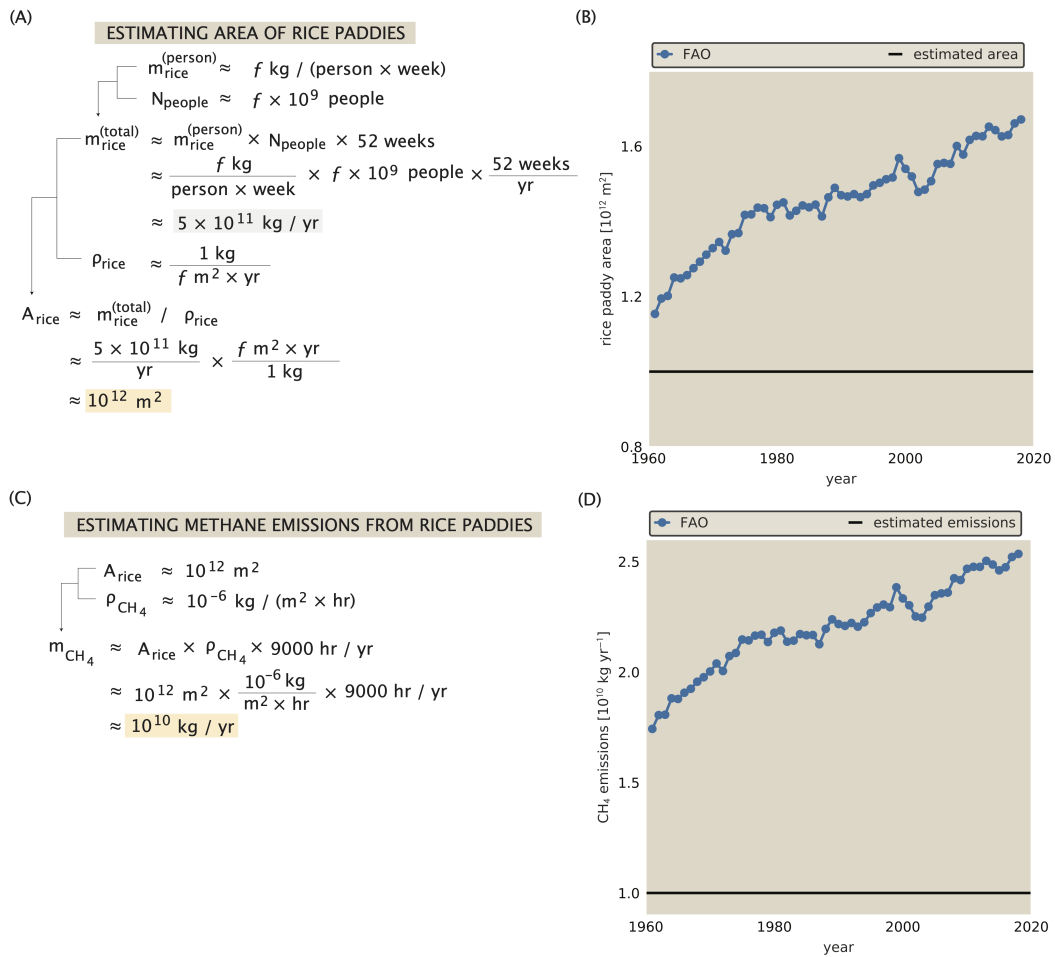


Figure 5.28: Estimate of methane emissions from rice paddies. (A) Estimate of the area of rice paddies required for human consumption, assuming a few billion people consume a few kg of rice per week. (B) Comparison of estimated rice paddy area with the value reported by the FAO. (C) Estimate of the amount of methane produced from these rice paddies using reported values of methane emissions from rice soil. (D) Comparison of estimated methane emissions from rice paddies with data from the FAO [237].

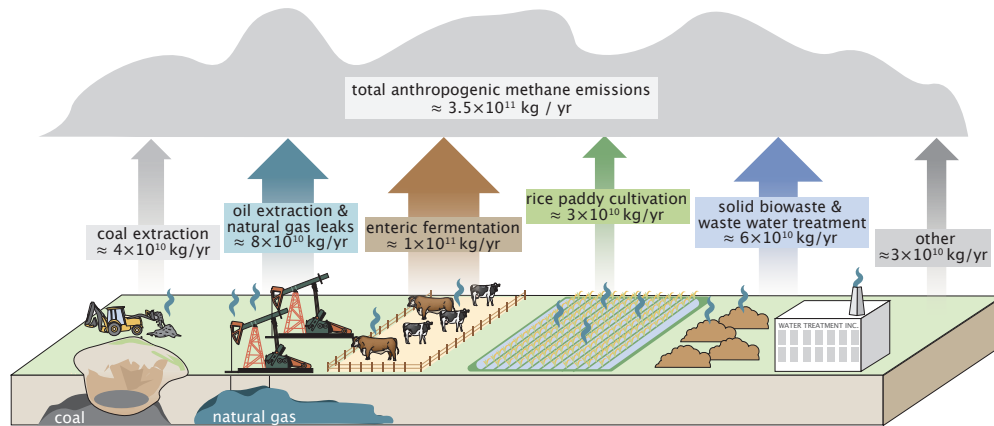


Figure 5.29: Anthropogenic sources of methane emissions. Anthropogenic emissions are dominated primarily by two sectors: fossil fuels extraction and agriculture. Unlike CO_2 , methane is not released as emissions from the combustion of fossil fuels, but rather in the storage and extraction process. Agricultural emissions are dominated by enteric fermentation in animal guts, but manure management and rice cultivation also contribute significantly. Solid and liquid bio waste, such as food waste as well as waste from processing pulp and paper, makes up the majority of the remaining emissions. Other miscellaneous factors include emissions from biofuels and forest clearing.

5.7.3 Total Anthropogenic Methane Emissions

While the above sources represent the largest individual sources of methane, there are many other, smaller contributing factors that add to this above total. The burning of biofuels, biomass burning for forest clearing, the treatment of palm oil, and many other smaller factors all contribute a total of around $\approx f \times 10^{10}$ kg / year to the total of anthropogenic emissions. Figure 5.29 summarizes the anthropogenic methane emission sources and their magnitude, which add to $\approx 3.5 \times 10^{11}$ kg CH_4 per year.

5.7.4 Natural CH_4 Sources

While anthropogenic emissions are a significant part of global methane fluxes, a similar amount of methane is emitted from natural sources, these sources are summarized in Figure 5.30. As with human agriculture, decomposition of biomass by anaerobic microbes is the dominant process by which methane is produced through biological mechanisms in the natural world [25]. Anaerobic environments

generally occur in flooded soils, similar to those of rice paddies. The largest anaerobic biome on Earth is its wetlands, which occupy $\approx 10^{13}$ m² and produce an estimated $\approx 2 \times 10^{11}$ kg/year of methane [25]. Rivers, ponds, lakes, and other freshwater bodies also tend to have anoxic environments in their soil beds, which contribute another $\approx 10^{11}$ kg/year of methane [25], shown in Fig. 5.30). While these are considered natural emissions, it is important to note that human activities are impacting the extent of wetlands and thereby influencing the magnitude of natural methane emissions.

The remaining biologically-produced methane comes from combustion of biomass (wildfires) and anaerobic animal digestion. One particularly large source of methane in the latter category is termites, which are individually responsible for roughly $\approx 10^{10}$ kg / year of methane production [25]. Digestion by wild animals amounts to $\approx 5 \times 10^9$ kg / year [25] (Fig. 5.30).

Another source of methane worth discussing is permafrost thaw. Permafrost is soil that remains frozen throughout the year, found in the Arctic and Antarctic. The permafrost layer typically sits underneath a surface-level ‘active layer,’ which freezes and thaws seasonally, and above lower layers of soil that are maintained at temperatures above freezing by geothermal heating. Although typically on the order of meters, the depth of the permafrost layer can reach several km in places, and the area of the layer is believed to extend nearly 2×10^7 km² [238].

Because it is permanently frozen, biomass, as well as simpler methane-containing compounds stored in permafrost layers do not undergo decomposition, or undergo decomposition at very slow rates. One consequence of this is that permafrost represents a large store of organic material that, when the permafrost layer thaws, can be released or decomposed anaerobically. The total volume of organic carbon stored in permafrost layers, is believed to be $\approx 1.5 \times 10^{15}$ kg, or around one half of all organic carbon stored in the soil [238], as illustrated in Fig. 5.30. When permafrost thaws, trapped biomass is exposed to more rapid decomposition, and trapped methane is able to seep through to the surface.

Significant amounts of permafrost thaw are already underway due to rising temperatures in the Arctic [125]. It is difficult to estimate precisely how this

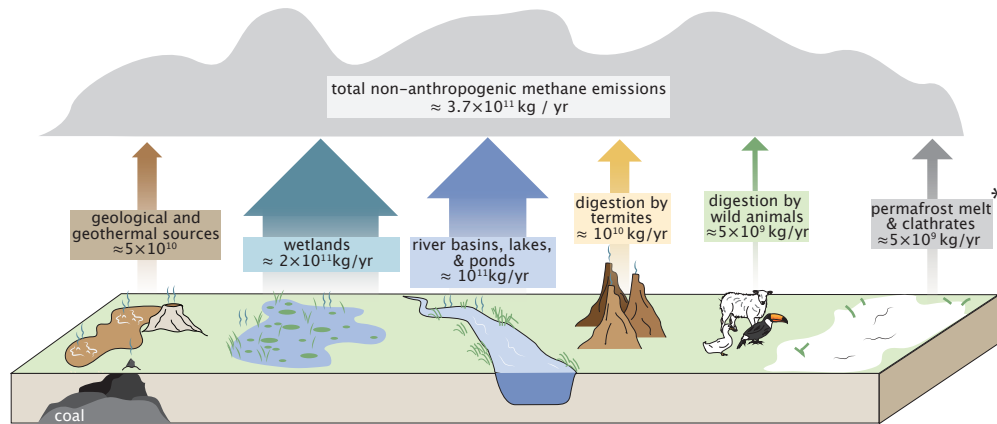


Figure 5.30: Non-anthropogenic sources of methane emissions. Natural methane emissions occur mostly in flooded areas, such as wetlands, river basins, lakes, and ponds, where the oxygen content in the soil is low, allowing methanogenic archaea to thrive. Methane trapped in hydrocarbon deposits is also brought to the surface through seepage or active processes like mud volcanoes. Other marginal sources of methane include animal digestion and wildfires. Due to uncertainty in the methane emissions from melting permafrost, the value has not been included in the total, indicated by an asterisk (*).

contributes to current methane fluxes, but some estimates place this number at 5×10^9 kg / year [25]. Some allow the possibility that annual emissions from thawing permafrost may reach as high as 6×10^{10} kg / yr by the end of the century [125]. Because of the small magnitude of methane emissions due to permafrost currently, we neglect it in our calculation of the total methane budget. However, there is a possibility that this number could increase significantly as temperatures increase, as permafrost thaws ever more rapidly and exposes more biomass to decay.

In total, natural sources contribute roughly $\approx 4 \times 10^{11}$ kg / year. Because of the wide uncertainty in a number of these estimates, we can approximate the total methane emitted through natural processes as being equal to the total amount of methane emitted anthropogenically.

5.7.5 Natural Methane Sinks

The discerning reader will have noticed that the estimates presented above for annual methane emissions, which sum to roughly 7×10^{11} kg/year, outweigh the estimated annual increase in atmospheric methane concentration ($\approx 2 \times 10^{10}$

kg/year) by a factor of $f \times 10$. Where does this discrepancy come from? The answer lies in the fact that methane emissions are only one component of the global methane balance. The other component is given by methane sinks, which consist of natural processes that remove methane from the atmosphere. In the pre-industrial era, the balance between natural methane emissions from wetlands and other sources was roughly balanced with the rate of methane removal via sinks, resulting in an atmospheric methane level that remained roughly constant over long periods of time.

The dominant way that methane is removed from the atmosphere is via chemical reaction with the highly reactive hydroxyl radical, OH [239]. Due to this reaction, a given molecule of methane has a lifetime of ≈ 10 years in the atmosphere before it is oxidized by OH, producing H₂O and CO₂. This means that each year, $\approx 10\%$ of all atmospheric methane is removed via this chemical reaction. The lifetime of methane is independent of the total mass of methane in the atmosphere [239]; thus, as methane concentrations increase, the rate at which methane is removed increases proportionally. Currently, the mass of methane in the atmosphere is $\approx 5 \times 10^{12}$ kg. One tenth of this is $\approx 5 \times 10^{11}$ kg that is removed by reacting with OH each year. Notably, this adds $\approx 7.5 \times 10^{11}$ kg CO₂/year, or roughly 2% of anthropogenic CO₂ emissions (HuID: 47200).

5.7.6 The Methane Number

In sum, anthropogenic sources of methane are currently approximately equal to natural sources of methane emissions. While methane is much less prevalent in the atmosphere than carbon dioxide, it is an extremely potent gas. Our contributions to the increasing concentration of methane in the atmosphere have important implications to surface warming and the changing chemistry of the atmosphere.

Our tabulation above shows that the precise global balance of methane sources and sinks is difficult to estimate from first principles. The life cycle of methane from emission to breakdown or trapping is affected by complex atmospheric physics and other considerations. However, even the relatively coarse picture that we have laid out above gives us a sense for the magnitude of the impact that humans

THE METHANE NUMBER

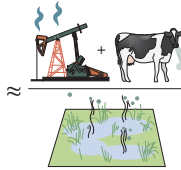
$$\text{Me} = \frac{\text{annual mass of anthropogenic CH}_4}{\text{annual mass of natural CH}_4} \approx \frac{\text{3.5} \times 10^{12} \text{ kg/yr}}{\text{3.6} \times 10^{12} \text{ kg/yr}} \approx 1$$


Figure 5.31: The Methane Number. While there are large uncertainties in both factors, the ratio of the annual mass of anthropogenic CH_4 emissions to emissions of CH_4 from natural sources is roughly 1.

have had on the global methane balance. We define the "Methane Number" as total anthropogenic methane emissions divided by total natural emissions, which is currently about one, as shown in Fig. 5.31. The total amount of methane emissions caused by humans annually ($\approx 3.5 \times 10^{12}$ kg/year) is roughly equal to the amount of methane emitted by all other natural sources ($\approx 3.6 \times 10^{12}$ kg/year). As noted above, it is important to consider how humans are also impacting the magnitude of natural emissions through our actions such as influencing the extent of wetlands.

5.8 Human Power Use

Next, we will shift focus from agricultural land use to human power consumption. From our cars to our homes, from our appliances to our factories, we require energy to power everything we use, build and consume. What are the major sources of human power consumption? How significant is the magnitude of our energy consumption in comparison to the natural sources of energy available to us? As we will see in this vignette, despite consuming an enormous amount of energy, the magnitude is dwarfed by the amount of energy that reaches the Earth's surface in the form of solar radiation.

Humanity has always used power to drive civilization. Historically, animal and human derived power used to be the major sources for tasks such as transportation and construction. Since the industrial revolution, however, both the amount of power consumed and the way we produce it has changed. Rather than using power derived from live animals, most of our power now comes from burning fossil fuels. This change in how we produce power has enabled humanity to consume an



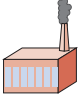
	CATEGORY	POWER	POWER PER CAPITA	FRACTION OF POWER USE
	 residential	$\approx 2 \text{ TW}$	$\approx 250 \text{ W / person}$	$\approx 15\%$
	 transportation	$\approx 2 \text{ TW}$	$\approx 250 \text{ W / person}$	$\approx 15\%$
	 industrial	$\approx 8 \text{ TW}$	$\approx 1000 \text{ W / person}$	$\approx 60\%$
	 commercial	$\approx 1 \text{ TW}$	$\approx 100 \text{ W / person}$	$\approx 8\%$

Figure 5.32: Breakdown of human power consumption by category. Total power consumption, per capita consumption, and fraction of the total are shown for each major category of human power use – residential, transportation, industrial, and commercial.

amount of power previously unfathomable. Today, we use power in nearly every aspect of our lives.

How much power does human civilization use? There are a few ways we can estimate its magnitude. One approach is to track energy directly, by looking at the consumption of fossil fuels and the generation of electricity through alternative energy sources. The U.S. Energy Information Agency (EIA) provides total energy consumption statistics for various sectors, and finds that human civilization consumes $\approx 20 \text{ TW}$ of power [HuID: 31373, 85317], or an average of $\approx 2500 \text{ W / person}$ [186]. The US average is significantly higher than this global average, at around $\approx 10^4 \text{ W / person}$ [240].

What are the major drivers of this power consumption? We can roughly divide all end-use power consumption into four sectors: transportation, industrial, commercial, and residential. Figure 5.32 shows the breakdown of human power consumption by each of these categories. Evidently, industrial dominates power consumption, accounting for $\approx 60\%$. Residential and transportation each make up $\approx 15\%$, and commercial uses account for the remaining $\approx 6\%$.

5.8.1 Transportation Power Use

The transportation sector involves all road vehicles like cars and trucks, as well as shipping, aviation, and rail. Before diving into the data, we can perform a back-of-the-envelope style calculation to try to get a sense of the size of this sector. For now, consider only passenger vehicles like cars. If we assume that each car is driven around $f \times 10$ km per day, and using a typical fuel economy of ≈ 10 km / liter (around 23 miles per gallon) [HuID: 30550], we arrive at around $\approx f$ liters of gasoline consumed per car per day, as outlined in Fig. 5.33. There are ≈ 1 billion cars on the road today [240]. This adds up to around $\approx f$ billion liters per day, or $f \times 10^6$ m³, globally. Gasoline is slightly less dense than water, weighing around 750 kg / m³, and has an energy density of around 4.5×10^7 J / kg. Thus, in total, the amount of energy consumed driving cars adds up to $\approx 10^{17}$ J / day, or around 1 TW. Dividing this by 8 billion people, we estimate that car usage alone is responsible for ≈ 125 W / person.

This estimate is fairly close to the truth. The entire transportation sector consumes around 2 TW, or roughly 250 W / person [186]. Approximately 45% of the total transportation power consumption is due to light-duty vehicles [241], which represents ≈ 110 W / person. Another 15% of the consumption is due to passenger aviation, buses and motorbikes, while the remaining $\approx 40\%$ is consumed by freight transportation [241]. This means that car power consumption is roughly equivalent to the power needed to ship and transport goods, and about three times the power used to fuel airplanes and buses. However, the amount of power consumed by industrial processes is much larger than the power used for transportation.

5.8.2 Industrial Power Use

The industrial sector involves energy used for manufacturing, mining, construction, and agriculture. This sector consumes ≈ 8 TW of power, or around 1 kW / person [186]. The breadth of activities covered by this sector is vast, and we cannot hope to provide even a cursory estimate for the scale of this sector in total. However, it is worthwhile to break down a few activities for the sake of gaining an intuition for the scales involved.

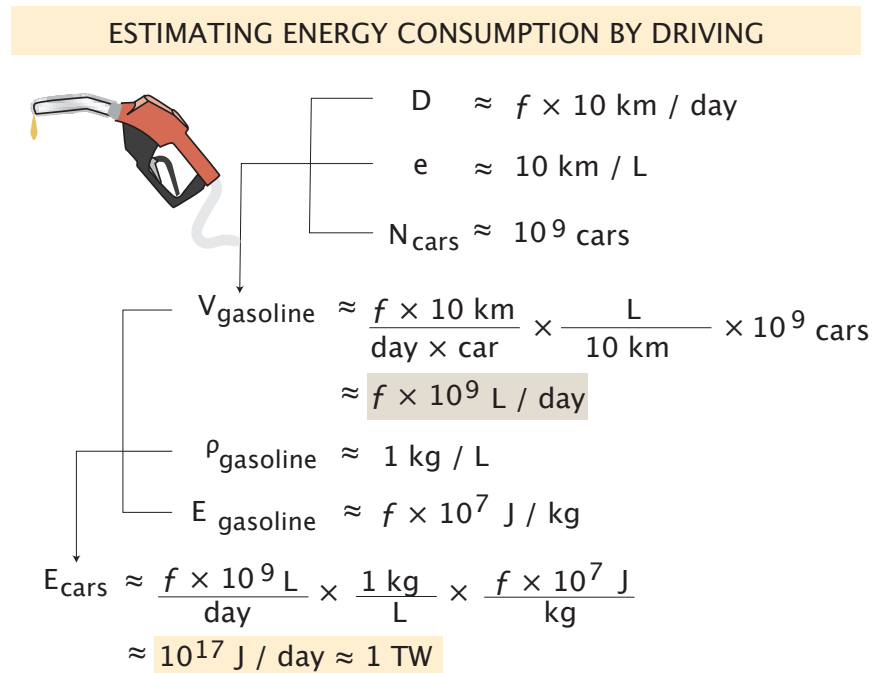


Figure 5.33: Order-of-magnitude estimate of the amount of power consumed by driving. We assume that on average, cars are driven $f \times 10$ km per day (which is about the average commute distance) to estimate the amount of energy consumed.

To start with, we can divide the industrial sector into three basic categories, each of which consume a similar amount of energy. The first is energy-intensive manufacturing, which covers most materials and chemical production, as well as the processing of food and petroleum. This consumes around 27% of the power used by this sector [241]. The second is non-energy-intensive manufacturing, which includes all manufacturing processes involving combining, but not producing, basic materials and parts. High-level industrial manufacturing like automotive and electronics production are covered in this category. This consumes around 39% of the power used by this sector [241]. Finally, we have the vast non-manufacturing sector, which covers things like agriculture, mining and construction. This consumes the remaining $\approx 34\%$ of the power consumed by this sector [241].

The first sector, energy-intensive manufacturing, is composed of a relatively small number of processes that make the fundamental building blocks for consumption or further manufacturing. We can thus get a sense of how much a few

essentials contribute to our total energy consumption. The production of paper, for instance, consumes $\approx 2.2 \times 10^{18}$ Joules of energy per year, or an average of $\approx 7 \times 10^{10}$ Watts [242]. This amounts to ≈ 8 Watts per person on the globe. Another component of high-energy manufacturing is food production. This consumes $\approx 1.2 \times 10^{18}$ Joules of energy per year, or an average of $\approx 4 \times 10^{10}$ Watts [242].

5.8.3 Residential Power Use

The final sector of energy usage we will consider is the residential sector, which encompasses the energy used to heat and power our homes and residences. This sector consumes ≈ 2 TW of power, or around 250 W / person. It is worth noting that residential power usage is severely unequal across the globe. For example, Americans consume roughly 2 kW [243] worth of household-related power per capita, or around eight times the global average.

What are the major drivers of this power usage? It would be difficult to estimate this from first principles based on considering the usage of a “typical” global human. The spectrum of individual energy consumption ranges from those in developed nations that enjoy unprecedented levels of comfort afforded by abundant energy, to the nearly one billion people on the planet without access to electricity. Rather than reconciling these two in the form of some attempt at locating the lifestyle of the “mean human,” we will instead examine the domestic consumption of a person living a “typical” North American lifestyle, for which the consumption patterns are most familiar to the authors.

Firstly, let us think about the energy consumption required by illumination. We can assume that a given person will typically have around one light bulb (give or take) on at all waking hours to illuminate the room that they currently reside in. Power consumption by light bulbs varies greatly; while a modern energy-saving LED light bulb may consume about 20 W, older incandescent designs often consume more. If we assume that most light bulbs are 100 Watt incandescents, and that a typical person is awake for ≈ 16 hours, this draws an average of $(16 \text{ hours}/24 \text{ hours}) \cdot (100 \text{ Watts}) \approx 70 \text{ W}$, or only about 3% of the residential energy usage of a typical American.

How about heating and cooling inside a house? The energy requirements for these systems are complex and difficult to estimate from first principles, depending on factors like quality of insulation, efficiency of air circulation, location-dependent weather patterns, and more.

For our illustrative example, consider that a typical American house has ≈ 2000 square feet, or ≈ 200 square meters. If we assume that this is a one-storey home roughly in a square, with a flat roof, and that a typical storey is about ≈ 3 meters high, this gives a roof area of $\approx 200 \text{ m}^2$ and a wall area of $(4 \text{ walls}) \times (3 \text{ m}) \times \sqrt{200 \text{ m}^2} \approx 200 \text{ m}^2$. The amount of heat that a house loses or gains through these surfaces is proportional to a) the area of the surfaces, b) the heat conductivity of each surface, and c) the temperature difference between the outside and the inside.

Suppose that an average house is maintained at a temperature on average ≈ 4 degrees Celsius different from the outside. Suppose also that each surface is made entirely of wood. Wood has a thermal conductivity of $\approx 0.2 \text{ W} / \text{m K}$ [244] meaning that a wall 1 meter thick will let 1 kWh worth of heat energy through per hour for every degree of difference between the inside and the outside temperature. If we assume that each surface is around 0.2 m thick, we obtain an average heat flux of $Q = (0.2 \text{ W} / \text{m K}) \times (4 \text{ K}) \times (200 \text{ m}^2) / (0.2 \text{ m}) \approx 800 \text{ Watts}$. This is $\approx 40\%$ of the budget for the average American household. While this model is heavily simplified and makes a number of unrealistic assumptions, detailed estimates show that heating and air conditioning contribute around 32% of the power consumption by the residential sector [245], which is similar to our simple estimate.

5.8.4 Energy Losses

There is one final category that accounts for the large discrepancy between end-use power consumption and total power consumption, which are losses incurred in electrical power generation and transmission due to inefficiencies in the electrical system. While there are modest losses associated with the transmission and distribution of electricity, the majority of this inefficiency is due to losses incurred converting the energy released at fossil fuel power plants into electricity. The size of these losses is around $\approx 8 \text{ TW}$, or roughly $1000 \text{ W} / \text{person}$ [186]. This amounts

to roughly 60% of all power expended in the generation of electricity.

5.8.5 Sizing up Incident Solar Power

We now turn to estimating the total amount of incident power the Earth's surface receives from the Sun. The Sun produces a staggering $\approx 3.8 \times 10^{26}$ watts of power [246], which radiates outwards uniformly in every direction. Some of this power hits the Earth as incoming shortwave radiation. This warms the planet which then emits outgoing longwave radiation back into space, diagrammed in Fig. 5.34 (A). To estimate the power output at a certain distance away from the Sun, we can assume that this power is distributed uniformly across the surface area of a sphere with radius equal to the distance from the Sun. The 'power density,' or power per unit area delivered by sunlight at that distance, is just the total power output of the Sun, divided by the area of this sphere (which is given by the formula $4\pi r^2$). The Earth lies roughly 10^{11} meters from the Sun, so the power density at this distance is $\approx f \times 10^3 \text{ W} / \text{m}^2$, as shown in Figure 5.34 (B). The total amount of power reaching the top of Earth's atmosphere is then given simply by this value of power density, multiplied by the area of the cross-section of the Earth that faces the sun, diagrammed in Fig. 5.34 (A). The total power reaching the Earth's atmosphere is then $\approx 10^{17} \text{ W}$. This corresponds to an average incident power of $\approx f \times 10^2 \text{ W} / \text{m}^2$ of Earth's surface. This power is not evenly distributed over time and space; polar latitudes receive less power than those at the equator, seasonal shifts change the amount of incident power a region receives, and of course at night no sunlight reaches Earth's surface.

5.8.5.1 Caveats in principle and in practice

While the cumulative amount of incident solar radiation is quite large, there are many factors that constrain the fraction of this energy we can gather, both in practice and even in principle. The first of these is that clouds and the atmosphere reflect around 30% of all incident light, and a further 20% is absorbed by the atmosphere as shown in Figure 5.35. This means that only $\approx 50\%$ of incident solar radiation reaches the Earth's surface for potential harvesting. Another significant

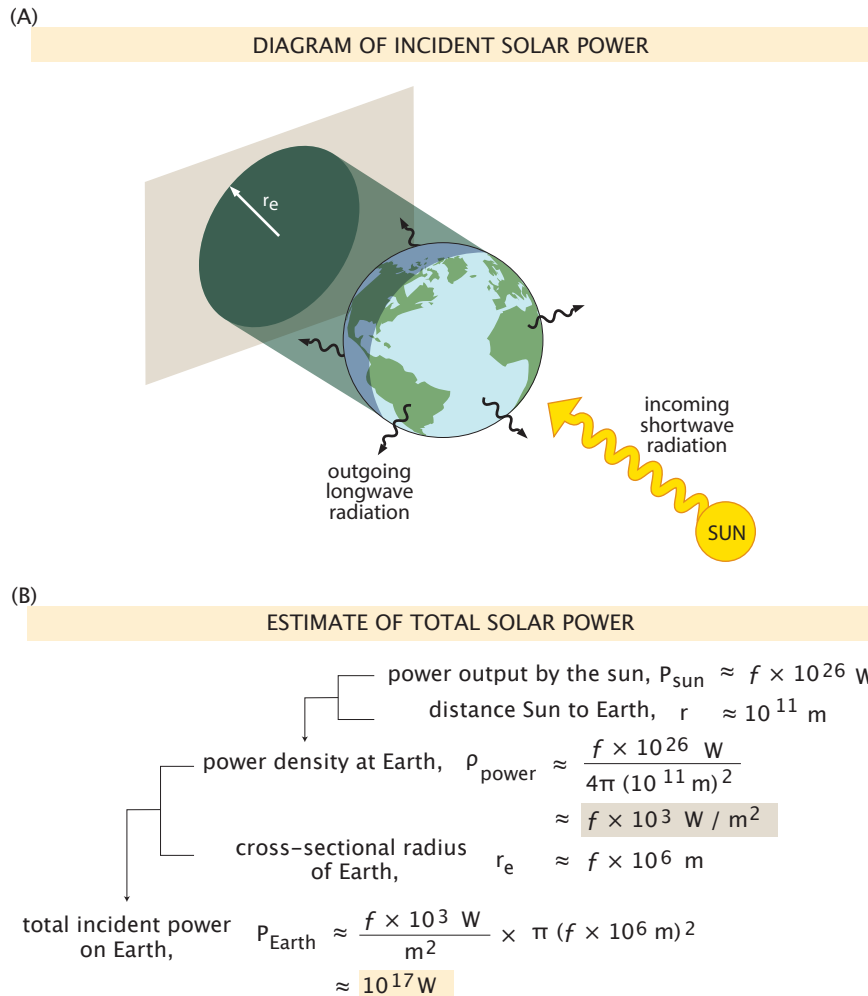


Figure 5.34: Earth's radiative energy balance. (A) The incoming radiation from the Sun is given by the product of the power density of sunlight and the area of Earth that the Sun sees. This incoming radiation is balanced by infrared, or longwave, radiation back to space. (B) Estimate of the incident power from the Sun on the Earth, comes to $\approx 10^{17}$ W. This is about ≈ 300 W / m^2 on average, but can vary from ≈ 1300 (with the Sun overhead) to 0 W / m^2 depending on the angle of the Sun with the surface.

limitation is that most of the Earth's area ($\approx 70\%$) is covered by oceans, which is a relatively difficult place to set up large solar collection projects. Additional constraints include the presence of crop land, natural ecosystems, and other land obstructions that would impede the development of solar harvesting projects in these areas.

In addition, there are constraints due to technological limitations. The dominant mode of solar power collection for the purpose of energy production are photovoltaic panels. Commercially available panels typically have efficiencies of $\approx 15\text{-}20\%$ [247]. There are limits on the performance of solar panels arising from the fundamental physics of photovoltaic devices, although depending on the design this can be as high as 60% or 80% [248]. The kind of panel that is most commonly used, known as a single-junction panel, has a theoretical maximum efficiency of only around $\approx 30\%$ [249].

Taking all of these multiples together, we see that the amount of solar energy we could harvest, in principle and in practice, is at most $(50\%) \times (30\%) \times (30\%) \approx 5\%$ of the incident power at the top of the atmosphere, and potentially much less than this based on other practical constraints on the placement of solar panels. Nevertheless, the enormous amount of energy received from the Sun means that there is still more than enough available solar capacity to power all of human civilization using solar energy.

5.8.6 The Solar Number

The solar number gives us a sense of the amount of energy that human civilization uses, compared with the total amount of energy available in principle from incident solar radiation. Human power usage measures up on average to about 19 TW , while total incident solar power measures up to $\approx 2 \times 10^5\text{ TW}$, giving a ratio of around 0.0001 , or one in ten-thousand, shown in Fig. 5.36. While there are many reasons why harnessing the vast quantity of energy delivered from the sun is not possible in practice, it gives a sense of the fact that there is a tremendous amount of available capacity in the form of incident solar power. This incident power offers us the option to replace our existing, largely fossil fuel based

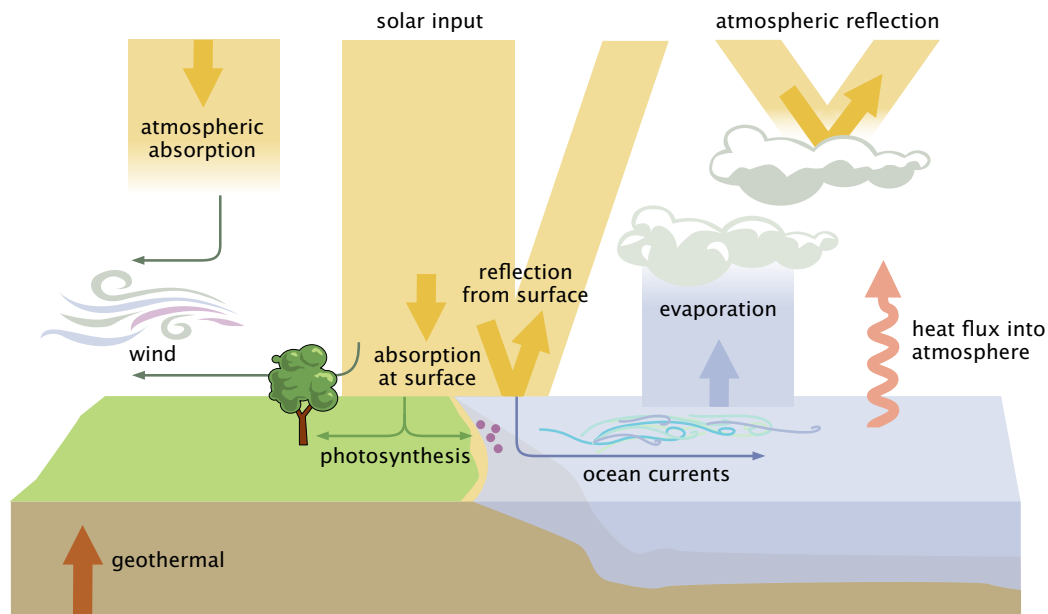


Figure 5.35: Decomposition of the incoming solar radiation into energy fluxes that fuel the dynamics of the atmosphere, ocean and life on Earth. From the $\approx 300 \text{ W} / \text{m}^2$ that the Earth receives from the Sun on average, about 30% is reflected back to space by clouds and the surface [250]. The atmosphere absorbs another $\approx 20\%$, such that only about $150 \text{ W} / \text{m}^2$ reach Earth's surface. This fraction of the total heats the surface, powers the ocean currents, the winds, and is used by photosynthetic organisms. The energy absorbed by the surface ultimately drives the heating of the atmosphere and the hydrological cycle by promoting water evaporation and cloud formation.

power generation infrastructure with photovoltaic generation capacity.

5.9 The Mass of Human-Made Materials

As we noted in the section on human power use, a large fraction of total power consumption is in the industrial sector, to create materials. In this section, we will look into some of the major human-made materials and their impact on the composition of mass on the Earth. Historically, the planet was covered in biomass, dominated by forests and other plant mass. However, humanity has progressively been altering the distribution of mass on the Earth by cutting down forests, rearing livestock, and cultivating crops. It is estimated that total biomass has decreased by about half due to the action of humans, primarily from forest loss [251]. Recently,

THE SOLAR NUMBER

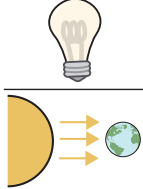
$$S_u = \frac{\text{annual human power usage}}{\text{annual incident solar power}} \approx \frac{\text{lightbulb}}{\text{sun}} \approx \frac{20 \text{ TW}}{2 \times 10^5 \text{ TW}} \approx 0.0001$$


Figure 5.36: The Solar Number. Human energy usage corresponds to only around one ten-thousandth of all of the incident power delivered by the sun to the top of Earth's atmosphere.

there has been enormous growth in the mass of inanimate human-made materials (termed anthropomass) produced, such as concrete, steel, and plastics, redefining how one should consider terrestrial mass [32].

Since around 1950, when cities and roads were built up in earnest, total anthropomass has been quickly growing. This mass is dominated by construction materials, including concrete, steel, and aggregates such as asphalt. Plastics and glass, despite composing materials we interact with everyday, make up only a small fraction of the total mass produced by humans. Aside from changing the surface composition of the Earth, producing these materials requires large amounts of energy and disposal at the end of their use presents challenges. Unlike biological mass, which is decomposed naturally and eventually becomes biomass for another organism or stored in the soil, human-made materials were designed to slowly if ever degrade. Thus, many enter landfills and accumulate as waste once out of use, and some ends up polluting rivers and oceans.

Thus, the generation of anthropomass is an important impact to consider in the evaluation of how humans are shaping the planet. In this section, we will estimate a couple forms of anthropomass – buildings and roads, and plastics, and size up their mass contribution against biomass on the Earth.

5.9.1 Buildings and Roads

One of the most apparent forms of anthropogenic mass is the various forms of buildings and roads that occupy our cities. We will estimate the mass of

construction materials that composes these buildings using our order-of-magnitude approach. In the "extent of human land use" section, we estimate the extent of urban land area to be $\approx 10^6 \text{ km}^2$ (10^{12} m^2). To estimate the mass of buildings and roads, we will assume that this urban area is covered in buildings and roads. The height of buildings vary widely, from single story homes (about 10 m) to many story office buildings (about 100 m), thus, to estimate the average height of a building, we will take the geometric average of these heights, $f \times 10 \text{ m}$, as shown in Figure 5.37 (A). Using these two estimates, we arrive at a volume of buildings of $\approx f \times 10^{13} \text{ m}^3$. However, not all of this volume is composed of construction materials, rather, much of the space is empty. We estimate that about 1 % of the total volume encompassed by buildings is filled with construction materials, and we assume that the density of these materials is a few times that of water. Thus, we estimate that the total mass of buildings and roads is $\approx 10^{15} \text{ kg}$ [Fig. 5.37 (A)].

This estimate is in good agreement with the data from Elhacham *et al.* [32]. In Figure 5.37 (B), we plot our estimate alongside the total mass of construction materials, including concrete, aggregates, bricks, and asphalt, which is approximately the mass of buildings and roads. Our estimate agrees well with the mass of construction materials in 2020. The data also show that the mass of construction materials has rapidly grown since the 1950s, similar to the rapid expansion of the urban population during that time. Plotting the mass of construction materials versus the urban population reveals that the increase in anthropomass has been somewhat linear with the growth in the urban population [Fig. 5.37 (C)].

5.9.2 Plastics

Another pervasive form of human-made mass is plastics, in the form of bags, water bottles, containers, to give just a few examples. Plastics were invented to be very resilient materials that are slowly, if ever, degraded naturally. Unlike biological materials, which are decomposed by bacteria and fungi and reused to build new biomass, plastics have to be melted down in an energy intensive process in order to be reused. Because of this, only a fraction of plastics are recycled, but rather most end up in landfills or in the oceans, where they negatively impact

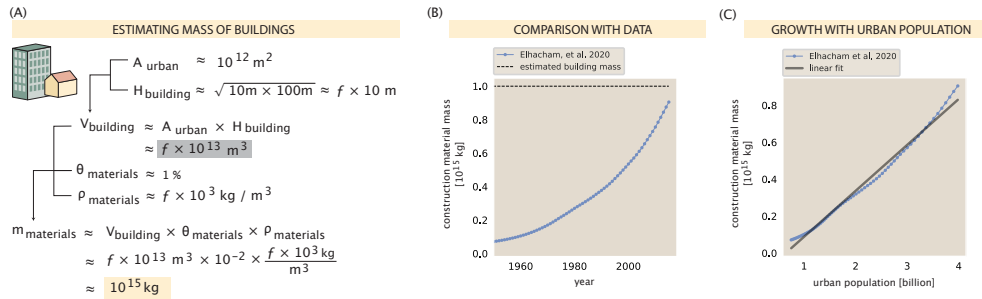


Figure 5.37: Estimate of the mass of buildings and roads. (A) Order-of-magnitude estimate of the mass of buildings and roads based on the extent of urban areas. (B) Comparison of our estimate with data from [32]. (C) Plot of the mass of construction materials versus the urban population and linear fit.

natural ecosystems.

We will estimate plastic production by considering our everyday use. The typical plastic water bottle has a mass of $\approx 100 \text{ g}$. We will assume that each person is responsible for the equivalent of 10 plastic water bottles a week, $\approx 1 \text{ kg}$. If a few billion people are each responsible for 1 kg of plastic, then we estimate yearly plastic production to be $\approx 10^{11} \text{ kg / yr}$, as shown in Figure 5.38 (A).

Comparing with the data, it is evident that plastic production has increased rapidly since the 1940s, when plastics were invented. Current annual production is $\approx 3 \times 10^{11} \text{ kg / yr}$, as estimated from the annual addition to plastic stocks in a material flows model of human-made material production [170] [Fig. 5.38 (B)]. Our estimate of $\approx 10^{11} \text{ kg / yr}$ of plastic production, is within a factor of a few of the current production rate, and about equal to the production rate in the 1990s. This implies that plastic production has about tripled over the past three decades.

5.9.3 The Anthropomass Number

In all, humans have generated an immense amount of materials in the last ≈ 70 years. How does this mass compare to global biomass? We can begin by making a few direct comparisons. The mass of all plastics generated is now about equal to the mass of all terrestrial and marine animals [Fig. 5.39 (A), top] [32]. Further, the mass of all buildings is about equal to the mass of all plants [Fig. 5.39

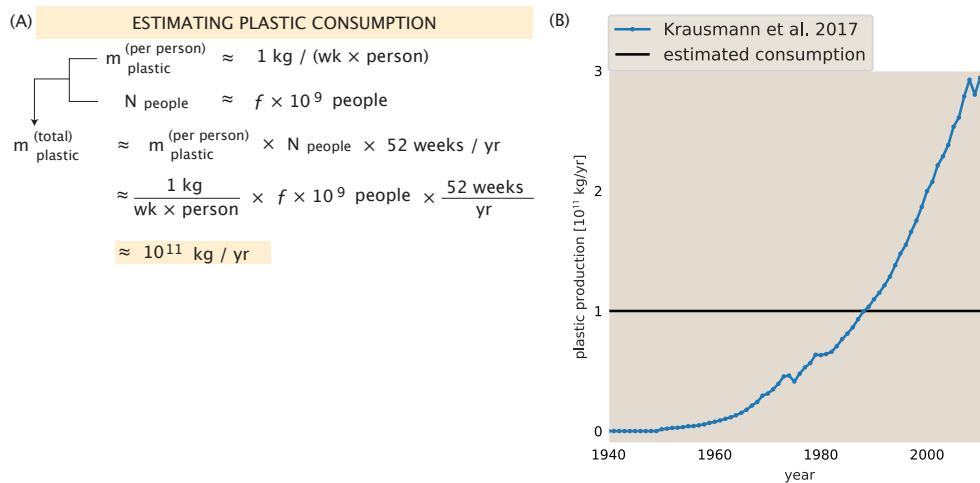


Figure 5.38: Order-of-magnitude estimate of plastic consumption. (A) Estimate of plastic consumption assuming each person is responsible for ≈ 1 kg of plastic per week. (B) Comparison of the estimate from (A) with data from [170] shows that our estimate is within a factor of a few of the actual mass of plastic production.

(A), bottom]. As shown in Figure 5.39 (B), waste, concrete, and aggregates are the largest contributors to total anthropomass. In all, currently, the mass of all human made materials is about equal to global biomass (by dry weight), as shown in Fig. 5.40. Biomass has remained relatively constant over the past 100 years while anthropomass has increased rapidly, driving this ratio from nearly zero in 1900 to one today and the anthropomass number will only continue to increase as we produce more and more materials, so that in all likelihood anthropomass will soon outweigh biomass [32].

5.10 Human Impact on CO₂ Emissions

Just as we discussed methane gas as a byproduct of human agricultural practices, in this section we will look into carbon dioxide emissions and how they related to human power consumption. To do so, we must first consider how power is generated. The coal-fired steam engine is an icon of the Industrial Revolution. James Watt's commercial release of a dual-chambered cylinder-condenser engine precipitated substantial changes in industries ranging from metallurgy to textile milling, and later enabled steam-powered locomotives, transforming trade and

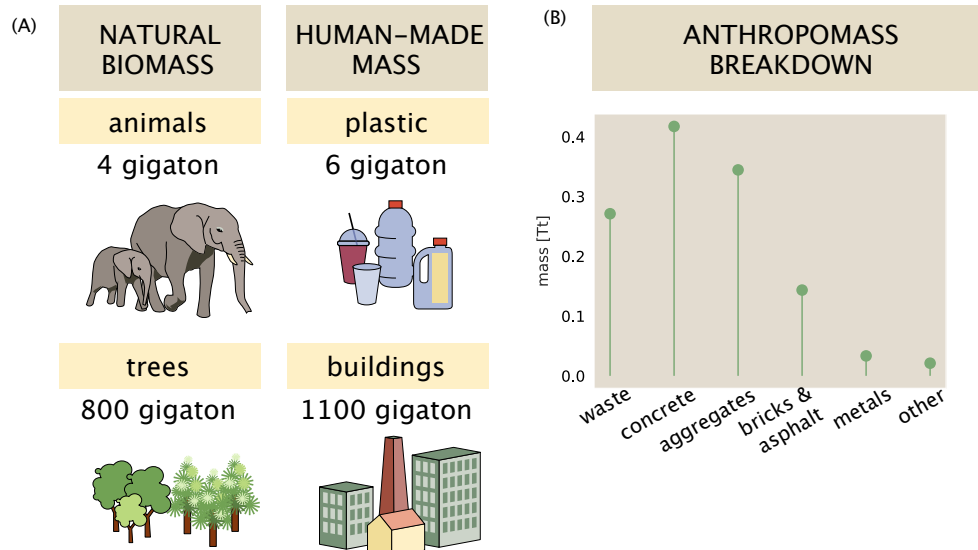


Figure 5.39: Comparison of anthropomass and biomass. (A) The mass of plastics are about equal to the mass of all terrestrial and marine animals on the planet, and the mass of buildings is currently greater than the mass of all trees. (B) Breaking down anthropomass by category shows that waste, concrete, and aggregates are the biggest contributors.

THE ANTHROPOMASS NUMBER

$$An = \frac{\text{total anthropomass}}{\text{total biomass}} = \frac{\text{[city icon]}}{\text{[trees icon] + [cows icon]}} \approx \frac{10^{15} \text{ kg}}{10^{15} \text{ kg}} \approx 1$$

Figure 5.40: The Anthropomass Number. We define the anthropomass number as the total mass of human-made materials divided by the mass of terrestrial wild mammals and birds, which is currently about one.

transportation forever. While there have been great improvements to the design and efficiency of steam engines, like Charles Parsons' invention of the steam turbine, the foundational premise is the same — some fuel (e.g. coal) is burned to transform water into steam, which then turns a wheel or a turbine, resulting in the generation of mechanical energy. Even today, a significant majority of our power ($\approx 80\%$) is produced by burning carbon-based fuels such as coal, oil, natural gas and wood (HuIDs: 10400; 42121; 49947; 31373). Combustion of these fuels produces CO₂ and other greenhouse gases like methane and nitrous oxide [252, 253]. Over the past ≈ 220 years, engines powered by the burning of carbon fuels have become ubiquitous, and the CO₂ released has measurably changed the chemistry of our atmosphere and oceans.

Figure 5.41 (A) shows the inferred (before 1958) and measured (1958 to present) atmospheric CO₂ concentrations over the past two millennia, revealing a rapid increase of $\approx 50\%$ since around 1800 [34]. The magnitude of CO₂ emissions by a growing human population is illustrated by the iconic “Keeling Curve” shown in Figure 5.41 (B & C), which plots the atmospheric CO₂ concentration measured from the North slope of Mauna Loa volcano in Hawaii since 1958 [254]. The curve is named for Charles David Keeling, who pioneered high-precision measurement of atmospheric CO₂ in the mid-20th century and set up infrared gas analysis at the Mauna Loa observatory [255]. Interestingly, Keeling observed oscillations in the CO₂ concentration, which reflect the seasonality of CO₂ uptake by plants. The Keeling Curve reveals that the current atmospheric CO₂ concentration is ≈ 420 parts per million (ppm), meaning that roughly one in every 2500 molecules in the atmosphere is CO₂. As shown in Figure 5.41 D, this corresponds to $\approx 3 \times 10^{15}$ kg CO₂ ($\approx 4 \times 10^{40}$ molecules of CO₂) currently in the atmosphere.

Contemplating the Keeling Curve alone would cause us to underestimate anthropogenic CO₂ emissions, however, as much of the CO₂ released is taken up by the oceans and terrestrial plants [Fig. 5.42 (A)]. This uptake partly offsets the amount of CO₂ released through anthropogenic activities, resulting in a smaller annual increase in atmospheric CO₂ than would occur without these natural sinks. While many natural processes produce and consume CO₂, they are mostly

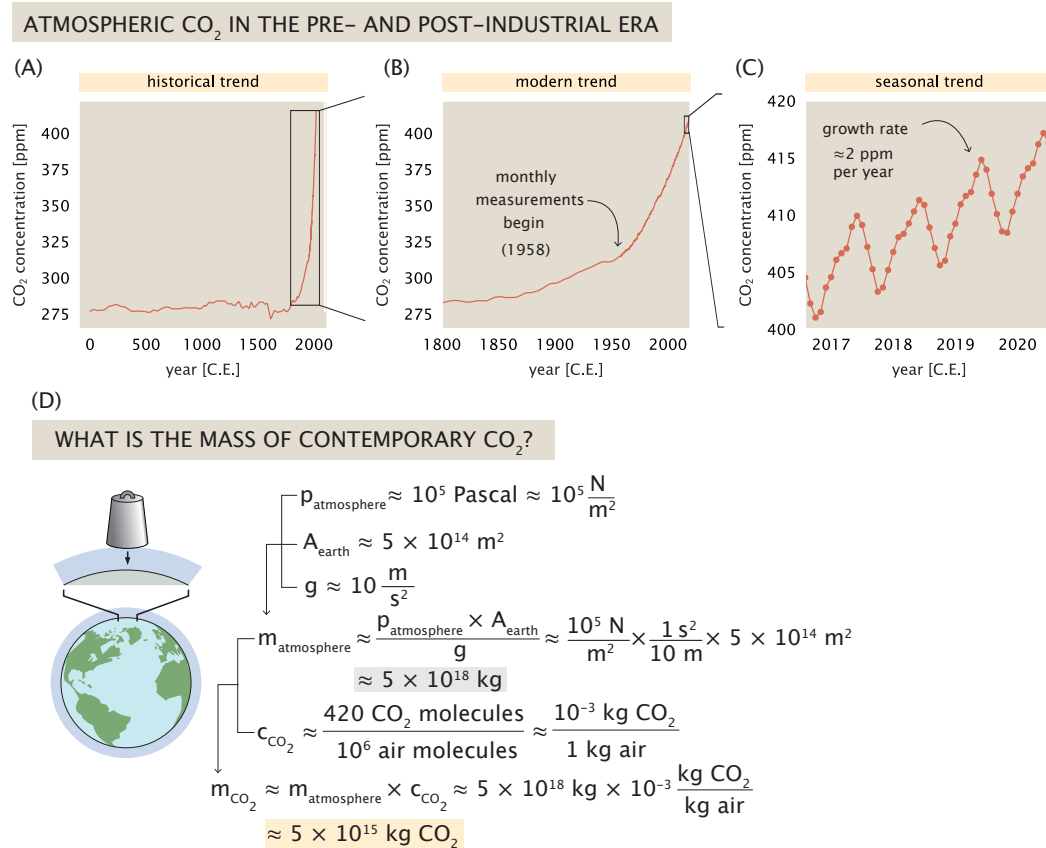


Figure 5.41: The concentration and mass of atmospheric CO₂. (A) The atmospheric CO₂ concentration over the past 2000 years. (B) The post-industrial atmospheric CO₂ concentration shown from 1800 through 2020. Data from 1800 to 1958 is inferred from ice core measurements at different depths (27). The data from 1958 through the present day is measured at the Mauna Loa outpost by infrared spectroscopy (4). (C) Recent measurements of the atmospheric CO₂ concentration from 2017-2020. Periodicity is due to CO₂ uptake by plants in the Northern hemisphere, which has greater land area than the Southern hemisphere. (D) Atmospheric pressure at sea level balances the weight of the atmosphere, so we can estimate the mass of the atmosphere using Earth's gravity and surface area. The total mass of atmospheric CO₂ is then estimated from its atmospheric concentration.

balanced with each other. For example, the net land sink is due to a balance between photosynthesis and respiration. Photosynthesis by terrestrial plants takes up $\approx 10^{14}$ kg CO₂ / year [227], but roughly 90% of this photosynthate is reconverted to CO₂ by the respiration of grazing organisms like bacteria and worms, resulting in a roughly 10-fold smaller net land sink of $\approx 10^{13}$ kg CO₂ per year (HuID: 74574; green bar in Fig. 5.42). After tallying the natural processes that produce and consume CO₂, a net natural sink remains, removing about half of human-caused CO₂ emissions from the atmosphere. This balance is represented by the purple bar in Figure 5.42 (B), showing that atmospheric CO₂ is currently increasing by $\approx 2 \times 10^{13}$ kg CO₂ per year (HuID: 98043).

Therefore, to estimate the magnitude of human CO₂ emissions, we must also consider natural CO₂ sinks. In this vignette, we will first estimate anthropogenic CO₂ emissions and then account for natural sinks to estimate the amount of CO₂ added to the atmosphere annually, as measured by the Keeling Curve.

5.10.1 Estimating Anthropogenic CO₂ Emissions

As diagrammed in Figure 5.43, the primary sources of anthropogenic CO₂ emissions are concrete production and the burning of fossil fuels (together $\approx 85\%$), and deforestation ($\approx 15\%$). To estimate anthropogenic CO₂ emissions due to fossil fuel burning, we will make several simplifying assumptions. First, we will assume that CO₂ emissions are proportional to total power consumption. This is a fair approximation since only about 20% of human energy consumption comes from sources other than fossil fuels as of 2018 (HuIDs: 29109, 29470, 31373). Further, though fossil fuels like coal, oil, and natural gas differ in their chemistry and practical uses, they all produce similar amounts of energy per unit of mass burned ($\approx f \times 10^7$ J / kg). Burning sugar (glucose) yields a similar amount of energy on a mass basis, as shown in Table 1. As the energy yield is related to the carbon content of fuels, we see that combustible fuels ranging from sugar to wood, natural gas and coal all release $\approx 10^{-7}$ kg CO₂ per Joule of energy released. Given this regularity in CO₂ and energy yields, we will simplify our calculations below by assuming that all fossil fuels produce $\approx 10^{-7}$ kg CO₂ per Joule.

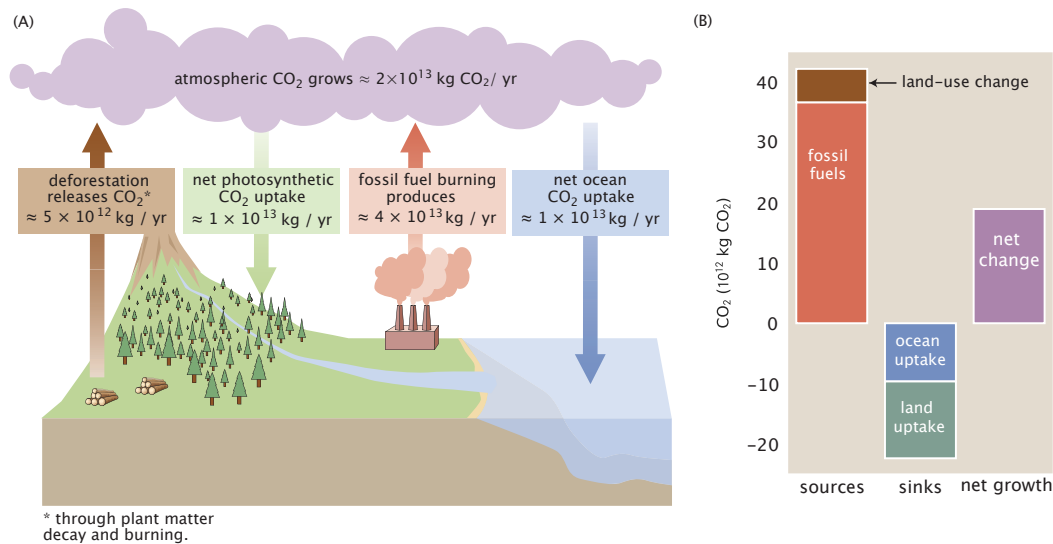


Figure 5.42: Major fluxes in the global carbon cycle. The fluxes driving the global carbon budget are due to a mix of biological, physical and human-related processes. The major processes drawing CO₂ out of the atmosphere are plant growth and oceanic uptake (HuIDs: 74574, 99089). Today, CO₂ emissions are due to two major anthropogenic sources: burning fossil fuels and making concrete ($\approx 85\%$), and destroying forests (so-called “land use change”, $\approx 15\%$) (HuIDs: 54608, 24789). (B) Summary of the major sources and sinks of CO₂, the result of which is a net growth of $\approx 20 \times 10^{12}$ kg CO₂ per year.

5.10.2 Estimating CO₂ Emissions from Fossil Fuels

Global CO₂ emissions from fossil fuels come from transportation, production of electricity, and industrial power use. To estimate global CO₂ emissions from burning fossil fuels, we will first estimate global use of fossil energy by thinking about cars and driving. Let’s assume that a few billion people have cars and drive like an American commuter, about 50 km / day. A typical car uses ≈ 10 L of gas per 100 km or ≈ 5 L for 50 km [260] [HuID: 30550]. Pure gasoline fuels have a density of ≈ 0.75 kg / L. Thus burning 1 kilogram (≈ 1.3 liters) of gasoline produces 45×10^6 J of energy or $\approx 35 \times 10^6$ J/L Fig. 5.43. Assuming that each car consumes 5 L of gasoline a day, we estimate an energy usage of $\approx f \times 10^{20}$ J per year, as shown in Figure 5.44 (A).

While the precise amount of CO₂ released per kilojoule of energy produced






SUBSTANCE	COMBUSTION ENERGY	CARBON CONTENT	CO ₂ YIELD
glucose 	$15.5 \times 10^3 \text{ kJ / kg}$	$\approx 40\%$	$\approx 10^{-4} \text{ kg CO}_2 / \text{kJ}$
dry wood 	$(15 - 18) \times 10^3 \text{ kJ / kg}$	$\approx 50\%$	$\approx 10^{-4} \text{ kg CO}_2 / \text{kJ}$
coal 	$(15 - 32) \times 10^3 \text{ kJ / kg}$	$\approx 70\% - 90\%$	$\approx (1-2) \times 10^{-4} \text{ kg CO}_2 / \text{kJ}$
natural gas 	$\approx 45 \times 10^3 \text{ kJ / kg}$	$\approx 75\%$	$\approx 6 \times 10^{-5} \text{ kg CO}_2 / \text{kJ}$
gasoline & diesel fuel 	$\approx 45 \times 10^3 \text{ kJ / kg}$	$\approx 80\%$	$\approx 7 \times 10^{-5} \text{ kg CO}_2 / \text{kJ}$

Figure 5.43: Energy and CO₂ produced from burning common fuel sources. Combustion energies are drawn from The Engineering Toolbox (“Wood and Bio Mass Heat,” n.d.) and the Handbook of Chemistry and Physics [256]. Carbon content can be calculated on a mass basis from the chemical formula of glucose (C₆H₁₂O₆) and is reported in the literature for dry wood [257] and coal [258]. Natural gas is composed mostly of methane (CH₄), which is 75% carbon by mass. Gasoline and diesel are both dominantly composed of alkanes (C_nH_{2n+2}) and alkenes (C_nH_{2n}) which have a carbon content of 75-85% [259]. As carbon has 12 atomic mass units and oxygen has 16, carbon makes up 27% of the mass of CO₂. Notice that the energies of combustion vary by only a factor of three [(1.5-4.5) × 10⁷ J/kg] and the amount of CO₂ produced per kilogram of fuel burned is about 10⁻⁷ kg CO₂/J for all fuels because the amount of energy produced by burning is determined in large part by the number of carbon atoms liberated from the fuel source.

depends on the type of fuel, typical yields are $\approx 10^{-7} \text{ kg CO}_2 / \text{J}$ (Fig. 5.43). Therefore, we estimate that $\approx f \times 10^{13} \text{ kg CO}_2$ are released every year from burning transportation fuels, as shown in Fig. 5.44 (A). Of course, treating the whole human population as American commuters is an overestimate of global commuting.

Another major source of fossil fuel emissions is production of electricity. We can estimate global electricity usage by assuming that each person uses a few lightbulbs worth of power ($f \times 100 \text{ W}$). Globally, this is equivalent to $\approx 10^{12} \text{ W}$ (Fig. 5.44 B). Our estimated value for electricity production is in good agreement with data that $\approx 3 \times 10^{19} \text{ J / yr}$ of energy is produced from fossil fuel burning power plants, as shown in Figure 5.44 (C). Assuming the same amount of carbon emissions per Joule as we used above for gasoline (Fig. 5.43), we arrive at an

estimate of $\approx f \times 10^{12}$ kg CO₂ / yr from electricity production, or about tenfold less than transportation emissions.

As shown in Fig. 5.44 C, our estimate of total fossil fuel emissions is remarkably similar to the values reported by much more detailed analyses [135], which give $\approx 3.5 \times 10^{13}$ kg CO₂ per year in 2018 (HuID: 54608). However, the U.S. Environmental Protection Agency notes that U.S. fossil fuel emissions are about equally due to transportation, electricity, and industrial processes. Thus, our estimate of transportation emissions must be an overestimate and our estimate of emissions due to electricity production must be an underestimate. However, both estimates are within a factor of a few of the data.

5.10.3 Estimating CO₂ Emissions from Land-Use Change

The other major anthropogenic source of CO₂ emissions is “land use change” – e.g. converting carbon-sequestering forests and prairies into farms and mines. Forests are removed by burning or cutting down the trees, which then decompose. Conversion of forests to other land uses, e.g. farmland, can also result in release of organic carbon sequestered in forest soils as CO₂ [261]. We can estimate the amount of CO₂ emitted from the extent of net annual anthropogenic deforestation and the carbon density of a typical forest. Net anthropogenic deforestation is the difference between gross deforestation (total forest removal) and forest regrowth in previously deforested areas. Annually, $\approx 10^{11}$ m² of land is deforested, roughly equal to the land area of Portugal or Cuba. The typical carbon content of a forest is $\approx f$ kg C / m² [153]. We therefore estimate that $\approx f \times 10^{11}$ kg C $\approx 10^{12}$ kg CO₂ is released annually due to land use changes, roughly in agreement with more detailed estimates arriving at $\approx 5 \times 10^{12}$ kg CO₂ / yr (HuID: 24789). Summing anthropogenic emissions due to fossil fuel burning and land use change, we estimate that total anthropogenic emissions are $\approx f \times 10^{13}$ kg CO₂ / yr, within a factor of a few of the reported value, $\approx 4 \times 10^{13}$ kg CO₂ / yr (HuID: 24789, 60670).

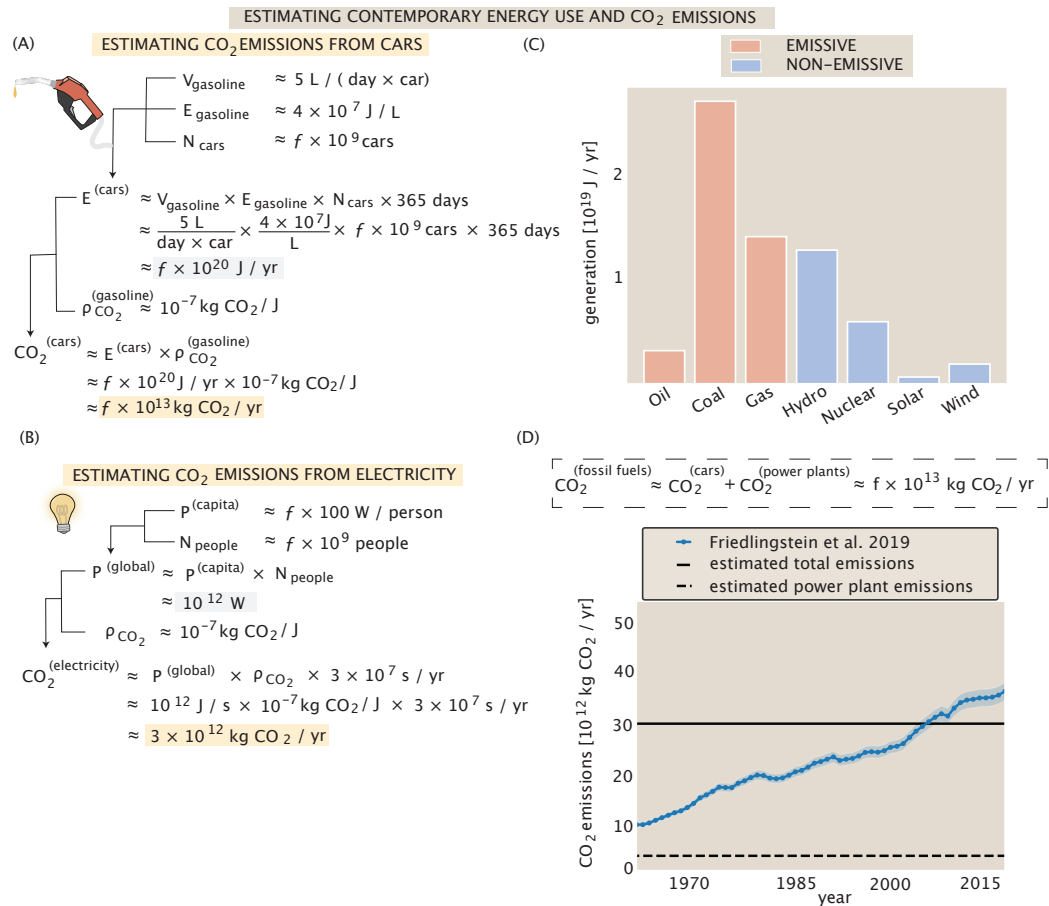


Figure 5.44: Estimating total human energy production and anthropogenic CO₂ emissions. (A) Estimate of global energy use and CO₂ emissions from car driving. We overestimate driving to account for other energy intensive activities. (B) Estimate of energy and CO₂ emissions due to electricity production. (C) Electricity production by fuel source [185]. Fossil fuel sources are colored red, non-emissive sources are in blue. (D) Comparison of estimated CO₂ emissions from fossil fuels with data from [135]. Shading indicates the uncertainty of the data.

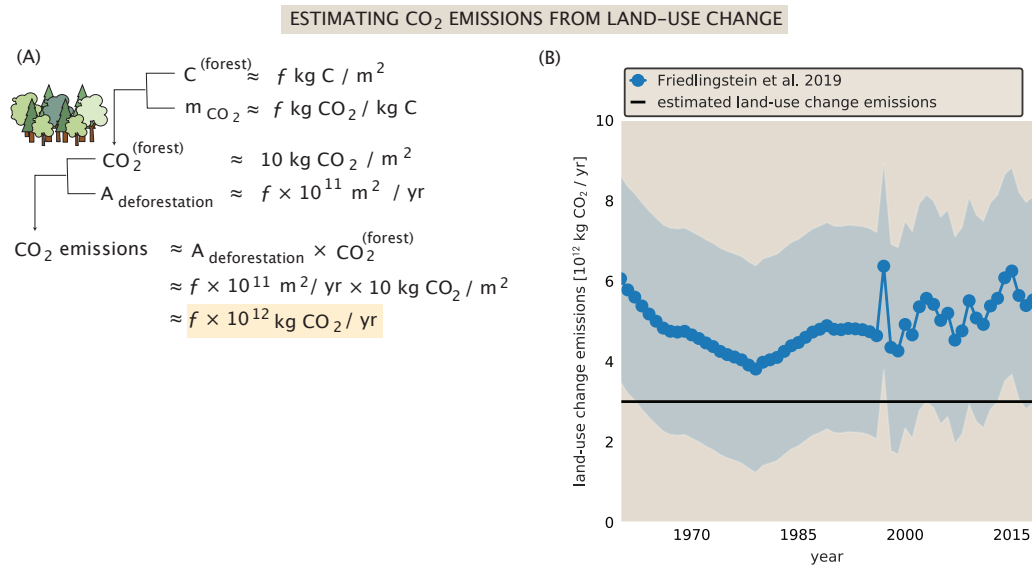


Figure 5.45: Estimation of CO₂ emissions from land-use change. (A) Estimate of CO₂ emitted from deforestation. We estimate emissions from deforestation based on the annual deforested area and the carbon density of forests. (B) Comparison of estimated land-use change emissions with data from [135].

5.10.4 How Much CO₂ Remains in the Atmosphere?

Can we reconcile the measurements summarized by the Keeling curve (Fig. 5.41) with our estimate of anthropogenic CO₂ emissions? Previously, we estimated that the total mass of CO₂ in the atmosphere is about 3×10^{15} kg CO₂, corresponding to ≈ 420 ppm. As shown in Fig. 5.41 C, atmospheric CO₂ has been increasing by ≈ 2 parts per million (ppm) per year for the last decade, which corresponds to $\approx 2 \times 10^{13}$ kg CO₂ / yr. This value is roughly half of total anthropogenic emissions $\approx 4 \times 10^{13}$ kg CO₂ / yr. The natural CO₂ sinks due to ocean uptake and photosynthesis by terrestrial plants must therefore take up the remaining half.

5.10.4.1 Plant CO₂ Uptake

How much CO₂ is sequestered by plant photosynthesis each year? We can estimate this quantity by considering the total amount of new plant growth

occurring each year. Somewhat confusingly, we will estimate the total amount of new plant growth by considering the destruction of forests, as we explain below.

Although humans have removed a tremendous amount of forest over the last ≈ 200 years and continue to deforest areas roughly the size of Portugal each year, satellite data suggest that total forested area may now be increasing through an even higher natural afforestation [262]. However, this suggestion is still contested by accounting methodologies that predict net deforestation [153]. Here we take a neutral stand and assume that net anthropogenic deforestation is roughly balanced by new forest growth. Assuming that new forest growth equals anthropogenic deforestation, we estimate $\approx 10^{11}$ m² / yr of new forest growth. Assuming the same forest carbon density as above, this amounts to $\approx 10^{12}$ kg CO₂ being taken out of the atmosphere by new forest growth every year.

This value is similar to others found in the scientific literature, which estimate that the sink due to forest growth represents between 5 and 60% of the net land sink of $\approx 1.3 \times 10^{13}$ kg of CO₂ [262, 263, 160]. The remainder is due to “CO₂ fertilization,” an observed effect wherein rising CO₂ levels enhance plant growth simply by making their primary growth substrate (CO₂) more available [264, 160]. Together, forest growth and CO₂ fertilization have a combined “global greening” effect [265], taking up $\approx 30\%$ of anthropogenic CO₂ emissions, each year [135].

5.10.5 Oceanic CO₂ uptake

CO₂ is fairly soluble in water, so increasing atmospheric CO₂ will tend to increase the “dissolved inorganic carbon” or DIC content of Earth’s oceans. Accurately computing the amount of CO₂ that makes it into the ocean requires a long quantitative detour through the chemistry of carbonates in seawater and the physics of mixing in the ocean. However, given the importance of this carbon sink, oceanographers have defined a single number known as the “Revelle factor” that we can use to estimate the magnitude of oceanic carbon uptake.

The Revelle factor, $R = ([\Delta\text{CO}_2]/[\text{CO}_2])(\Delta\text{DIC}/\text{DIC})$, is a unitless factor describing how proportional changes in atmospheric CO₂ ($[\Delta\text{CO}_2]/[\text{CO}_2]$) translates into changes in DIC concentration ($\Delta\text{DIC}/\text{DIC}$) in the surface ocean (top ≈ 10 m),

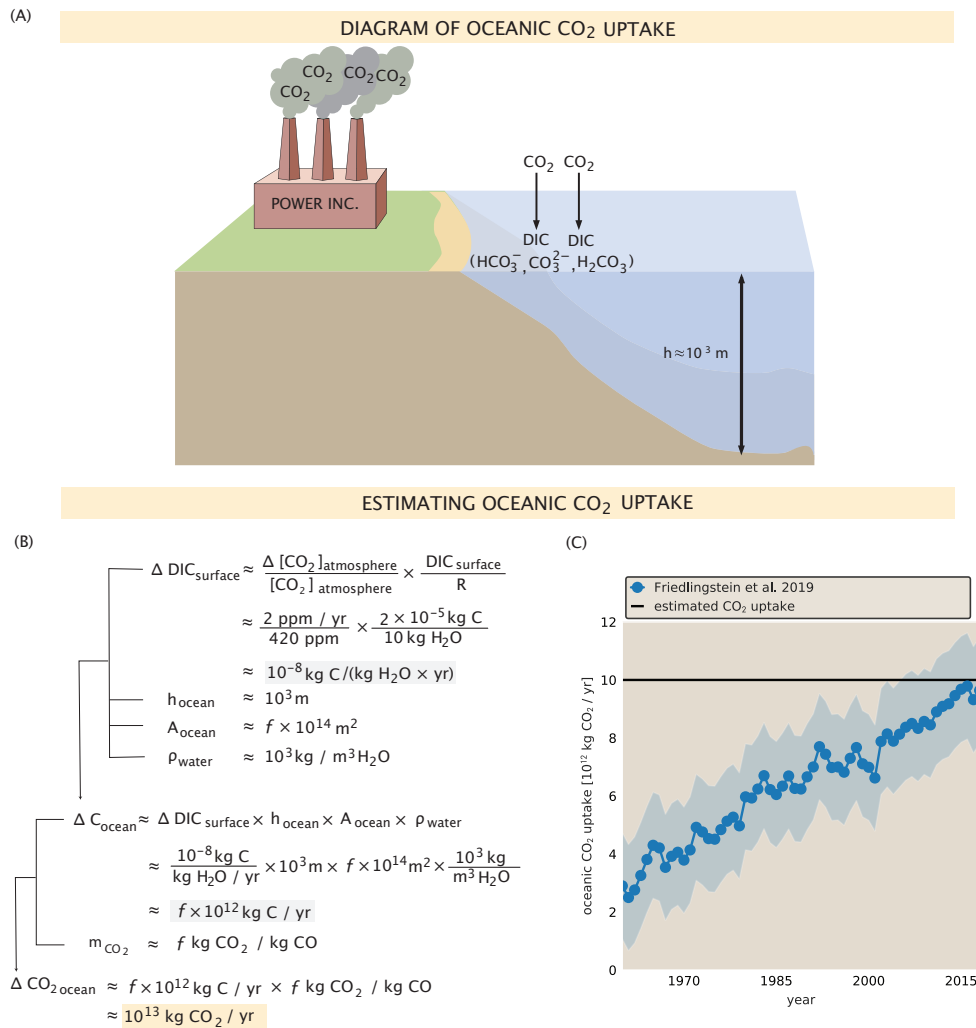


Figure 5.46: Estimating oceanic carbon uptake using the Revelle factor. (A) Diagram of CO₂ emissions being absorbed by the ocean to make dissolved inorganic carbon (DIC). (B) Estimate of the amount of dissolved carbon in the ocean using the Revelle Factor, current atmospheric CO₂ growth rate, and current CO₂ concentration in the atmosphere. (C) Comparison of estimate of carbon uptake by the oceans with data from [135].

as shown in Figure 5.46. The average value of the Revelle factor across Earth's oceans is ≈ 10 [266], meaning that a 10% change in atmospheric CO₂ becomes a $\approx 1\%$ change in surface DIC. Notably, Roger Revelle first emphasized the importance of this conversion factor in an article submitted in 1956, a few months after hiring a young Charles D. Keeling [267].

Using the current level of atmospheric CO₂ and the average DIC content of oceans, DIC $\approx 2 \times 10^{-5}$ kg C / kg of seawater [266], we estimate the change in inorganic carbon at the surface of the ocean is $\approx 10^{-8}$ kg C / kg of seawater per year, as shown in Figure 5.46 A and B. The total mass of oceanic carbon uptake is given by this concentration multiplied by the mass of the upper ocean. We estimate the mass of the upper ocean by multiplying the ocean surface area by the relevant height (≈ 1000 m), and the density of water. In this way, we estimate that $\approx 10^{13}$ kg CO₂ per year is taken up by the oceans.

This value is very close to estimates obtained from detailed studies of Earth's carbon inventory, which yield $\approx 10^{13}$ kg CO₂ / yr (HuID: 99089). This oceanic uptake represents $\approx 25\%$ of annual CO₂ emissions [135], meaning that the oceans and terrestrial plants together take up one-half of annual human emissions, as we estimated by comparing the atmospheric growth rate and anthropogenic emissions above.

5.10.6 How do Human CO₂ Emissions Compare to Natural Processes?

For the purpose of understanding the magnitude of human CO₂ emissions, we define the “CO₂ Number,” which compares annual anthropogenic CO₂ emissions to all natural sinks of atmospheric CO₂ 5.47. There are many natural sources of atmospheric CO₂, including respiration, outgassing from tropical warm waters, and volcanism [268]. However, most of these natural sources are approximately balanced by natural sinks of similar magnitude, such that the dominant sources of atmospheric CO₂ currently are human-caused. As shown in Figure 5.42, and estimated above, the dominant sources of anthropogenic CO₂ emissions are the burning of fossil fuels, the production of concrete, and deforestation. The major natural sinks of CO₂ are photosynthesis by land plants and uptake by Earth's oceans. As estimated above, these processes together absorb $\approx 50\%$ of anthropogenic CO₂ emissions, summing to $\approx 2 \times 10^{13}$ kg CO₂. The “CO₂ number” reflects the fact that human-caused CO₂ emissions currently outpace the natural processes removing CO₂ from the atmosphere by a factor of ≈ 2 . As a result, atmospheric CO₂ levels must rise over time. This is clearly captured by the Keeling curve (Figure 5.41),

THE CO₂ NUMBER

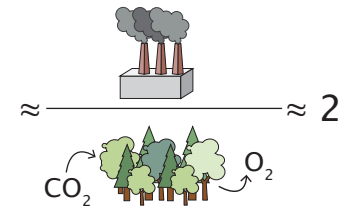
$$\text{CO}_2 = \frac{\text{annual mass of anthropogenic CO}_2}{\text{annual mass of naturally removed CO}_2} \approx \frac{\text{factories}}{\text{trees}} \approx 2$$


Figure 5.47: The CO₂ number. The CO₂ number compares the mass of annual anthropogenic CO₂ emissions to the mass of CO₂ that can be naturally removed by plants and the oceans annually. Anthropogenic emissions are currently twice the natural sink.

and leads to a strengthening of the greenhouse effect.

5.11 Conclusions

In this work, we explored several axes upon which humanity is a dominant force shaping the Earth, including water, land, power, and atmospheric chemistry. While detailed data on these impacts have been reported extensively, it can be difficult to have an intuition for the magnitude of human influence unless one is an expert in the field. We show that simple order-of-magnitude estimates can be useful for understanding these magnitudes, starting with simple common knowledge, such as how much chicken an average diet consists of. Furthermore, exploring the Anthropocene in this way reveals key connections between various axes of human influence, such as between land use, agriculture, water use, and nitrogen production.

Quantitatively understanding the myriad ways that humans are impacting the state of the Earth system is key to making plans for the future. Any attempts to alter the trajectory of the human-Earth system must be made in light of such an understanding. With the deluge of careful studies, gaining a full picture of this system can be daunting; we hope to facilitate wider knowledge about the human-Earth system, demonstrating that many of the impacts humanity is having can be understood simply by considering the lifestyle of an "average" human.

5.12 Supplementary Information

5.12.1 Choice of dimensionless ratios

We contextualized human impacts on several axes of the human-Earth system by formulating dimensionless ratios that compare the magnitude of anthropogenic influence to a natural analog. Often, several choices could have been made for the numerator and denominator for these numbers. In the following sections, we describe how we chose each ratio and alternative choices that could have been made.

The Terra Number. This number summarizes the extent of human terrestrial land use and is defined as total terrestrial anthropogenic land use divided by all terrestrial land area and is about one-third. Alternatively, the denominator for this number could have been total habitable land area, which would make the number larger. However, defining habitable land area is not simple and the extent of habitable area will be changing as a result of climate change as arctic areas warm and potentially become more habitable while deserts likely become less habitable. In order to have a consistent denominator for this number, therefore, we chose total terrestrial land area.

The Barnyard Number. The Barnyard Number reflects an impact that human agriculture is having on the biological composition of the Earth. We define this number as the total mass of livestock divided by the mass of all terrestrial mammals and birds. The choice to use biomass as the metric is derived from Bar-On, Phillips, and Milo [124], in which they quantified the biomass distribution of the Earth. They chose biomass rather than number of organisms because it gives a better comparison across taxa and an understanding of the distribution of elements such as carbon stocks. We agree with this and thus chose to use biomass to give a perspective on the magnitude of human livestock cultivation.

Another option for this number could have been to compare population sizes. However, an estimate of wild biomass is better constrained than estimating population sizes. In fact, wild mammals is the only category of organisms that Bar-On et al. do not provide an estimate for the abundance of individuals for [124]. The

estimate they report for wild birds is $\approx 10^{11}$ individuals, which is larger than the population of livestock by about an order of magnitude.

Another consideration for this number was which organisms to include in the denominator. As human livestock are predominantly mammals and poultry, using the mass of wild mammals and birds was an obvious choice to compare similar populations to each other. Other options could have been to use total biomass, which is dominated by plant mass and thus would not be a good representation of the transformation of animal mass. Another option could have been total animal biomass, which insect mass would have a large impact on. Estimating insect biomass, however, is difficult and improvements are still being made in this estimate.

The Water Number. The water number reflects the magnitude of human water use, compared to the amount of potentially available freshwater. There were two obvious choices for the numerator for this ratio: human water withdrawal, or human water consumption. The distinction between these two is subtle but important. Water withdrawal includes all water used by humans, including water that is taken from the natural source (rivers or lakes usually), used, and then may be returned. An example of this is water used to cool thermoelectric power plants. This water is typically withdrawn from its source, flowed through the cooling tower, and then returned to the source. In contrast, consumed water is used and not returned directly to its source, such as water used for irrigation. To assess the magnitude of human influence on the water cycle, we determined that withdrawn water is the more representative value as it encompasses all human water use and although withdrawn water includes water that is returned to the source, this use still can have important ecological effects as it is usually returned at a different temperature or purity.

For the denominator of the water number, we chose the volume of global river discharge. We chose this value because it represents the maximum amount of renewable freshwater available for human use [222]. In theory, the maximum renewable freshwater is the difference between terrestrial precipitation and evapotranspiration, but some of this goes into groundwater. Thus, river discharge, which

is a majority of the net between precipitation and evapotranspiration, and only loses about 10% to groundwater, is a better measure of the available freshwater.

The water number, then, comes to about 0.05 globally. While this may seem like a small impact, this number varies widely across the globe. As trade/transportation of water between countries and continents is not easy, the local amount of water use compared to available freshwater is key. To quantify this issue, the water scarcity index, which is defined as $R_W S = (W - S)/Q$, where W is total annual water withdrawal, S is the volume of water used from desalinated water, and Q is annual renewable freshwater resources [222]. Highly water stressed areas are considered those that have a scarcity index of 0.4 or larger; by this metric, about one third of the human population lives in highly stressed areas.

The Nitrogen Number. Nitrogen is a key element, required for vital biological molecules such as DNA and protein. The Nitrogen Number considers the impact that humanity is having on the availability of nitrogen and fluxes through the natural cycle. We define this number as the mass of synthetically fixed nitrogen through the Haber-Bosch process to the mass of terrestrial biologically fixed nitrogen. While there are multiple ways that human activities lead to nitrogen fixation, the Haber-Bosch process is the dominant, intentional form. Including the other anthropogenic contributions to nitrogen fixation would have little effect on the magnitude of the numerator. Similarly, biological nitrogen fixation is the major form of natural nitrogen fixation, including other values such as the amount fixed by lightning and rock weathering would have a modest effect on the magnitude of the denominator. Furthermore, a vast majority of the nitrogen fixed through the Haber-Bosch process goes towards fertilizer for crops; natural biological nitrogen fixation generally occurs in symbiosis with plants, thus the effect of these two inputs are similar in that they provide nitrogen to plants more or less directly.

Another major natural input to the nitrogen cycle is marine biological nitrogen fixation, which is about equal in magnitude to terrestrial biological nitrogen fixation. We chose not to use this value, though, because the terrestrial and marine cycles are relatively independent of each other. Some runoff of nitrogen from the land does become input into the ocean, but little goes the other way except through

consumption of seafoods. Thus, to contextualize human influence on the nitrogen cycle, we decided to only compare terrestrial fixation values.

The Methane Number. The Methane Number reveals the magnitude of human impact on emissions of methane, a key greenhouse gas. We define this number as the ratio of anthropogenic methane emissions to natural emissions. Unlike carbon dioxide (CO_2), which is naturally sequestered by the land and oceans, the major sink of methane is its chemical conversion to water and carbon dioxide. The extent of this reaction depends on the amount of methane in the atmosphere and therefore is not constant in magnitude. In addition, estimates of the magnitude of this sink are not very well constrained.

Another option for this number is the current atmospheric methane concentration divided by the pre-Industrial methane concentration. As described in the main text, this would be about two, meaning that the concentration of methane in the atmosphere has about doubled in the last ≈ 200 years. While informative, this ratio does not tell us about the annual contribution of human activities to methane fluxes. Formulating the methane number as we have reveals that human activities are approximately doubling the amount of methane released into the atmosphere annually. One other caveat to this number is that natural methane emissions are also being impacted by human activities. Natural emissions are mainly from wetlands, the extent of which is being altered by human land use and climate change. Estimating the magnitude of anthropogenic impacts on these emissions is challenging, however. Thus, we decided that using all natural emissions is an informative denominator while keeping in mind that this number is also impacted by humanity.

The Solar Number. The Solar Number compares the magnitude of current human power use to the magnitude of incident power from the Sun. This number is formulated to demonstrate that while human power use is large, the amount of power coming from the Sun is about 10,000 times larger. Currently, the vast majority of human power use is derived from burning fossil fuels; the Solar Number emphasizes the amount of power available from alternative sources.

While solar power is a major renewable source of energy, other renewable options exist, including wind and geothermal.

The Anthropomass Number. This number represents the contribution of human made mass to the composition of the land surface. While the land used to be dominated by biological materials, the mass of human made materials (anthropomass) has been increasing rapidly and now is a major impact on the surface of the Earth [32]. To summarize this impact, we compare the current anthropomass to the current biomass (by dry weight).

There are a few minor subtleties about the choice of the numerator. The numerator we chose is the mass of all human made materials (concrete, steel, plastics, lumber, etc). However, this does not account for the mass of humans or livestock which are also impacts of humanity on the mass composition of the Earth. Including these masses would not significantly change the anthropomass number, however, and are less well known than the mass of inanimate materials.

The CO₂ Number. The CO₂ Number is defined as the magnitude of human carbon dioxide emissions divided by the natural sinks. In this way, the number tells us how much excess carbon dioxide is going into the atmosphere. If the CO₂ Number were less than one, then the natural sinks could sequester all anthropogenic emissions. On the other hand, if greater than one, then emissions are outweighing the natural sinks, leading to the accumulation of carbon dioxide in the atmosphere. Currently, this number is about two, meaning that anthropogenic emissions are twice the magnitude of natural sinks.

Similar to the methane number, an alternative formulation would be the current atmospheric CO₂ concentration to the pre-Industrial concentration. While this formulation would reflect the integrated effect of human activity on carbon dioxide concentration over the past 200 years, it would not tell us about the current annual dynamics. Thus, we prefer the form of the CO₂ described in the main text, as it informs us about how atmospheric CO₂ is changing yearly.

Chapter 6

Conclusions and Future Work

The work done in this thesis explored the behavior of two dynamic non-equilibrium systems – the cellular cytoskeleton and the human-Earth system. Both of these systems are pushed away from equilibrium by forcing action of certain constituents. In the case of the cytoskeleton, motor proteins consume chemical fuel to walk along and reorganize filamentous proteins, forming highly ordered structures. Human action is pushing the Earth system away from the stable state of the Holocene, the previous geological epoch, to a new state with difficult to predict consequences. Although very different, both of these systems share that they are both pushed out of equilibrium by a forcing actor. Furthermore, in both cases, we would like to be able to predict the state they are moving towards and where they might end up. This requires not only having information about the initial conditions of the system, but also ways to quantitatively measure various properties over time.

The work described in this thesis took steps towards understanding motor-microtubule assemblies such as those vital for proper cellular functioning. Progress towards this end was limited previously by a lack of control over the creation of organized structures, rather relying on spontaneous self-assembly. In Chapter 1 of this thesis, we attempted to overcome this limitation by engineering a system wherein motor activity is dictated by shining light patterns on the sample. By dictating regions of motor activity in space and time, we were now able to prescribe

initial conditions and ask how varying those conditions affects the dynamics and the final steady-state achieved. In this way, we demonstrated controlled structure formation and measured how sizes and speeds of these structures depended on the size of the active region.

Then, we went beyond this and asked how the details of the motor-microtubule interactions affect the properties of the assemblies, as described in Chapter 2. In order to approach this question, we formed structures with the same activation region but varying the motor protein used. These motors differ in their speeds, processivities (how many steps they take before falling off the microtubule), directionalities, and forces they are able to exert. We found that both the size of steady-state asters and network contractile rates depend on these motor properties and developed a model for the distribution of motors within a steady-state aster that also depends on the microscopic motor characteristics. This work took a step towards understanding how the details of the interactions between motors and microtubules are translated to the much larger scale of the assemblies that they form.

Much work still remains to more fully understand the active system of motor proteins and their constituent filaments. For example, in the work described herein, we developed a model for the steady-state aster but not the dynamics of how the distributions of motors and microtubules change over time to form this structure. A model that predicts this time-evolution would represent a deeper grasp of how these collections form organized structures from disordered states. Another property we have yet to measure in this system is the spatiotemporal energy consumption by the motor proteins. Energy consumption is a key aspect of this system, pushing it out of equilibrium; knowing where, when, and how much energy is consumed throughout structure formation will yield further insight into this process.

On the planetary side, the work we have done has been towards creating a more complete picture of human impacts on the Earth and to make the scale of human influence intuitive and relatable for a wide audience. Human action is altering the state of the Earth system, resulting in changes to the composition of

life on the planet, chemistry of the atmosphere, pH and height of the oceans, and biogeochemistry. Myriad reports exist that go in depth on the human influence of one of these aspects of the human-Earth system. However, truly understanding the system requires a more holistic view that integrates this information and highlights the various interconnections between seemingly distinct axes. For example, knowing how much water humanity uses requires an understanding of agricultural practices, production of power, and domestic water uses.

To facilitate wider knowledge about quantitative measures of human influence on the Earth, we established the Human Impacts Database, described in Chapter 3. In this database, we curated and compiled > 300 values from governmental, intergovernmental, industrial, and scientific reports on human impacts. This searchable database makes finding values easy and fast, rather than having to read through long, complicated reports. We also provide links to the source of the value as well as a "tidy" version of the data, stored in our GitHub repository. We hope that this resource will be of use to researchers and anyone interested in learning about the Anthropocene.

In Chapter 4, we dove deeper into several key axes of human influence, using order-of-magnitude estimates to gain an intuition for each aspect and the scale of humanity's perturbation to the natural system. In doing so, we demonstrated that the magnitude of human action, although seemingly unfathomable, such as 10^{13} m² of crop land, or producing 20 TW of power, can be understood by considering simple facts such as how much plant mass an average diet consists of or how much energy is consumed by commuting to work. This work revealed some of the interconnections between human impacts on the Earth and that humanity now rivals the influence of natural processes in shaping our planet.

Bibliography

- [1] Marc Kirschner, John Gerhart, and Tim Mitchison. “Molecular “Vitalism””. en. In: *Cell* 100.1 (Jan. 2000), pp. 79–88. ISSN: 00928674. DOI: 10.1016/S0092-8674(00)81685-2.
- [2] A.M. Turing. “The Chemical Basis of Morphogenesis”. In: *Philosophical Transaction of the Royal Society of London. Series B, Biological Sciences* 237.641 (1952), pp. 37–72.
- [3] Zibo Chen and Michael B. Elowitz. “Programmable protein circuit design”. en. In: *Cell* 184.9 (Apr. 2021), pp. 2284–2301. ISSN: 00928674. DOI: 10.1016/j.cell.2021.03.007.
- [4] Rob Phillips. “Schrödinger’s What Is Life? at 75”. en. In: *Cell Systems* 12.6 (June 2021), pp. 465–476. ISSN: 24054712. DOI: 10.1016/j.cels.2021.05.013.
- [5] Alexandra F. Long, Pooja Suresh, and Sophie Dumont. “Individual kinetochore-fibers locally dissipate force to maintain robust mammalian spindle structure”. en. In: *Journal of Cell Biology* 219.8 (Aug. 2020), e201911090. ISSN: 0021-9525, 1540-8140. DOI: 10.1083/jcb.201911090.
- [6] David Oriola, Daniel J. Needleman, and Jan Brugués. “The Physics of the Metaphase Spindle”. en. In: *Annual Review of Biophysics* 47.1 (May 2018), pp. 655–673. ISSN: 1936-122X, 1936-1238. DOI: 10.1146/annurev-biophys-060414-034107.
- [7] Tamás Vicsek et al. “Novel Type of Phase Transition in a System of Self-Driven Particles”. In: *Physical Review Letters* 75.6 (1995), pp. 1226–1229. DOI: 10.1103/PhysRevLett.75.1226.

- [8] Eshel Ben-Jacob et al. “Generic modelling of cooperative growth patterns in bacterial colonies”. en. In: *Nature* 368.6466 (Mar. 1994), pp. 46–49. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/368046a0.
- [9] John Toner and Yuhai Tu. “Flocks, herds, and schools: A quantitative theory of flocking”. en. In: *Physical Review E* 58.4 (Oct. 1998), pp. 4828–4858. ISSN: 1063-651X, 1095-3787. DOI: 10.1103/PhysRevE.58.4828.
- [10] John Toner, Yuhai Tu, and Sriram Ramaswamy. “Hydrodynamics and phases of flocks”. en. In: *Annals of Physics* 318.1 (July 2005), pp. 170–244. ISSN: 00034916. DOI: 10.1016/j.aop.2005.04.011.
- [11] M. C. Marchetti et al. “Hydrodynamics of soft active matter”. en. In: *Reviews of Modern Physics* 85.3 (July 2013), pp. 1143–1189. ISSN: 0034-6861, 1539-0756. DOI: 10.1103/RevModPhys.85.1143.
- [12] I. Theurkauff et al. “Dynamic Clustering in Active Colloidal Suspensions with Chemical Signaling”. en. In: *Physical Review Letters* 108.26 (June 2012), p. 268303. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.108.268303.
- [13] S. C. Takatori and J. F. Brady. “Towards a thermodynamics of active matter”. en. In: *Physical Review E* 91.3 (Mar. 2015), p. 032117. ISSN: 1539-3755, 1550-2376. DOI: 10.1103/PhysRevE.91.032117.
- [14] Sho C. Takatori and John F. Brady. “Forces, stresses and the (thermo?) dynamics of active matter”. en. In: *Current Opinion in Colloid & Interface Science* 21 (Feb. 2016), pp. 24–33. ISSN: 13590294. DOI: 10.1016/j.cocis.2015.12.003.
- [15] F. J. Nédélec et al. “Self-organization of microtubules and motors”. In: *Nature* 389.6648 (1997), p. 305. ISSN: 1476-4687. DOI: 10.1038/38532.
- [16] Volker Schaller et al. “Polar patterns of driven filaments”. en. In: *Nature* 467.7311 (Sept. 2010), pp. 73–77. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature09312.
- [17] Stephen J. DeCamp et al. “Orientational order of motile defects in active nematics”. In: *Nature Materials* 14.11 (2015), p. 1110. ISSN: 1476-4660. DOI: 10.1038/nmat4387.
- [18] Daniel Needleman and Zvonimir Dogic. “Active matter at the interface between materials science and cell biology”. In: *Nature Reviews Materials* 2.9 (2017), p. 17048. ISSN: 2058-8437. DOI: 10.1038/natrevmats.2017.48.

- [19] Thomas Surrey et al. “Physical Properties Determining Self-Organization of Motors and Microtubules”. In: *Science* 292.5519 (2001), pp. 1167–1171. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1059758.
- [20] H. Y. Lee and M. Kardar. “Macroscopic equations for pattern formation in mixtures of microtubules and molecular motors”. In: *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics* 64.5 (2001), p. 056113. ISSN: 1539-3755. DOI: 10.1103/PhysRevE.64.056113.
- [21] S. Sankararaman, G. I. Menon, and P. S. Kumar. “Self-organized pattern formation in motor–microtubule mixtures”. In: *Physical Review E* 70.3 (2004), p. 031905.
- [22] F Julicher et al. “Active behavior of the Cytoskeleton”. en. In: *Physics Reports* 449.1-3 (Sept. 2007), pp. 3–28. ISSN: 03701573. DOI: 10.1016/j.physrep.2007.02.018.
- [23] Tyler D. Ross et al. “Controlling organization and forces in active matter through optically defined boundaries”. In: *Nature* 572.7768 (Aug. 2019), pp. 224–229. ISSN: 1476-4687. DOI: 10.1038/s41586-019-1447-1.
- [24] Woodward W. Fischer, James Hemp, and Jena E. Johnson. “Evolution of Oxygenic Photosynthesis”. en. In: *Annual Review of Earth and Planetary Sciences* 44.1 (June 2016), pp. 647–683. ISSN: 0084-6597, 1545-4495. DOI: 10.1146/annurev-earth-060313-054810.
- [25] Marielle Saunois et al. “The Global Methane Budget 2000–2017”. en. In: *Earth System Science Data* 12.3 (July 2020), pp. 1561–1623. ISSN: 1866-3516. DOI: 10.5194/essd-12-1561-2020.
- [26] Hanqin Tian et al. “A comprehensive quantification of global nitrous oxide sources and sinks”. en. In: *Nature* 586.7828 (Oct. 2020), pp. 248–256. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-020-2780-0.
- [27] NOAA Global Monitoring Laboratory Earth System Research Laboratories. *Trends in Atmospheric Carbon Dioxide*. 2020.
- [28] NOAA Global Monitoring Laboratory Earth System Research Laboratories. *GLobal CH₄ monthly means*. 2020.
- [29] Magdalena A. Balmaseda, Kevin E. Trenberth, and Erland Källén. “Distinctive climate signals in reanalysis of global ocean heat content”. en. In: *Geophysical Research Letters* 40.9 (May 2013), pp. 1754–1759. ISSN: 0094-8276, 1944-8007. DOI: 10.1002/grl.50382.

- [30] Lijing Cheng et al. “Improved estimates of ocean heat content from 1960 to 2015”. en. In: *Science Advances* 3.3 (Mar. 2017), e1601545. ISSN: 2375-2548. DOI: 10.1126/sciadv.1601545.
- [31] Yinon M. Bar-On and Ron Milo. “The Biomass Composition of the Oceans: A Blueprint of Our Blue Planet”. en. In: *Cell* 179.7 (Dec. 2019), pp. 1451–1454. ISSN: 00928674. DOI: 10.1016/j.cell.2019.11.018.
- [32] Emily Elhacham et al. “Global human-made mass exceeds all living biomass”. en. In: *Nature* 588.7838 (Dec. 2020), pp. 442–444. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-020-3010-5.
- [33] Paul J. Crutzen. “Geology of mankind”. en. In: *Nature* 415.6867 (Jan. 2002), pp. 23–23. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/415023a.
- [34] C. MacFarling Meure et al. “Law Dome CO₂, CH₄ and N₂O ice core records extended to 2000 years BP”. en. In: *Geophysical Research Letters* 33.14 (2006), p. L14810. ISSN: 0094-8276. DOI: 10.1029/2006GL026152.
- [35] T. M. Lenton et al. “Tipping elements in the Earth’s climate system”. en. In: *Proceedings of the National Academy of Sciences* 105.6 (Feb. 2008), pp. 1786–1793. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.0705414105.
- [36] Yongyang Cai, Timothy M. Lenton, and Thomas S. Lontzek. “Risk of multiple interacting tipping points should encourage rapid CO₂ emission reduction”. en. In: *Nature Climate Change* 6.5 (May 2016), pp. 520–525. ISSN: 1758-678X, 1758-6798. DOI: 10.1038/nclimate2964.
- [37] Johan Rockström et al. “A safe operating space for humanity”. en. In: *Nature* 461.7263 (Sept. 2009), pp. 472–475. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/461472a.
- [38] Yue Qin et al. “Flexibility and intensity of global water use”. en. In: *Nature Sustainability* 2.6 (June 2019), pp. 515–523. ISSN: 2398-9629. DOI: 10.1038/s41893-019-0294-2.
- [39] M. C. Marchetti et al. “Hydrodynamics of soft active matter”. In: *Rev. Mod. Phys.* 85 (2013), pp. 1143–1189. DOI: 10.1103/RevModPhys.85.1143.
- [40] Sophie Dumont and Manu Prakash. “Emergent mechanics of biological structures”. In: *Molecular Biology of the Cell* 25.22 (2014), pp. 3461–3465. ISSN: 1059-1524, 1939-4586. DOI: 10.1091/mbc.E14-03-0784.
- [41] Tim Sanchez et al. “Spontaneous motion in hierarchically assembled active matter”. In: *Nature* 491.7424 (2012), p. 431. ISSN: 1476-4687. DOI: 10.1038/nature11591.

- [42] Kun-Ta Wu et al. “Transition from turbulent to coherent flows in confined three-dimensional active fluids”. In: *Science* 355.6331 (2017), eaal1979. ISSN: 0036-8075, 1095-9203.
- [43] Antoine Bricard et al. “Emergence of macroscopic directed motion in populations of motile colloids”. In: *Nature* 503.7474 (2013), pp. 95–98. ISSN: 1476-4687. DOI: 10.1038/nature12673.
- [44] François Nédélec, Thomas Surrey, and A. C. Maggs. “Dynamic Concentration of Motors in Microtubule Arrays”. In: *Physical Review Letters* 86.14 (2001), pp. 3192–3195. DOI: 10.1103/PhysRevLett.86.3192.
- [45] Felix C. Keber et al. “Topology and dynamics of active nematic vesicles”. In: *Science* 345.6201 (2014), pp. 1135–1139. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1254784.
- [46] Susumu Aoyama, Masahiko Shimoike, and Yuichi Hiratsuka. “Self-organized optical device driven by motor proteins”. In: *Proceedings of the National Academy of Sciences* 110.41 (2013), pp. 16408–16413. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1306281110.
- [47] Gurkan Guntas et al. “Engineering an improved light-induced dimer (iLID) for controlling the localization and activity of signaling proteins”. In: *Proceedings of the National Academy of Sciences of the United States of America* 112.1 (2015), pp. 112–117. ISSN: 1091-6490. DOI: 10.1073/pnas.1417910112.
- [48] Matthias Schuppler et al. “Boundaries steer the contraction of active gels”. In: *Nature Communications* 7 (2016), p. 13120. ISSN: 2041-1723. DOI: 10.1038/ncomms13120.
- [49] Julio M Belmonte, Maria Leptin, and François Nédélec. “A theory that predicts behaviors of disordered cytoskeletal networks”. In: *Molecular Systems Biology* 13.9 (2017), p. 941. DOI: 10.15252/msb.20177796. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.15252/msb.20177796>.
- [50] Peter J Foster et al. “Active contraction of microtubule networks”. In: *Elife* 4 (2015), e10837.
- [51] Matthew C. Good et al. “Cytoplasmic Volume Modulates Spindle Size During Embryogenesis”. In: *Science* 342.6160 (2013), pp. 856–860. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1243147.
- [52] Orion D Weiner et al. “An Actin-Based Wave Generator Organizes Cell Motility”. In: *PLOS Biology* 5.9 (2007), pp. 1–11. DOI: 10.1371/journal.pbio.0050221.

- [53] Margaret L. Gardel et al. “Mechanical Integration of Actin and Adhesion Dynamics in Cell Migration”. In: *Annual Review of Cell and Developmental Biology* 26.1 (2010), pp. 315–333.
- [54] William E Theurkauf. “Premature microtubule-dependent cytoplasmic streaming in cappuccino and spire mutant oocytes”. In: *Science* 265.5181 (1994), pp. 2093–2096.
- [55] Sujoy Ganguly et al. “Cytoplasmic streaming in *Drosophila* oocytes varies with kinesin activity and correlates with the microtubule cytoskeleton architecture”. In: *Proceedings of the National Academy of Sciences* 109.38 (2012), pp. 15109–15114.
- [56] Raymond E Goldstein, Idan Tuval, and Jan-Willem van de Meent. “Microfluidics of cytoplasmic streaming and its implications for intracellular transport”. In: *Proceedings of the National Academy of Sciences* 105.10 (2008), pp. 3663–3667.
- [57] Knut Drescher et al. “Fluid dynamics and noise in bacterial cell–cell and cell–surface scattering”. In: *Proceedings of the National Academy of Sciences* 108.27 (2011), pp. 10940–10945.
- [58] Knut Drescher et al. “Direct Measurement of the Flow Field around Swimming Microorganisms”. In: *Phys. Rev. Lett.* 105 (2010), p. 168101. DOI: 10.1103/PhysRevLett.105.168101.
- [59] Bing He et al. “Apical constriction drives tissue-scale hydrodynamic flow to mediate cell elongation”. In: *Nature* 508 (2014), p. 392.
- [60] Tamar Shinar et al. “A model of cytoplasmically driven microtubule-based motion in the single-celled *Caenorhabditis elegans* embryo”. In: *Proceedings of the National Academy of Sciences* 108.26 (2011), pp. 10508–10513. ISSN: 0027-8424. DOI: 10.1073/pnas.1017369108. eprint: <https://www.pnas.org/content/108/26/10508.full.pdf>.
- [61] Matthäus Mittasch et al. “Non-invasive perturbations of intracellular flow reveal physical principles of cell organization”. In: *Nature Cell Biology* 20.3 (2018), pp. 344–351. ISSN: 1476-4679. DOI: 10.1038/s41556-017-0032-9.
- [62] Nefeli Georgoulia. *Tubulin Polymerization with GTP/GMPCPP/Taxol*. 2012.
- [63] Melissa K Gardner, Marija Zanic, and Jonathon Howard. “Microtubule catastrophe and rescue”. In: *Current Opinion in Cell Biology* 25.1 (2013). Cell architecture, pp. 14–22. ISSN: 0955-0674. DOI: <https://doi.org/10.1016/j.ceb.2012.09.006>.

- [64] A. W. C. Lau, A. Prasad, and Z. Dogic. “Condensation of isolated semi-flexible filaments driven by depletion interactions”. In: *EPL (Europhysics Letters)* 87.4 (2009), p. 48006.
- [65] Arthur Edelstein et al. “Computer Control of Microscopes Using μ Manager”. In: *Current Protocols in Molecular Biology* 92.1 (2010), pp. 14.20.1–14.20.17.
- [66] Martin Lenz et al. “Contractile Units in Disordered Actomyosin Bundles Arise from F-Actin Buckling”. In: *Phys. Rev. Lett.* 108 (23 June 2012), p. 238107. DOI: 10.1103/PhysRevLett.108.238107.
- [67] Roderick P. Tas et al. “Guided by Light: Optical Control of Microtubule Gliding Assays”. In: *Nano Letters* 0.0 (2018). PMID: 30449112, null. DOI: 10.1021/acs.nanolett.8b03011.
- [68] Hideki Nakamura et al. “Intracellular production of hydrogels and synthetic RNA granules by multivalent molecular interactions”. In: *Nature materials* 17.1 (2018), p. 79.
- [69] Heath E Johnson et al. “The spatiotemporal limits of developmental Erk signaling”. In: *Developmental cell* 40.2 (2017), pp. 185–192.
- [70] Hayretin Yumerefendi et al. “Light-induced nuclear export reveals rapid dynamics of epigenetic modifications”. In: *Nature chemical biology* 12.6 (2016), p. 399.
- [71] Maria M. Tirado and José García de la Torre. “Translational friction coefficients of rigid, symmetric top macromolecules. Application to circular cylinders”. In: *The Journal of Chemical Physics* 71.6 (1979), pp. 2581–2587. DOI: 10.1063/1.438613.
- [72] K Svoboda and S M Block. “Biological Applications of Optical Forces”. In: *Annual Review of Biophysics and Biomolecular Structure* 23.1 (1994), pp. 247–285.
- [73] Mohammed Mahamdeh, Citlali Perez Campos, and Erik Schäffer. “Under-Filling Trapping Objectives Optimizes the use of Available Laser Power in Optical Tweezers”. In: *Optics express* 19 (2011), pp. 11759–68. DOI: 10.1364/OE.19.011759.
- [74] HG Maas, A Gruen, and D Papantoniou. “Particle tracking velocimetry in three-dimensional flows”. In: *Experiments in Fluids* 15.2 (1993), pp. 133–146.
- [75] Thomas Schmidt et al. “Imaging of single molecule diffusion”. In: *Proceedings of the National Academy of Sciences* 93.7 (1996), pp. 2926–2929.

- [76] Jörn Dunkel et al. “Fluid dynamics of bacterial turbulence”. In: *Physical review letters* 110.22 (2013), p. 228102.
- [77] John Happel and Howard Brenner. *Low Reynolds number hydrodynamics: with special applications to particulate media*. Vol. 1. Springer Science & Business Media, 2012.
- [78] Allen T Chwang and T Yao-Tsu Wu. “Hydromechanics of low-Reynolds-number flow. Part 2. Singularity method for Stokes flows”. In: *Journal of Fluid Mechanics* 67.4 (1975), pp. 787–815.
- [79] Sangtae Kim and Seppo J Karrila. *Microhydrodynamics: principles and selected applications*. Courier Corporation, 2013.
- [80] Ronald L Panton. *Incompressible flow*. John Wiley & Sons, 2006.
- [81] Giacomo Frangipane et al. “Dynamic density shaping of photokinetic *E. coli*”. In: *eLife* 7 (Aug. 14, 2018). Ed. by Raymond E Goldstein and Naama Barkai, e36608. ISSN: 2050-084X. DOI: 10.7554/eLife.36608.
- [82] Jochen Arlt et al. “Painting with light-powered bacteria”. In: *Nature Communications* 9.1 (Feb. 22, 2018), p. 768. ISSN: 2041-1723. DOI: 10.1038/s41467-018-03161-8.
- [83] Gabriel Popkin. “The physics of life”. en. In: *Nature* 529.7584 (Jan. 2016), pp. 16–18. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/529016a.
- [84] Bill Wickstead and Keith Gull. “A “Holistic” Kinesin Phylogeny Reveals New Kinesin Families and Predicts Protein Functions”. en. In: *Molecular Biology of the Cell* 17.4 (Apr. 2006), pp. 1734–1743. ISSN: 1059-1524, 1939-4586. DOI: 10.1091/mbc.e05-11-1090.
- [85] Barbara J. Mann and Patricia Wadsworth. “Kinesin-5 Regulation and Function in Mitosis”. en. In: *Trends in Cell Biology* 29.1 (Jan. 2019), pp. 66–79. ISSN: 09628924. DOI: 10.1016/j.tcb.2018.08.004.
- [86] S. A. Endow et al. “Mutants of the *Drosophila ncd* microtubule motor protein cause centrosomal and spindle pole defects in mitosis”. eng. In: *Journal of Cell Science* 107 (Pt 4) (Apr. 1994), pp. 859–867. ISSN: 0021-9533.
- [87] Nenad Pavin and Iva M. Tolić. “Mechanobiology of the Mitotic Spindle”. en. In: *Developmental Cell* 56.2 (Jan. 2021), pp. 192–201. ISSN: 15345807. DOI: 10.1016/j.devcel.2020.11.003.
- [88] Maya I. Anjur-Dietrich, Colm P. Kelleher, and Daniel J. Needleman. “Mechanical Mechanisms of Chromosome Segregation”. en. In: *Cells* 10.2 (Feb. 2021), p. 465. ISSN: 2073-4409. DOI: 10.3390/cells10020465.

- [89] R. Grover et al. “Transport efficiency of membrane-anchored kinesin-1 motors depends on motor density and diffusivity”. In: *Proceedings of the National Academy of Sciences* 113.46 (2016), E7185–E7193.
- [90] N Hirokawa et al. “Kinesin associates with anterogradely transported membranous organelles in vivo.” en. In: *Journal of Cell Biology* 114.2 (July 1991), pp. 295–302. ISSN: 0021-9525, 1540-8140. DOI: 10.1083/jcb.114.2.295.
- [91] Johanna Roostalu et al. “Determinants of Polar versus Nematic Organization in Networks of Dynamic Microtubules and Mitotic Motors”. en. In: *Cell* 175.3 (Oct. 2018), 796–808.e14. ISSN: 00928674. DOI: 10.1016/j.cell.2018.09.029.
- [92] Steven M. Block, Lawrence S. B. Goldstein, and Bruce J. Schnapp. “Bead movement by single kinesin molecules studied with optical tweezers”. en. In: *Nature* 348.6299 (Nov. 1990), pp. 348–352. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/348348a0.
- [93] Ryan B Case et al. “The Directional Preference of Kinesin Motors Is Specified by an Element outside of the Motor Catalytic Domain”. en. In: *Cell* 90.5 (Sept. 1997), pp. 959–966. ISSN: 00928674. DOI: 10.1016/S0092-8674(00)80360-8.
- [94] I.M.-T.C. Crevel, A. Lockhart, and R.A. Cross. “Kinetic evidence for low chemical processivity in *ncd* and *Eg5*”. en. In: *Journal of Molecular Biology* 273.1 (Oct. 1997), pp. 160–170. ISSN: 00222836. DOI: 10.1006/jmbi.1997.1319.
- [95] Kelly A. Foster and Susan P. Gilbert. “Kinetic Studies of Dimeric *Ncd*: Evidence That *Ncd* Is Not Processive †”. en. In: *Biochemistry* 39.7 (Feb. 2000), pp. 1784–1791. ISSN: 0006-2960, 1520-4995. DOI: 10.1021/bi991500b.
- [96] Richard A. Walker, Edward D. Salmon, and Sharyn A. Endow. “The *Drosophila* claret segregation protein is a minus-end directed motor molecule”. en. In: *Nature* 347.6295 (Oct. 1990), pp. 780–782. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/347780a0.
- [97] Megan T. Valentine et al. “Individual dimers of the mitotic kinesin motor *Eg5* step processively and support substantial loads *in vitro*”. en. In: *Nature Cell Biology* 8.5 (May 2006), pp. 470–476. ISSN: 1465-7392, 1476-4679. DOI: 10.1038/ncb1394.
- [98] Christian Hentrich and Thomas Surrey. “Microtubule organization by the antagonistic mitotic motors kinesin-5 and kinesin-14”. en. In: *The Journal*

- of Cell Biology* 189.3 (May 2010), pp. 465–480. ISSN: 0021-9525, 1540-8140. DOI: 10.1083/jcb.200910125.
- [99] I. S. Aranson and L. S. Tsimring. “Theory of self-assembly of microtubules and motors”. In: *Physical Review E* 74.3 (2006), p. 031915.
- [100] M. T. Valentine and S. P. Gilbert. “To step or not to step? How biochemistry and mechanics influence processivity in Kinesin and Eg5”. In: *Current Opinion in Cell Biology* 19.1 (2007), pp. 75–81.
- [101] K. Y. Furuta et al. “Measuring collective transport by defined numbers of processive and nonprocessive kinesin motors”. In: *Proceedings of the National Academy of Sciences* 110.2 (2013), pp. 501–506.
- [102] Stephen R. Norris et al. “Microtubule minus-end aster organization is driven by processive HSET-tubulin clusters”. en. In: *Nature Communications* 9.1 (Dec. 2018), p. 2659. ISSN: 2041-1723. DOI: 10.1038/s41467-018-04991-2.
- [103] E. Jonsson et al. “Clustering of a kinesin-14 motor enables processive retrograde microtubule-based transport in plants”. In: *Nature Plants* 1.7 (2015), pp. 1–7.
- [104] L. S. Ferro et al. “Kinesin and dynein use distinct mechanisms to bypass obstacles”. In: *Elife* 8 (2019), e48629.
- [105] R. Schneider et al. “Kinesin-1 motors can circumvent permanent roadblocks by side-shifting to neighboring protofilaments”. In: *Biophysical journal* 108.9 (2015), pp. 2249–2257.
- [106] Gina A. Monzon et al. “Stable tug-of-war between kinesin-1 and cytoplasmic dynein upon different ATP and roadblock concentrations”. en. In: *Journal of Cell Science* 133.22 (Nov. 2020), jcs249938. ISSN: 1477-9137, 0021-9533. DOI: 10.1242/jcs.249938.
- [107] N. D. Derr et al. “Tug-of-War in Motor Protein Ensembles Revealed with a Programmable DNA Origami Scaffold”. en. In: *Science* 338.6107 (Nov. 2012), pp. 662–665. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1226734.
- [108] François Nédélec. “Computer simulations reveal motor properties generating stable antiparallel microtubule interactions”. en. In: *Journal of Cell Biology* 158.6 (Sept. 2002), pp. 1005–1015. ISSN: 1540-8140, 0021-9525. DOI: 10.1083/jcb.200202051.
- [109] R. Oldenbourg and Mei G. “New polarized light microscope with precision universal compensator”. In: 180.2 (1995), pp. 140–147.

- [110] Michael P. N. Juniper et al. “Spherical network contraction forms microtubule asters in confinement”. en. In: *Soft Matter* 14.6 (2018), pp. 901–909. ISSN: 1744-683X, 1744-6848. DOI: 10.1039/C7SM01718A.
- [111] H. R. Belsham and C. T. Friel. “Identification of key residues that regulate the interaction of kinesins with microtubule ends”. In: *Cytoskeleton* 76.7-8 (2019), pp. 440–446.
- [112] Malcolm S. W. Hodgskiss et al. “A productivity collapse to end Earth’s Great Oxidation”. en. In: *Proceedings of the National Academy of Sciences* 116.35 (Aug. 2019), pp. 17207–17212. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1900325116.
- [113] Ashley P. Gumsley et al. “Timing and tempo of the Great Oxidation Event”. en. In: *Proceedings of the National Academy of Sciences* 114.8 (Feb. 2017), pp. 1811–1816. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1608824114.
- [114] Alex L. Sessions et al. “The Continuing Puzzle of the Great Oxidation Event”. en. In: *Current Biology* 19.14 (July 2009), R567–R574. ISSN: 09609822. DOI: 10.1016/j.cub.2009.05.054.
- [115] James A. Estes et al. “Trophic Downgrading of Planet Earth”. en. In: *Science* 333.6040 (July 2011), pp. 301–306. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1205106.
- [116] A. M. Springer et al. “Sequential megafaunal collapse in the North Pacific Ocean: An ongoing legacy of industrial whaling?” en. In: *Proceedings of the National Academy of Sciences* 100.21 (Oct. 2003), pp. 12223–12228. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1635156100.
- [117] Sarah L. Hale and John L. Koprowski. “Ecosystem-level effects of keystone species reintroduction: a literature review: Effects of keystone species reintroduction”. en. In: *Restoration Ecology* 26.3 (May 2018), pp. 439–445. ISSN: 10612971. DOI: 10.1111/rec.12684.
- [118] Ricardo M. Holdo et al. “A Disease-Mediated Trophic Cascade in the Serengeti and its Implications for Ecosystem C”. en. In: *PLoS Biology* 7.9 (Sept. 2009). Ed. by Georgina M. Mace, e1000210. ISSN: 1545-7885. DOI: 10.1371/journal.pbio.1000210.
- [119] Gerardo Ceballos et al. “Accelerated modern human-induced species losses: Entering the sixth mass extinction”. en. In: *Science Advances* 1.5 (June 2015), e1400253. ISSN: 2375-2548. DOI: 10.1126/sciadv.1400253.

- [120] A. D. Davidson et al. “Multiple ecological pathways to extinction in mammals”. en. In: *Proceedings of the National Academy of Sciences* 106.26 (June 2009), pp. 10702–10705. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.0901956106.
- [121] Daniel Fortin et al. “WOLVES INFLUENCE ELK MOVEMENTS: BEHAVIOR SHAPES A TROPHIC CASCADE IN YELLOWSTONE NATIONAL PARK”. en. In: *Ecology* 86.5 (May 2005), pp. 1320–1330. ISSN: 0012-9658. DOI: 10.1890/04-0953.
- [122] Lusha M. Tronstad et al. “Introduced Lake Trout Produced a Four-Level Trophic Cascade in Yellowstone Lake”. en. In: *Transactions of the American Fisheries Society* 139.5 (Sept. 2010), pp. 1536–1550. ISSN: 0002-8487, 1548-8659. DOI: 10.1577/T09-151.1.
- [123] Stephen R. Palumbi. “Humans as the World’s Greatest Evolutionary Force”. en. In: *Science* 293.5536 (Sept. 2001), pp. 1786–1790. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.293.5536.1786.
- [124] Yinon M. Bar-On, Rob Phillips, and Ron Milo. “The biomass distribution on Earth”. en. In: *Proceedings of the National Academy of Sciences* 115.25 (June 2018), pp. 6506–6511. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1711842115.
- [125] Intergovernmental Panel on Climate Change (IPCC). *Special Report on the Ocean and Cryosphere in a Changing Climate*. 2019.
- [126] Hai-Tao Peng et al. “Summer albedo variations in the Arctic Sea ice region from 1982 to 2015”. en. In: *International Journal of Climatology* 40.6 (May 2020), pp. 3008–3020. ISSN: 0899-8418, 1097-0088. DOI: 10.1002/joc.6379.
- [127] F. Marcianesi, G. Aulicino, and P. Wadhams. “Arctic sea ice and snow cover albedo variability and trends during the last three decades”. en. In: *Polar Science* 28 (June 2021), p. 100617. ISSN: 18739652. DOI: 10.1016/j.polar.2020.100617.
- [128] Jérémie Mouginot et al. “Forty-six years of Greenland Ice Sheet mass balance from 1972 to 2018”. en. In: *Proceedings of the National Academy of Sciences* 116.19 (May 2019), pp. 9239–9244. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1904242116.
- [129] Joseph A. MacGregor et al. “A synthesis of the basal thermal state of the Greenland Ice Sheet”. en. In: *Journal of Geophysical Research: Earth Surface* 121.7 (July 2016), pp. 1328–1350. ISSN: 2169-9003, 2169-9011. DOI: 10.1002/2015JF003803.

- [130] R Kwok. “Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018)”. In: *Environmental Research Letters* 13.10 (Oct. 2018), p. 105005. ISSN: 1748-9326. DOI: 10.1088/1748-9326/aae3ec.
- [131] Robert A. Massom et al. “Antarctic ice shelf disintegration triggered by sea ice loss and ocean swell”. en. In: *Nature* 558.7710 (June 2018), pp. 383–389. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-018-0212-1.
- [132] Dirk Notz and Julienne Stroeve. “Observed Arctic sea-ice loss directly follows anthropogenic CO₂ emission”. en. In: *Science* 354.6313 (Nov. 2016), pp. 747–750. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.aag2345.
- [133] Julienne Stroeve and Dirk Notz. “Changing state of Arctic sea ice across all seasons”. In: *Environmental Research Letters* 13.10 (Sept. 2018), p. 103001. ISSN: 1748-9326. DOI: 10.1088/1748-9326/aade56.
- [134] P. R. Goode et al. “Earth’s Albedo 1998–2017 as Measured From Earthshine”. en. In: *Geophysical Research Letters* 48.17 (Sept. 2021). ISSN: 0094-8276, 1944-8007. DOI: 10.1029/2021GL094888.
- [135] Pierre Friedlingstein et al. “Global Carbon Budget 2019”. en. In: *Earth System Science Data* 11.4 (Dec. 2019), pp. 1783–1838. ISSN: 1866-3516. DOI: 10.5194/essd-11-1783-2019.
- [136] R. A. Houghton and Alexander A. Nassikas. “Global and regional fluxes of carbon from land use and land cover change 1850-2015: Carbon Emissions From Land Use”. en. In: *Global Biogeochemical Cycles* 31.3 (Mar. 2017), pp. 456–472. ISSN: 08866236. DOI: 10.1002/2016GB005546.
- [137] Eberhard Hansis, Steven J. Davis, and Julia Pongratz. “Relevance of methodological choices for accounting of land use change carbon fluxes”. en. In: *Global Biogeochemical Cycles* 29.8 (Aug. 2015), pp. 1230–1246. ISSN: 0886-6236, 1944-9224. DOI: 10.1002/2014GB004997.
- [138] Charles D. Keeling. “The Concentration and Isotopic Abundances of Carbon Dioxide in the Atmosphere”. en. In: *Tellus* 12.2 (May 1960), pp. 200–203. ISSN: 00402826, 21533490. DOI: 10.1111/j.2153-3490.1960.tb01300.x.
- [139] Jin-Ho Yoon et al. “Extreme Fire Season in California: A Glimpse Into the Future?” en. In: *Bulletin of the American Meteorological Society* 96.12 (Dec. 2015), S5–S9. ISSN: 0003-0007, 1520-0477. DOI: 10.1175/BAMS-D-15-00114.1.

- [140] Richard Seager et al. “Causes of the 2011–14 California Drought*”. en. In: *Journal of Climate* 28.18 (Sept. 2015), pp. 6997–7024. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-14-00860.1.
- [141] Ron Milo et al. “BioNumbers-the database of key numbers in molecular and cell biology”. In: *Nucleic Acids Res.* 38 (2010), pp. D750–D753.
- [142] Ron Milo and Rob Phillips. *Cell biology by the numbers*. Garland Science, 2016.
- [143] Hadley Wickham. “Tidy Data”. en. In: *Journal of Statistical Software* 59.10 (2014). ISSN: 1548-7660. DOI: 10.18637/jss.v059.i10.
- [144] Scripps Institution of Oceanography et al. “Understanding Ocean Acidification Impacts on Organismal to Ecological Scales”. In: *Oceanography* 25.2 (June 2015), pp. 16–27. ISSN: 10428275. DOI: 10.5670/oceanog.2015.27.
- [145] Thomas Frederikse et al. “The causes of sea-level rise since 1900”. en. In: *Nature* 584.7821 (Aug. 2020), pp. 393–397. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-020-2591-3.
- [146] Jonathan A. Patz et al. “Impact of regional climate change on human health”. en. In: *Nature* 438.7066 (Nov. 2005), pp. 310–317. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature04188.
- [147] David M. J. S. Bowman et al. “Fire in the Earth System”. en. In: *Science* 324.5926 (Apr. 2009), pp. 481–484. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1163886.
- [148] Cristina Santín et al. “Pyrogenic organic matter production from wildfires: a missing sink in the global carbon cycle”. en. In: *Global Change Biology* 21.4 (Apr. 2015), pp. 1621–1633. ISSN: 1354-1013, 1365-2486. DOI: 10.1111/gcb.12800.
- [149] Philip G. Curtis et al. “Classifying drivers of global forest loss”. en. In: *Science* 361.6407 (Sept. 2018), pp. 1108–1111. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.aau3445.
- [150] Théophile Olivier et al. “Urbanization and agricultural intensification destabilize animal communities differently than diversity loss”. en. In: *Nature Communications* 11.1 (Dec. 2020), p. 2686. ISSN: 2041-1723. DOI: 10.1038/s41467-020-16240-6.
- [151] Tim Newbold et al. “Global effects of land use on local terrestrial biodiversity”. en. In: *Nature* 520.7545 (Apr. 2015), pp. 45–50. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature14324.

- [152] Jaia Syvitski et al. “Extraordinary human energy consumption and resultant geological impacts beginning around 1950 CE initiated the proposed Anthropocene Epoch”. en. In: *Communications Earth & Environment* 1.1 (Dec. 2020), p. 32. ISSN: 2662-4435. DOI: 10.1038/s43247-020-00029-y.
- [153] FAO. *State of the World’s Forests 2020*. 2020.
- [154] Jan C. Minx et al. “A comprehensive and synthetic dataset for global, regional, and national greenhouse gas emissions by sector 1970–2018 with an extension to 2019”. en. In: *Earth System Science Data* 13.11 (Nov. 2021), pp. 5213–5252. ISSN: 1866-3516. DOI: 10.5194/essd-13-5213-2021.
- [155] IPCC. *Global Warming of 1.5°C*. 2018.
- [156] IPCC. *The Ocean and Cryosphere in a Changing Climate*. 2019.
- [157] Will Steffen et al. “Planetary boundaries: Guiding human development on a changing planet”. en. In: *Science* 347.6223 (Feb. 2015), p. 1259855. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1259855.
- [158] Colin P. Morice et al. “Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set: THE HADCRUT4 DATASET”. en. In: *Journal of Geophysical Research: Atmospheres* 117.D8 (Apr. 2012), n/a–n/a. ISSN: 01480227. DOI: 10.1029/2011JD017187.
- [159] Nathan J. L. Lenssen et al. “Improvements in the GISTEMP Uncertainty Model”. en. In: *Journal of Geophysical Research: Atmospheres* 124.12 (June 2019), pp. 6307–6326. ISSN: 2169-897X, 2169-8996. DOI: 10.1029/2018JD029522.
- [160] Huai-Min Zhang et al. “Updated Temperature Data Give a Sharper View of Climate Trends”. In: *Eos* 100 (July 2019). ISSN: 2324-9250. DOI: 10.1029/2019E0128229.
- [161] M. Zemp et al. “Global glacier mass changes and their contributions to sea-level rise from 1961 to 2016”. en. In: *Nature* 568.7752 (Apr. 2019), pp. 382–386. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-019-1071-0.
- [162] Bert Wouters, Alex S. Gardner, and Geir Moholdt. “Global Glacier Mass Loss During the GRACE Satellite Mission (2002-2016)”. In: *Frontiers in Earth Science* 7 (May 2019), p. 96. ISSN: 2296-6463. DOI: 10.3389/feart.2019.00096.

- [163] NASA/JPL. *JPL GRACE and GRACE-FO Mascon Ocean, Ice, and Hydrology Equivalent Water Height Release 06 Version 02*. Type: dataset. 2019. DOI: 10.5067/TEMSC-3MJ62.
- [164] Axel Schweiger et al. “Uncertainty in modeled Arctic sea ice volume”. en. In: *Journal of Geophysical Research* 116 (Sept. 2011), p. C00D06. ISSN: 0148-0227. DOI: 10.1029/2011JC007084.
- [165] Josefino C. Comiso, Walter N. Meier, and Robert Gersten. “Variability and trends in the Arctic Sea ice cover: Results from different techniques: TRENDS IN THE ARCTIC SEA ICE COVER”. en. In: *Journal of Geophysical Research: Oceans* 122.8 (Aug. 2017), pp. 6883–6900. ISSN: 21699275. DOI: 10.1002/2017JC012768.
- [166] F.; Knowles K.; Meier W.; Savoie M.; Windnagel A. Fetterer. *Sea Ice Index, Version 3*. en. Type: dataset. 2017. DOI: 10.7265/N5K072F8.
- [167] *Mineral Commodity Summaries*. en. Tech. rep. 2020.
- [168] Sabbie A Miller, Arpad Horvath, and Paulo J M Monteiro. “Readily implementable techniques can cut annual CO₂ emissions from the production of concrete by over 20%”. In: *Environmental Research Letters* 11.7 (July 2016), p. 074029. ISSN: 1748-9326. DOI: 10.1088/1748-9326/11/7/074029.
- [169] Paulo J. M. Monteiro, Sabbie A. Miller, and Arpad Horvath. “Towards sustainable concrete”. en. In: *Nature Materials* 16.7 (July 2017), pp. 698–699. ISSN: 1476-1122, 1476-4660. DOI: 10.1038/nmat4930.
- [170] Fridolin Krausmann et al. “Global socioeconomic material stocks rise 23-fold over the 20th century and require half of annual resource use”. en. In: *Proceedings of the National Academy of Sciences* 114.8 (Feb. 2017), pp. 1880–1885. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1613773114.
- [171] World Steel Association. *World Steel in Figures*. 2020.
- [172] Roland Geyer, Jenna R. Jambeck, and Kara Lavender Law. “Production, use, and fate of all plastics ever made”. en. In: *Science Advances* 3.7 (July 2017), e1700782. ISSN: 2375-2548. DOI: 10.1126/sciadv.1700782.
- [173] Food and Agriculture Organization. *Food and Agriculture Organization of the United Nations Statistical Database - Live Animals*. 2020.
- [174] International Fertilizer Association. *Ammonia Production and Trade Tables by Region*. 2020.

- [175] Collin Smith, Alfred K. Hill, and Laura Torrente-Murciano. “Current and future role of Haber–Bosch ammonia in a carbon-free energy landscape”. en. In: *Energy & Environmental Science* 13.2 (2020), pp. 331–344. ISSN: 1754-5692, 1754-5706. DOI: 10.1039/C9EE02873K.
- [176] Copernicus Marine Service. *Global Mean Sea Water pH*. 2020.
- [177] European Environment Agency. *CLIM 043*. 2020.
- [178] Food and Agriculture Organization (FAO) of the United Nations Statistical Database. *Land Use*. 2020.
- [179] European Commission. Joint Research Centre. *GHSL data package 2019: public release GHS P2019*. eng. LU: Publications Office, 2019.
- [180] Xiaoping Liu et al. “High-resolution multi-temporal mapping of global urban land using Landsat images based on the Google Earth Engine Platform”. en. In: *Remote Sensing of Environment* 209 (May 2018), pp. 227–239. ISSN: 00344257. DOI: 10.1016/j.rse.2018.02.055.
- [181] G. Grill et al. “Mapping the world’s free-flowing rivers”. en. In: *Nature* 569.7755 (May 2019), pp. 215–221. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-019-1111-9.
- [182] Food and Agriculture Organization. *Annual population*. 2020.
- [183] Christopher Jones and Daniel M. Kammen. “Spatial Distribution of U.S. Household Carbon Footprints Reveals Suburbanization Undermines Greenhouse Gas Benefits of Urban Population Density”. en. In: *Environmental Science & Technology* 48.2 (Jan. 2014), pp. 895–902. ISSN: 0013-936X, 1520-5851. DOI: 10.1021/es4034364.
- [184] Food and Agriculture Organization. *AQUASTAT*. 2020.
- [185] British Petroleum. *Statistical Review of World Energy*. 2020.
- [186] US Energy Information Agency. *International*. 2020.
- [187] M. C. Hansen et al. “High-Resolution Global Maps of 21st-Century Forest Cover Change”. en. In: *Science* 342.6160 (Nov. 2013), pp. 850–853. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1244693.
- [188] S. Levitus et al. “World ocean heat content and thermosteric sea level change (0-2000 m), 1955-2010: WORLD OCEAN HEAT CONTENT”. en. In: *Geophysical Research Letters* 39.10 (May 2012), n/a–n/a. ISSN: 00948276. DOI: 10.1029/2012GL051106.

- [189] Dean Roemmich et al. “Unabated planetary warming and its ocean structure since 2006”. en. In: *Nature Climate Change* 5.3 (Mar. 2015), pp. 240–245. ISSN: 1758-678X, 1758-6798. DOI: 10.1038/nclimate2513.
- [190] Gary J. Hancock et al. “The release and persistence of radioactive anthropogenic nuclides”. en. In: *Geological Society, London, Special Publications* 395.1 (2014), pp. 265–281. ISSN: 0305-8719, 2041-4927. DOI: 10.1144/SP395.15.
- [191] United Nations Scientific Committee on the Effects of Atomic Radiation. *UNSCEAR 2000 Report Vol. 1: Sources and effects of ionizing radiation*. 2020.
- [192] David M. Taylor. “Environmental plutonium — creation of the universe to twenty-first century mankind”. en. In: *Radioactivity in the Environment*. Vol. 1. Elsevier, 2001, pp. 1–14. ISBN: 978-0-08-043425-4. DOI: 10.1016/S1569-4860(01)80003-6.
- [193] IUCN. *Red List of Threatened Species*. 2020.
- [194] Anthony H Cooper et al. “Humans are the most significant global geomorphological driving force of the 21st century”. en. In: *The Anthropocene Review* 5.3 (Dec. 2018), pp. 222–229. ISSN: 2053-0196, 2053-020X. DOI: 10.1177/2053019618800234.
- [195] Zhengang Wang and Kristof Van Oost. “Modeling global anthropogenic erosion in the Holocene”. en. In: *The Holocene* 29.3 (Mar. 2019), pp. 367–379. ISSN: 0959-6836, 1477-0911. DOI: 10.1177/0959683618816499.
- [196] Pasquale Borrelli et al. “Land use and climate change impacts on global soil erosion by water (2015-2070)”. en. In: *Proceedings of the National Academy of Sciences* 117.36 (Sept. 2020), pp. 21994–22001. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.2001403117.
- [197] E. Suzuki. *World’s population will continue to grow and will reach nearly 10 billion by 2050*. 2019.
- [198] Population Division UNWUP - United Nations Department of Economic and Social Affairs. *World Urbanization Prospects: The 2018 Revision, Online Edition*. 2018.
- [199] Center For International Earth Science Information Network-CIESIN-Columbia University. *Low Elevation Coastal Zone (LECZ) Urban-Rural Population and Land Area Estimates, Version 2*. Type: dataset. 2013. DOI: 10.7927/H4MW2F2J.

- [200] USDA-ERS. *What is Rural?* 2019.
- [201] Jason W. Haegele et al. “Row Arrangement, Phosphorus Fertility, and Hybrid Contributions to Managing Increased Plant Density of Maize”. en. In: *Agronomy Journal* 106.5 (Sept. 2014), pp. 1838–1846. ISSN: 0002-1962, 1435-0645. DOI: 10.2134/agronj2013.0382.
- [202] G. Eshel et al. “Land, irrigation water, greenhouse gas, and reactive nitrogen burdens of meat, eggs, and dairy production in the United States”. en. In: *Proceedings of the National Academy of Sciences* 111.33 (Aug. 2014), pp. 11996–12001. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1402183111.
- [203] Alon Shepon et al. “The opportunity cost of animal based diets exceeds all food losses”. en. In: *Proceedings of the National Academy of Sciences* 115.15 (Apr. 2018), pp. 3804–3809. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1713820115.
- [204] Food and Agriculture Organization. *Crops and livestock products*. 2020.
- [205] Catherine I Birney et al. “An assessment of individual foodprints attributed to diets and food waste in the United States”. In: *Environmental Research Letters* 12.10 (Oct. 2017), p. 105008. ISSN: 1748-9326. DOI: 10.1088/1748-9326/aa8494.
- [206] M.J. Zuidhof et al. “Growth, efficiency, and yield of commercial broilers from 1957, 1978, and 2005”. en. In: *Poultry Science* 93.12 (Dec. 2014), pp. 2970–2982. ISSN: 00325791. DOI: 10.3382/ps.2014-04291.
- [207] N.M. Bello, J.S. Stevenson, and R.J. Tempelman. “Invited review: Milk production and reproductive performance: Modern interdisciplinary insights into an enduring axiom”. en. In: *Journal of Dairy Science* 95.10 (Oct. 2012), pp. 5461–5475. ISSN: 00220302. DOI: 10.3168/jds.2012-5564.
- [208] and Forestry OKFFA (Oklahoma Department of Agriculture Food. *How much meat (beef)?* 2020.
- [209] Ryota Hosomi, Munehiro Yoshida, and Kenji Fukunaga. “Seafood Consumption and Components for Health”. In: *Global Journal of Health Science* 4.3 (Apr. 2012), p72. ISSN: 1916-9744, 1916-9736. DOI: 10.5539/gjhs.v4n3p72.
- [210] Food and Agriculture Organization. *The state of the world’s land and water resources for food and agriculture*. 2011.
- [211] MaP testing. *Toilet Types*.

- [212] Consumer Reports. *Yes, your washing machine is using enough water*. 2017.
- [213] Tongjun Zhang. “Methods of Improving the Efficiency of Thermal Power Plants”. In: *Journal of Physics: Conference Series* 1449.1 (Jan. 2020), p. 012001. DOI: 10.1088/1742-6596/1449/1/012001.
- [214] C E Raptis, M T H van Vliet, and S Pfister. “Global thermal pollution of rivers from thermoelectric power plants”. In: *Environmental Research Letters* 11.10 (Oct. 2016), p. 104011. ISSN: 1748-9326. DOI: 10.1088/1748-9326/11/10/104011.
- [215] Jordan Macknick et al. *Review of Operational Water Consumption and Withdrawal Factors for Electricity Generating Technologies*. en. Tech. rep. NREL/TP-6A20-50900, 1009674. Mar. 2011, NREL/TP-6A20-50900, 1009674. DOI: 10.2172/1009674.
- [216] Energy and Water in a Warming World Initiative. *Freshwater use by U.S. power plants*. 2011.
- [217] Food. *Geo-referenced Database on Dams*. 2020.
- [218] USGS. *Three Gorges Dam: the world’s largest hydroelectric plant*.
- [219] *Hydrology of the Yangtze river*.
- [220] K.J. Earp and M.T. Moreo. *Evaporation from Lake Mead and Lake Mohave, Nevada and Arizona, 2010-2019*.
- [221] Laura Scherer and Stephan Pfister. “Global water footprint assessment of hydropower”. en. In: *Renewable Energy* 99 (Dec. 2016), pp. 711–720. ISSN: 09601481. DOI: 10.1016/j.renene.2016.07.021.
- [222] Taikan Oki and Shinjiro Kanae. “Global Hydrological Cycles and World Water Resources”. en. In: *Science* 313.5790 (Aug. 2006), pp. 1068–1072. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1128845.
- [223] Marcel M. M. Kuypers, Hannah K. Marchant, and Boran Kartal. “The microbial nitrogen-cycling network”. en. In: *Nature Reviews Microbiology* 16.5 (May 2018), pp. 263–276. ISSN: 1740-1526, 1740-1534. DOI: 10.1038/nrmicro.2018.9.
- [224] Ben Johnson and Colin Goldblatt. “The nitrogen budget of Earth”. en. In: *Earth-Science Reviews* 148 (Sept. 2015), pp. 150–173. ISSN: 00128252. DOI: 10.1016/j.earscirev.2015.05.006.

- [225] B. Z. Houlton, S. L. Morford, and R. A. Dahlgren. “Convergent evidence for widespread rock nitrogen sources in Earth’s surface environment”. en. In: *Science* 360.6384 (Apr. 2018), pp. 58–62. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.aan4399.
- [226] Brian M. Hoffman et al. “Mechanism of Nitrogen Fixation by Nitrogenase: The Next Stage”. en. In: *Chemical Reviews* 114.8 (Apr. 2014), pp. 4041–4062. ISSN: 0009-2665, 1520-6890. DOI: 10.1021/cr400641x.
- [227] Yangjian Zhang et al. “Global pattern of NPP to GPP ratio derived from MODIS data: effects of ecosystem type, geographical location and climate”. en. In: *Global Ecology and Biogeography* 18.3 (May 2009), pp. 280–290. ISSN: 1466822X, 14668238. DOI: 10.1111/j.1466-8238.2008.00442.x.
- [228] Xinning Zhang, Bess B. Ward, and Daniel M. Sigman. “Global Nitrogen Cycle: Critical Enzymes, Organisms, and Processes for Nitrogen Budgets and Dynamics”. en. In: *Chemical Reviews* 120.12 (June 2020), pp. 5308–5351. ISSN: 0009-2665, 1520-6890. DOI: 10.1021/acs.chemrev.9b00613.
- [229] U. Schumann and H. Huntrieser. “The global lightning-induced nitrogen oxides source”. en. In: *Atmospheric Chemistry and Physics* 7.14 (July 2007), pp. 3823–3907. ISSN: 1680-7324. DOI: 10.5194/acp-7-3823-2007.
- [230] Christopher R. Maggio, Thomas C. Marshall, and Maribeth Stolzenburg. “Estimations of charge transferred and energy released by lightning flashes”. en. In: *Journal of Geophysical Research* 114.D14 (July 2009), p. D14203. ISSN: 0148-0227. DOI: 10.1029/2008JD011506.
- [231] Hugh J. Christian. “Global frequency and distribution of lightning as observed from space by the Optical Transient Detector”. en. In: *Journal of Geophysical Research* 108.D1 (2003), p. 4005. ISSN: 0148-0227. DOI: 10.1029/2002JD002347.
- [232] Nicolas Gruber and James N. Galloway. “An Earth-system perspective of the global nitrogen cycle”. en. In: *Nature* 451.7176 (Jan. 2008), pp. 293–296. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature06592.
- [233] James N. Galloway et al. “The Nitrogen Cascade”. en. In: *BioScience* 53.4 (2003), p. 341. ISSN: 0006-3568. DOI: 10.1641/0006-3568(2003)053[0341:TNC]2.0.CO;2.
- [234] M. Meinshausen et al. “Historical greenhouse gas concentrations for climate modelling (CMIP6)”. In: *Geoscientific Model Development* 10.5 (2017), pp. 2057–2116. DOI: 10.5194/gmd-10-2057-2017.

- [235] D. M. Etheridge et al. “Atmospheric methane between 1000 A.D. and present: Evidence of anthropogenic emissions and climatic variability”. en. In: *Journal of Geophysical Research: Atmospheres* 103.D13 (July 1998), pp. 15979–15993. ISSN: 01480227. DOI: 10.1029/98JD00923.
- [236] Andy Thorpe. “Enteric fermentation and ruminant eructation: the role (and control?) of methane in the climate change debate”. en. In: *Climatic Change* 93.3-4 (Apr. 2009), pp. 407–431. ISSN: 0165-0009, 1573-1480. DOI: 10.1007/s10584-008-9506-x.
- [237] Food and Agriculture Organization. *Emissions Totals - FAOSTAT*. 2020.
- [238] C. Tarnocai et al. “Soil organic carbon pools in the northern circumpolar permafrost region: SOIL ORGANIC CARBON POOLS”. en. In: *Global Biogeochemical Cycles* 23.2 (June 2009), n/a–n/a. ISSN: 08866236. DOI: 10.1029/2008GB003327.
- [239] J.H. Seinfeld and S.N. Pandis. *Atmospheric Chemistry and Physics*. Wiley, 2016.
- [240] US Energy Information Agency. *International Energy Outlook*. 2021.
- [241] US Energy Information Agency. “Transportation”. In: *International Energy Outlook*. 2016.
- [242] US Energy Information Agency. *Manufacturing Energy Consumption Survey*. 2018.
- [243] US Energy Information Agency. “Energy consumption by sector”. In: *Monthly Energy Review*. 2021.
- [244] B. M. Suleiman et al. “Thermal conductivity and diffusivity of wood”. In: *Wood Science and Technology* 33.6 (Dec. 1999), pp. 465–473. ISSN: 0043-7719, 1432-5225. DOI: 10.1007/s002260050130.
- [245] US Energy Information Agency. *Residential Energy Consumption Survey*. 2015.
- [246] D.R. Williams. *NASA Sun Fact Sheet*. 2021.
- [247] University of Michigan Center for Sustainable Systems. *Photovoltaic energy factsheet*. 2021.
- [248] A. De Vos and H. Pauwels. “On the thermodynamic limit of photovoltaic energy conversion”. en. In: *Applied physics* 25.2 (June 1981), pp. 119–125. ISSN: 0340-3793, 1432-0630. DOI: 10.1007/BF00901283.

- [249] William Shockley and Hans J. Queisser. “Detailed Balance Limit of Efficiency of p - n Junction Solar Cells”. en. In: *Journal of Applied Physics* 32.3 (Mar. 1961), pp. 510–519. ISSN: 0021-8979, 1089-7550. DOI: 10.1063/1.1736034.
- [250] Clare E. Singer et al. “Top-of-atmosphere albedo bias from neglecting three-dimensional cloud radiative effects”. In: *Journal of the Atmospheric Sciences* (Sept. 2021). ISSN: 0022-4928, 1520-0469. DOI: 10.1175/JAS-D-21-0032.1.
- [251] Karl-Heinz Erb et al. “Unexpectedly large impact of forest management and grazing on global vegetation biomass”. en. In: *Nature* 553.7686 (Jan. 2018), pp. 73–76. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature25138.
- [252] Hannah Ritchie and Max Roser. “Energy”. In: *Our World in Data* (2020). <https://ourworldindata.org/energy>.
- [253] V. Smil. *Energy transitions: global and national perspectives*. 2nd. Praeger, 2016.
- [254] Charles D. Keeling et al. “Atmospheric CO₂ and ¹³CO₂ Exchange with the Terrestrial Biosphere and Oceans from 1978 to 2000: Observations and Carbon Cycle Implications”. In: *A History of Atmospheric CO₂ and Its Effects on Plants, Animals, and Ecosystems*. Ed. by I.T. Baldwin et al. New York, NY: Springer New York, 2005, pp. 83–113. ISBN: 978-0-387-27048-7. DOI: 10.1007/0-387-27048-5_5.
- [255] Daniel C. Harris. “Charles David Keeling and the Story of Atmospheric CO₂ Measurements”. en. In: *Analytical Chemistry* 82.19 (Oct. 2010), pp. 7865–7870. ISSN: 0003-2700, 1520-6882. DOI: 10.1021/ac1001492.
- [256] WM Haynes. *CRC Handbook of Chemistry and Physics*. CRC Press, 2014.
- [257] Mounir Chaouch et al. “Use of wood elemental composition to predict heat treatment intensity and decay resistance of different softwood and hardwood species”. en. In: *Polymer Degradation and Stability* 95.12 (Dec. 2010), pp. 2255–2259. ISSN: 01413910. DOI: 10.1016/j.polyimdeggradstab.2010.09.010.
- [258] Lan Yi et al. “Prediction of elemental composition of coal using proximate analysis”. en. In: *Fuel* 193 (Apr. 2017), pp. 315–321. ISSN: 00162361. DOI: 10.1016/j.fuel.2016.12.044.
- [259] J. Burri et al. “Gasoline composition determined by ¹H NMR spectroscopy”. en. In: *Fuel* 83.2 (Jan. 2004), pp. 187–193. ISSN: 00162361. DOI: 10.1016/S0016-2361(03)00261-8.
- [260] US Energy Information Agency. *Petroleum Supply Annual*. 2020.

- [261] Lei Deng et al. “Global patterns of the effects of land-use changes on soil carbon stocks”. en. In: *Global Ecology and Conservation* 5 (Jan. 2016), pp. 127–138. ISSN: 23519894. DOI: 10.1016/j.gecco.2015.12.004.
- [262] Xiao-Peng Song et al. “Global land change from 1982 to 2016”. en. In: *Nature* 560.7720 (Aug. 2018), pp. 639–643. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-018-0411-9.
- [263] David Schimel, Britton B. Stephens, and Joshua B. Fisher. “Effect of increasing CO₂ on the terrestrial carbon cycle”. en. In: *Proceedings of the National Academy of Sciences* 112.2 (Jan. 2015), pp. 436–441. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1407302112.
- [264] Elizabeth A. Ainsworth and Stephen P. Long. “What have we learned from 15 years of free-air CO₂ enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO₂”. en. In: *New Phytologist* 165.2 (Feb. 2005), pp. 351–372. ISSN: 0028-646X, 1469-8137. DOI: 10.1111/j.1469-8137.2004.01224.x.
- [265] Shilong Piao et al. “Characteristics, drivers and feedbacks of global greening”. en. In: *Nature Reviews Earth & Environment* 1.1 (Jan. 2020), pp. 14–27. ISSN: 2662-138X. DOI: 10.1038/s43017-019-0001-x.
- [266] Christopher L. Sabine et al. “The Oceanic Sink for Anthropogenic CO₂”. en. In: *Science* 305.5682 (July 2004), pp. 367–371. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1097403.
- [267] Charles D. Keeling. “Rewards and Penalties of Monitoring the Earth”. en. In: *Annual Review of Energy and the Environment* 23.1 (Nov. 1998), pp. 25–82. ISSN: 1056-3466. DOI: 10.1146/annurev.energy.23.1.25.
- [268] “Carbon and Other Biogeochemical Cycles”. In: *Climate Change 2013 - The Physical Science Basis*. Ed. by Intergovernmental Panel on Climate Change. Cambridge: Cambridge University Press, 2014, pp. 465–570. ISBN: 978-1-107-41532-4. DOI: 10.1017/CB09781107415324.015.