
Contents

Preface	xvii
CHAPTER 1 ■ Introduction to modelling	1
1.1 WHAT IS MODELLING?	1
1.1.1 What are models?	2
1.2 WHY BUILD MODELS?	3
1.2.1 Why model biological systems?	5
1.2.2 Why systems biology?	6
1.3 CHALLENGES IN MODELLING BIOLOGICAL SYSTEMS	7
1.4 THE PRACTICE OF MODELLING	9
1.4.1 Scope of the model	10
1.4.2 Making assumptions	11
1.4.3 Modelling paradigms	11
1.4.4 Building the model	13
1.4.5 Model analysis, debugging and (in)validation	17
1.4.6 Simulating the model	18
1.5 EXAMPLES OF MODELS	19
1.5.1 Lotka–Volterra predator–prey model	19
1.5.2 SIR model: A classic example	20
1.6 TROUBLESHOOTING	22
1.6.1 Clarity of scope and objectives	22
1.6.2 The breakdown of assumptions	22
1.6.3 Is my model fit for purpose?	23
1.6.4 Handling uncertainties	23
EXERCISES	23
REFERENCES	24
FURTHER READING	26

PART I Static Modelling

CHAPTER 2 ■ Introduction to graph theory	29
2.1 BASICS	29
2.1.1 History of graph theory	29
2.1.2 Examples of graphs	31
2.2 WHY GRAPHS?	32
2.3 TYPES OF GRAPHS	33
2.3.1 Simple vs. non-simple graphs	34
2.3.2 Directed vs. undirected graphs	34
2.3.3 Weighted vs. unweighted graphs	35
2.3.4 Other graph types	35
2.3.5 Hypergraphs	38
2.4 COMPUTATIONAL REPRESENTATIONS OF GRAPHS	39
2.4.1 Data structures	39
2.4.2 Adjacency matrix	39
2.4.3 The Laplacian matrix	41
2.5 GRAPH REPRESENTATIONS OF BIOLOGICAL NETWORKS	42
2.5.1 Networks of protein interactions and functional associations	42
2.5.2 Signalling networks	43
2.5.3 Protein structure networks	45
2.5.4 Gene regulatory networks	45
2.5.5 Metabolic networks	46
2.6 COMMON CHALLENGES & TROUBLESHOOTING	49
2.6.1 Choosing a representation	49
2.6.2 Loading and creating graphs	51
2.7 SOFTWARE TOOLS	51
EXERCISES	52
REFERENCES	54
FURTHER READING	56
CHAPTER 3 ■ Structure of networks	57
3.1 NETWORK PARAMETERS	57

3.1.1	Fundamental parameters	58
3.1.2	Measures of centrality	62
3.1.3	Mixing patterns: Assortativity	65
3.2	CANONICAL NETWORK MODELS	66
3.2.1	Erdős–Rényi (ER) network model	67
3.2.2	Small-world networks	69
3.2.3	Scale-free networks	71
3.2.4	Other models of network generation	74
3.3	COMMUNITY DETECTION	75
3.3.1	Modularity maximisation	76
3.3.2	Similarity-based clustering	77
3.3.3	Girvan–Newman algorithm	78
3.3.4	Other methods	78
3.3.5	Community detection in biological networks	79
3.4	NETWORK MOTIFS	79
3.4.1	Randomising networks	80
3.5	PERTURBATIONS TO NETWORKS	81
3.5.1	Quantifying effects of perturbation	82
3.5.2	Network structure and attack strategies	82
3.6	TROUBLESHOOTING	83
3.6.1	Is your network <i>really</i> scale-free?	83
3.7	SOFTWARE TOOLS	84
	EXERCISES	85
	REFERENCES	86
	FURTHER READING	90
CHAPTER 4 ■ Applications of network biology		91
4.1	THE CENTRALITY–LETHALITY HYPOTHESIS	92
4.1.1	Predicting essential genes using network measures	92
4.2	NETWORKS AND MODULES IN DISEASE	93
4.2.1	Disease networks	93
4.2.2	Identification of disease modules	95
4.2.3	Edgetic perturbation models	97
4.3	DIFFERENTIAL NETWORK ANALYSIS	97
4.4	DISEASE SPREADING ON NETWORKS	99

4.4.1	Percolation-based models	99
4.4.2	Agent-based simulations	100
4.5	MOLECULAR GRAPHS AND THEIR APPLICATIONS	101
4.5.1	Retrosynthesis	102
4.6	PROTEIN STRUCTURE AND CONFORMATIONAL NETWORKS	104
4.6.1	Protein folding pathways	104
4.7	LINK PREDICTION	107
	EXERCISES	107
	REFERENCES	108
	FURTHER READING	112
PART II Dynamic Modelling		
CHAPTER 5 ■ Introduction to dynamic modelling		115
<hr/>		
5.1	CONSTRUCTING DYNAMIC MODELS	116
5.1.1	Modelling a generic biochemical system	116
5.2	MASS-ACTION KINETIC MODELS	117
5.3	MODELLING ENZYME KINETICS	118
5.3.1	The Michaelis–Menten model	118
5.3.2	Co-operativity: Hill kinetics	122
5.3.3	An illustrative example: a three-node oscillator	123
5.4	GENERALISED RATE EQUATIONS	124
5.4.1	Biochemical systems theory	125
5.5	SOLVING ODEs	125
5.6	TROUBLESHOOTING	127
5.6.1	Handling stiff equations	127
5.6.2	Handling uncertainty	127
5.7	SOFTWARE TOOLS	128
	EXERCISES	128
	REFERENCES	129
	FURTHER READING	130
CHAPTER 6 ■ Parameter estimation		131
<hr/>		
6.1	DATA-DRIVEN MECHANISTIC MODELLING: AN OVERVIEW	131
6.1.1	Pre-processing the data	133

6.1.2	Model identification	134
6.2	SETTING UP AN OPTIMISATION PROBLEM	134
6.2.1	Linear regression	135
6.2.2	Least squares	135
6.2.3	Maximum likelihood estimation	138
6.3	ALGORITHMS FOR OPTIMISATION	139
6.3.1	Desiderata	139
6.3.2	Gradient-based methods	139
6.3.3	Direct search methods	140
6.3.4	Evolutionary algorithms	143
6.4	POST-REGRESSION DIAGNOSTICS	150
6.4.1	Model selection	150
6.4.2	Sensitivity and robustness of biological models	151
6.5	TROUBLESHOOTING	153
6.5.1	Regularisation	153
6.5.2	Sloppiness	153
6.5.3	Choosing a search algorithm	154
6.5.4	Model reduction	154
6.5.5	The curse of dimensionality	154
6.6	SOFTWARE TOOLS	155
	EXERCISES	155
	REFERENCES	157
	FURTHER READING	160
CHAPTER 7 ■ Discrete dynamic models: Boolean networks		161
<hr/>		
7.1	INTRODUCTION	161
7.2	BOOLEAN NETWORKS: TRANSFER FUNCTIONS	162
7.2.1	Characterising Boolean network dynamics	164
7.2.2	Synchronous vs. asynchronous updates	165
7.3	OTHER PARADIGMS	166
7.3.1	Probabilistic Boolean networks	166
7.3.2	Logical interaction hypergraphs	166
7.3.3	Generalised logical networks	166
7.3.4	Petri nets	166
7.4	APPLICATIONS	167

7.5	TROUBLESHOOTING	167
7.6	SOFTWARE TOOLS	167
	EXERCISES	168
	REFERENCES	169
	FURTHER READING	170

PART III Constraint-based Modelling

CHAPTER	8 ■ Introduction to constraint-based modelling	173
8.1	WHAT ARE CONSTRAINTS?	174
8.1.1	Types of constraints	175
8.1.2	Mathematical representation of constraints	176
8.1.3	Why are constraints useful?	177
8.2	THE STOICHIOMETRIC MATRIX	177
8.3	STEADY-STATE MASS BALANCE: FLUX BALANCE ANALYSIS	178
8.4	THE OBJECTIVE FUNCTION	181
8.4.1	The biomass objective function	182
8.5	OPTIMISATION TO COMPUTE FLUX DISTRIBUTION	183
8.6	AN ILLUSTRATION	185
8.7	FLUX VARIABILITY ANALYSIS (FVA)	187
8.8	UNDERSTANDING FBA	187
8.8.1	Blocked reactions and dead-end metabolites	189
8.8.2	Caps in metabolic networks	190
8.8.3	Multiple solutions	191
8.8.4	Loops	191
8.8.5	Parsimonious FBA (pFBA)	192
8.8.6	ATP maintenance fluxes	193
8.9	TROUBLESHOOTING	194
8.9.1	Zero growth rate	194
8.9.2	Objective values vs. flux values	194
8.10	SOFTWARE TOOLS	195
	EXERCISES	195
	REFERENCES	197
	FURTHER READING	200

CHAPTER 9 ■ Extending constraint-based approaches	201
9.1 MINIMISATION OF METABOLIC ADJUSTMENT (MoMA)	202
9.1.1 Fitting experimentally measured fluxes	203
9.2 REGULATORY ON-OFF MINIMISATION (ROOM)	204
9.2.1 ROOM vs. MoMA	205
9.3 BI-LEVEL OPTIMISATION	205
9.3.1 OptKnock	205
9.4 INTEGRATING REGULATORY INFORMATION	206
9.4.1 Embedding regulatory logic: Regulatory FBA (rFBA)	206
9.4.2 Informing metabolic models with omic data	207
9.4.3 Tissue-specific models	209
9.5 COMPARTMENTALISED MODELS	209
9.6 DYNAMIC FLUX BALANCE ANALYSIS (dFBA)	210
9.7 ^{13}C -MFA	212
9.8 ELEMENTARY FLUX MODES AND EXTREME PATHWAYS	213
9.8.1 Computing EFMs and EPs	216
9.8.2 Applications	216
EXERCISES	216
REFERENCES	218
FURTHER READING	222
CHAPTER 10 ■ Perturbations to metabolic networks	223
10.1 KNOCK-OUTS	224
10.1.1 Gene deletions vs. reaction deletions	225
10.2 SYNTHETIC LETHALS	225
10.2.1 Exhaustive enumeration	226
10.2.2 Bi-level optimisation	227
10.2.3 Fast-SL: Massively pruning the search space	229
10.3 OVER-EXPRESSION	232
10.3.1 Flux Scanning based on Enforced Objective Flux (FSEOF)	232
10.4 OTHER PERTURBATIONS	233
10.5 EVALUATING AND RANKING PERTURBATIONS	233
10.6 APPLICATIONS OF CONSTRAINT-BASED MODELS	234
10.6.1 Metabolic engineering	235

10.6.2	Drug target identification	236
10.7	LIMITATIONS OF CONSTRAINT-BASED APPROACHES	237
10.7.1	Incorrect predictions	237
10.8	TROUBLESHOOTING	239
10.8.1	Interpreting gene deletion simulations	239
10.9	SOFTWARE TOOLS	239
	EXERCISES	240
	REFERENCES	241
	FURTHER READING	246

PART IV Advanced Topics

CHAPTER 11 ■ Modelling cellular interactions 249

11.1	MICROBIAL COMMUNITIES	249
11.1.1	Network-based approaches	250
11.1.2	Population-based and agent-based approaches	255
11.1.3	Constraint-based approaches	256
11.2	HOST-PATHOGEN INTERACTIONS (HPIs)	260
11.2.1	Network models	260
11.2.2	Dynamic models	262
11.2.3	Constraint-based models	263
11.3	SUMMARY	265
11.4	SOFTWARE TOOLS	265
	EXERCISES	266
	REFERENCES	267
	FURTHER READING	274

CHAPTER 12 ■ Designing biological circuits 275

12.1	WHAT IS SYNTHETIC BIOLOGY?	275
12.1.1	From LEGO bricks to biobricks	276
12.2	CLASSIC CIRCUIT DESIGN EXPERIMENTS	277
12.2.1	Designing an oscillator: The repressilator	278
12.2.2	Toggle switch	279
12.3	DESIGNING MODULES	280
12.3.1	Exploring the design space	280

12.3.2	Systems-theoretic approaches	282
12.3.3	Automating circuit design	282
12.4	DESIGN PRINCIPLES OF BIOLOGICAL NETWORKS	282
12.4.1	Redundancy	283
12.4.2	Modularity	283
12.4.3	Exaptation	284
12.4.4	Robustness	284
12.5	COMPUTING WITH CELLS	284
12.5.1	Adleman's classic experiment	285
12.5.2	Examples of circuits that can compute	285
12.5.3	DNA data storage	286
12.6	CHALLENGES	286
12.7	SOFTWARE TOOLS	287
	EXERCISES	287
	REFERENCES	288
	FURTHER READING	292
CHAPTER 13 ■ Robustness and evolvability of biological systems		293
13.1	ROBUSTNESS IN BIOLOGICAL SYSTEMS	294
13.1.1	Key mechanisms	294
13.1.2	Hierarchies and protocols	295
13.1.3	Organising principles	296
13.2	GENOTYPE SPACES AND GENOTYPE NETWORKS	298
13.2.1	Genotype spaces	298
13.2.2	Genotype–phenotype mapping	298
13.3	QUANTIFYING ROBUSTNESS AND EVOLVABILITY	301
13.4	SOFTWARE TOOLS	305
	EXERCISES	305
	REFERENCES	306
	FURTHER READING	307
CHAPTER 14 ■ Epilogue: The road ahead		309

Online-only appendices	325
APPENDIX A ■ Introduction to key biological concepts	
APPENDIX B ■ Reconstruction of biological networks	
APPENDIX C ■ Databases for systems biology	
APPENDIX D ■ Software tools compendium	
APPENDIX E ■ MATLAB for systems biology	
Index	327