2241153

MICHIGAN STATE UNIVERSITY LIBRARIES

3 1293 00571 2207

V. 1

LIBRARY
Michigan State
University

This is to certify that the

dissertation entitled

POPULATION DYNAMICS OF BIOLOGICAL SYSTEMS
USING THE HEAT UNIT HYPOTHESIS

presented by

Francis Lynn Severance

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Systems Science

Robert O. Barr, Jr.
Major professor

Date February 4, 1988

MSU is an Affirmative Action/Equal Opportunity Institution

O-12771



RETURNING MATERIALS:
Place in book drop to remove this checkout from your record. FINES will be charged if book is returned after the date stamped below.

POPULATION DYNAMICS OF BIOLOGICAL SYSTEMS USING THE HEAT UNIT HYPOTHESIS

Volume I

By

Francis Lynn Severance

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Electrical Engineering

1988

ABSTRACT

POPULATION DYNAMICS OF BIOLOGICAL SYSTEMS USING THE HEAT UNIT HYPOTHESIS

By

Francis Lynn Severance

A mathematical treatment of heat-unit based physiological time systems is developed for applications in population dynamics models. By defining an appropriate biological activity function it is possible to generalize on the chronological time based models traditionally used. Further, it is shown that a general class of chronological time bases is equivalent to the physiological time base under a classical description. Systems using this methodology are analyzed and solved. Also, numerical techniques are developed, computed and illustrated to show system characteristics.

Employing the discrete data collection methods used by the National Weather Bureau, data reduction formulae are derived for both historical validation and stochastic weather generation models. It is shown that these reduction formulae accurately represent the physiological time base generally, but most importantly during the critical spring seasons. Numerical solutions validate the accuracy of this approach. A procedure is developed by which the endogenous parameters, such as temperature thresholds and growth rates,

can

det

exa

By

red

is

and

an

pre

ana

can be accurately determined. Formulae are derived in detail, prescriptive algorithms listed and illustrative examples are given throughout.

By invoking a stochastic weather generator, the data reduction formulae produce a realistic population model. It is shown that these models are well behaved statistically and that long term means and variation are controllable. As an illustration, a passive control example using a predator-prey model is given to demonstrate the above results and analyze system characteristics.

The

follo

the c

a

a

a

Se

he

ACKNOWLEDGMENTS

The author wishes to express special appreciation to the following individuals who have significantly contributed to the completion of this work.

Professor Robert O. Barr, of the Electrical Engineering and Systems Science department at Michigan State University, served as major professor and primary advisor. It was under his tutelage, encouragement and assistance that this work came to fruition.

Dr. Harvey Davis, Professor of Mathematics, Dr. Stuart Gage, Professor of Entomology, Dr. John Kreer, Professor of Electrical Engineering and Dr. Tom Manetsch, Professor of Systems Science, each from Michigan State University served as committee members and provided immeasurable technical and moral support.

Professor Lois Stegman, of Western Michigan University, serving as an outside reader, provided considerable help in the final critique.

TABLE OF CONTENTS

OVERVIEW OF RESEARCH 1

INTRODUCTION 1

GENERALITY OF NON CHRONOLOGICAL TIME 5

THE HEAT UNIT HYPOTHESIS 8

APPLICATION PROBLEMS 14

SYSTEM DEFINITION 19

TIME SCALE AND HORIZONS 21

THE SYSTEM METHODOLOGY 29

THE RESEARCH OBJECTIVES 32

Determination of the temperature profile . . 32
Parameter determination 34
Foundations and solutions 35
Stochastic modelling 36

SUMMARY 38

LITERATURE SURVEY 39

INTRODUCTION 39

HISTORICAL PERSPECTIVE 41

MODERN APPLICATIONS 48

THEORETICAL AND MODELING QUESTIONS 57

THE COMPUTATION OF HEAT UNITS IN BIOLOGICAL SYSTEMS . . 61

INTRODUCTION 61

THE SINUSOID MODEL 64

STATEMENT OF THE PROBLEM 65

APPLYING HIGH-LOW DATA TO THE SINUSOID MODEL . . . 67

THE STATISTICAL INTERPRETATION 70

THE NATURE OF THE HEAT UNIT MODEL 77

THE CASE OF DISCRETE DATA 81

DEGREE-DAYS IN THE SINUSOID MODEL 87

THE DOUBLE THRESHOLD MODEL 92

APPLICATION OF THE SINUSOID MODEL 103

NUMERICAL COMPUTATION OF THE PHYSIOLOGICAL TIME . 106

CONCLUSION 112

SOLVING SYSTEMS USING THE HEAT UNIT HYPOTHESIS 113

INTRODUCTION 113

AN ELEMENTARY INTERPRETATION 115

THE HEAT UNIT HYPOTHESIS 118

EXAMPLE ONE 127

EXAMPLE (cont'd) 148

RUNGE-KUTTA METHODS 153

CURVE FITTING METHODS 158

EXAMPLE TWO 159

PARAMETER

N

P

C

T

T

C

T

F

C

S

H

E

NONDAILY

N

T

T

T

E

E

T

T

Sum

PARAMETER ESTIMATION 174

INTRODUCTION 174

ENDOGENOUS AND EXOGENOUS SYSTEM PARAMETERS 177

DEFINING THE ENDOGENOUS SYSTEM PARAMETERS 179

TYPE II MODELS 180

TYPE III MODELS 186

TYPE IV MODELS 189

DETERMINING THE ENDOGENOUS PARAMETERS 198

IDENTIFICATION FOR TYPE I MODELS 200

PERIODIC TEMPERATURE PROFILES 202

CONSTANT MAGNITUDE TEMPERATURE PROFILES 205

SINUSOIDAL TEMPERATURE PROFILES 208

THE DATA-INDEPENDENT APPROXIMATION 214

EXAMPLE 221

NONANALYTIC OPTIMIZATION METHODS 227

INTRODUCTION 227

THE ENDOGENOUS TYPE I PROBLEM 228

THE ENDOGENOUS TYPE III MODEL 257

THE POWELL OPTIMIZATION PROCEDURE 259

ENDOGENOUS TYPE II AND IV MODELS 263

EXOGENOUS MODELS 270

THE LOGISTIC MODEL 271

THE LOTKA-VOLTERRA MODEL 274

SUMMARY OF NUMERICAL TECHNIQUES 280

STOCHAST

IN

TH

ST

TH

TH

EV

EL

FE

EL

ST

ST

ST

EX

CO

CONTROL

IN

TH

TH

CO

AP

A

AP

CO

STOCHASTIC WEATHER GENERATION 282

INTRODUCTION 282

THE STOCHASTIC DRIVER 284

STOCHASTIC TEMPERATURE GENERATORS 287

THE DETERMINISTIC TEMPERATURE GENERATOR 290

THEORETICAL DISTRIBUTION TEMPERATURE GENERATORS . 291

EMPIRICAL DISTRIBUTION TEMPERATURE GENERATORS . . 298

EXAMPLE ONE 302

FEEDBACK GENERATORS 305

EXAMPLE TWO 315

STATISTICAL PROPERTIES OF PHYSIOLOGICAL TIME . . . 317

STEADY STATE PHYSIOLOGICAL TIME PROPERTIES 326

STATISTICAL PROPERTIES OF POPULATION SIZE 348

EXAMPLE THREE 349

CONCLUSION 356

CONTROL SYSTEMS 360

INTRODUCTION 360

THE COMPARTMENTAL MODEL 363

THE LOTKA-VOLTERRA MODEL 369

CONTROL PROBLEMS 372

THE CROP INTRODUCTION TIME PROBLEM 377

A CONTROL EXAMPLE 383

IMPLEMENTING CONTROLS 414

CONCLUSION 445

CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER STUDY . . . 446

INTRODUCTION 446

RESULTS OF THIS STUDY 448

RECOMMENDATIONS FOR FURTHER STUDY 451

Theoretical Generalizations 451
Theoretical Extensions. 452
Applying the results 454
Applications 455

CONCLUSION 457

Table 2.1

Table 2.2

Table 2.3

Table 3.1

Table 3.2

Table 3.3

Table 3.4

Table 3.5

Table 3.6

Table 3.7

Table 3.8

Table 4.1

Table 4.2

Table 4.3

LIST OF TABLES

Table 2.1:	List of publications reviewed in study .	50
Table 2.2:	Growth of Relevant Articles	51
Table 2.3:	Classification of publication by application	51
Table 3.1	EIGHT-POINT temperature record for a single daily cycle.	66
Table 3.2	HIGH-LOW temperature record for a single daily cycle.	66
Table 3.3:	EIGHT-POINT temperature data	73
Table 3.4:	HIGH-LOW Temperature data	74
Table 3.5:	MEAN Calculations using EIGHT-POINT data, HIGH-LOW data (with the SINUSOID model)	75
Table 3.6:	STANDARD DEVIATION Calculations using EIGHT-POINT data, HIGH-LOW data (with the SINUSOID model)	76
Table 3.7:	DEGREE-DAY Calculations using Straight HIGH-LOW data (Compared against EIGHT-POINT data)	85
Table 3.8:	DEGREE-DAY Calculations using HIGH-LOW with SINUSOID Model (Compared against EIGHT-POINT data)	105
Table 4.1:	Absolute error obtained by solving a population model using the Euler method.	150
Table 4.2:	Relative error obtained by solving a population model using the Euler method.	150
Table 4.3:	Absolute error obtained by solving a population model using the Runge-Kutta method.	156

Table 4.4: Comparison of integration techniques of the logistic model. 168

Table 5.1 Summary of Type I model characteristics. 191

Table 5.2: Summary of Type II model characteristics. 192

Table 5.3: Summary of Type III model characteristics. 193

Table 5.4 Summary of Type IV model characteristics. 194

Table 5.5 Summary of general continuous model characteristics. 196

Table 5.6: Summary Physiological time results for the sinusoidal temperature profile model. 212

Table 5.7 Summary Action set Norm results for the sinusoidal temperature model. 213

Table 5.8 Threshold T_0 as a function of parameter a determined by Algorithm 5.5. 223

Table 5.9 Parameter a as a function of threshold T_0 as determined by equation 5.47. . . 223

Table 6.1: Sample population data used in Type I model example. 235

Table 6.2: Results of simulations using Algorithms 6.1 on the data of Table 6.2 for 40 different temperature thresholds T_0. . 236

Table 6.3 Optimal results using the first five data points of Table 6.1. 244

Table 6.4 Optimal results using the first eight data points of Table 6.1. 244

Table 6.5 Optimal results using the first eleven data points of Table 6.1. 247

Table 6.6 Optimal results using the first fourteen data points of Table 6.1. 249

Table 6.7 Optimal results using the seventeen data points of Table 6.1. 251

Table

Table

Table

Table

Table

Table 6.8 Optimal rms values along with the corresponding T_0 and r values as a function of number of data points q. . . 253

Table 7.1: Daily Mean Temperature Empirical History 289

Table 7.2: Daily Temperature Range Empirical History 289

Table 7.3: Mean/Standard deviation comparisons for each temperature measure using various generators. (Using 100 trials.) . . . 304

Table 8.1 Values of parameters in the introduction time control problem. 387

Figure 1
Figure 1
Figure 1
Figure 1
Figure 1
Figure 1
Figure 3
Figure 3
Figure 3
Figure 3
Figure 3
Figure 3
Figure 3
Figure 4

LIST OF FIGURES

Figure 1.1:	A possible multiple threshold model under the heat unit hypothesis using constant thresholds.	12
Figure 1.2:	A possible multiple threshold model under the heat unit hypothesis using time varying thresholds.	13
Figure 1.3	The average mean temperature for Lansing, Michigan (1961 to 1985)	25
Figure 1.4	The average temperature range for Lansing, Michigan (1961 to 1985)	26
Figure 1.5	The standard deviation of mean temperatures for Lansing, Michigan (1961 to 1985)	27
Figure 1.6	The standard deviation of temperature ranges for Lansing, Michigan (1961 to 1985).	28
Figure 3.1	An idealized daily temperature profile.	69
Figure 3.2	The Biological Activity $\beta(T)$.	80
Figure 3.3	Relationship of physiological time to the temperature profile T(t).	80
Figure 3.4	All possible cases for the sinusoidal heat unit model - TYPE I.	86
Figure 3.5	The active region for the TYPE III model.	100
Figure 3.6a:	The first four cases of a TYPE III model using the sinusoidal approximation. . .	101
Figure 3.6b:	Cases 5 and 6 for the TYPE III model using a sinusoidal approximation. . . .	102
Figure 4.2	The relationship between fundamental heat-unit functions.	123

Figure
Figure
Figure
Figure
Figure
Figure
Figure
Figure
Figure
Figure
Figure
Figure
Figure
Figure

Figure 4.3: A cubic temperature profile. 124

Figure 4.4: A family of physiological time curves for various thresholds. 131

Figure 4.5: A family of physiological horizon times for various initial times. 132

Figure 4.6: A family of physiological horizon times for various thresholds. 133

Figure 4.7: A family of exponential model population curves for various growth rates. 137

Figure 4.8: A family of classical logistic model population curves for various growth rates. 138

Figure 4.9a: A family of heat-unit based logistic model population curves for various growth rates. 139

Figure 4.9b: A family of heat-unit based logistic model population curves for various growth rates. 140

Figure 4.10: The absolute error in a population model using a Eulerian approximation as a function of threshold. δ=1. 151

Figure 4.11: The relative error in a population model using a Eulerian approximation as a function of threshold. δ=1. 152

Figure 4.12: The absolute error of a population model using a Runge-Kutta approximation as a function of threshold. δ=1. 157

Figure 4.13: The empirical temperature profile for Lansing, Michigan (1985) superimposed on the sinusoidal curve fitting approximation. 165

Figure 4.14: Physiological time as computed using empirical data and a sinusoidal curve fitting approximation. 166

Figure 4.15: Population size at horizon (t_f=300 days) using an empirical temperature profile and a sinusoidal curve fitting profile as a function of threshold. . . 167

Figure 4

Figure 4

Figure 4.

Figure 4.

Figure 5.1

Figure 5.2

Figure 5.3

Figure 5.4

Figure 5.5

Figure 5.6

Figure 5.7

Figure 6.1

Figure 6.2

Figure 6.3

Figure 6.4

Figure 4.16: Population metric as computed using empirical temperature profile and a sinusoidal curve fitting approximation. 170

Figure 4.17a: Difference between population metrics as computed by the empirical temperature profile and a curve fitting approximation. $T_0=35^o$ F. 171

Figure 4.17b: Difference between population metrics as computed by the empirical temperature profile and a curve fitting approximation. $T_0=40^o$ F. 172

Figure 4.17c: Difference between population metrics as computed by the empirical temperature profile and a curve fitting approximation. $T_0=45^o$ F. 173

Figure 5.1 The Biological Activity function for a Type II model. 185

Figure 5.2 The Biological Activity function for a general piecewise continuous model. . . 185

Figure 5.3 The Biological Activity function for a Type III model. 188

Figure 5.4 The Biological Activity function for a Type IV model. 188

Figure 5.5 A typical periodic temperature profile and the corresponding physiological time and action sets. 204

Figure 5.6 Relationship between the Physiological time $\tau(t)$ and the Action set Norm $\|\Gamma(t)\|$. 220

Figure 5.7: Plots of Tables 5.6 and 5.7 showing point of intersection for data independent solution. 224

Figure 6.1: Graphs of objective rms as a function of T_0, reproducing Table 7.5. 237

Figure 6.2 Graph comparing empirical data against modelled population metric. q=5. . . . 245

Figure 6.3 Graph comparing empirical data against modelled population metric. q=8. . . . 246

Figure 6.4 Graph comparing empirical data against modelled population metric. q=11. . . . 248

Figure
Figure
Figure
Figure
Figure
Figure
Figure
Figure
Figure
Figure
Figure
Figur
Figur
Figur
Figur
Figur
Figur

Figure 6.5 Graph comparing empirical data against modelled population metric. q=14. . . . 250

Figure 6.6 Graph comparing empirical data against modelled population metric. q=17. . . . 252

Figure 6.7: Plot of objective rms against horizon time q. 254

Figure 6.8: Plot of optimal threshold T_0 against horizon time q. 255

Figure 6.9 Plot of optimal rate r against horizon time q. 256

Figure 7.1: Mean Temperature Distribution for the June 1 Time Slice. 294

Figure 7.2: Temperature Range Distribution for the June 1 Time Slice. 295

Figure 7.3: The deterministic generator. 297

Figure 7.4: The theoretic distribution fitting generator. 297

Figure 7.5: The empirical distribution preserving generator. 301

Figure 7.6a: Mean temperature distribution for the June 1 time slice using the deterministic generator. 308

Figure 7.6b: Mean temperature distribution for the June 1 time slice using the theoretic distribution fitting generator. 309

Figure 7.6c: Mean temperature distribution for the June 1 time slice using the empirical distribution fitting generator. 310

Figure 7.7a: Temperature range distribution for the June 1 time slice using the deterministic generator. 311

Figure 7.7b: Temperature range distribution for the June 1 time slice using the theoretic distribution fitting generator. 312

Figure 7.7c: Temperature range distribution for the June 1 time slice using the empirical distribution fitting generator. 313

Figure 7.8: The Feedback Temperature Generator . . . 314

Figure 7.9: Plot of objective rms against parameter a for a feedback temperature generator. Lansing, Michigan (1985) . . 316

Figure 7.10a: Mean physiological time plotted against chronological time for various temperature thresholds. 320

Figure 7.10b: Mean physiological time plotted against chronological time for various temperature thresholds. 321

Figure 7.10c: Mean physiological time plotted against chronological time for various temperature thresholds. 322

Figure 7.11a: Standard deviation of physiological time plotted against chronological time for various temperature thresholds. 323

Figure 7.11b: Standard deviation of physiological time plotted against chronological time for various temperature thresholds. 324

Figure 7.11c: Standard deviation of physiological time plotted against chronological time for various temperature thresholds. 325

Figure 7.12: Mean final physiological time as a function of threshold for various initial times. T=40° F. 344

Figure 7.13: Standard deviation of final physiological time as a function of threshold for various initial times. T=40° F. 345

Figure 7.14: Graph of the function defined in equation 7.44 along with various approximations to f(z). 346

Figure 7.15: The effective ratio over k=5 cycles using various temperature thresholds. . 347

Figure 7.16a: Mean population metric as a function of time for various thresholds. 350

Figure 7.16b: Mean population metric as a function of time for various thresholds. 351

Figure 7.16c: Mean population metric as a function of time for various thresholds. 352

Figure 7.17a: Standard deviation of the population metric as a function of time for various thresholds. 353

Figure 7.17b: Standard deviation of the population metric as a function of time for various thresholds. 354

Figure 7.17c: Standard deviation of the population metric as a function of time for various thresholds. 355

Figure 8.1: Block diagram of the heat-unit population model. 362

Figure 8.2: Block diagram of the compartmental model. 362

Figure 8.3: Block diagram of the Leslie model. . . . 368

Figure 8.4: Block diagram of the predator/prey interaction model. 368

Figure 8.5: The biological activity function defined by equation [8.16]. 381

Figure 8.6: Graph of the Biological Activity $\beta_1(T)$ of equation [8.20]. 382

Figure 8.7: Graph of the Biological Activity $\beta_2(T)$ of equation [8.21]. 382

Figure 8.8: Population metric over time with uncontrolled predator, uncontrolled prey. $T_{10}=15^{o}$, $T_{11}=40^{o}$ F. 388

Figure 8.9: Population metric over time with uncontrolled predator, uncontrolled prey. $T_{10}=25^{o}$, $T_{11}=40^{o}$ F. 389

Figure 8.10a Harvest predator/prey mean population metrics as a function of introduction time with an uncontrolled predator. $T_0=15^{o}$, $T_1=40^{o}$ F. 390

Figure 8.10b Harvest predator/prey mean population metrics as a function of introduction time with an uncontrolled predator. $T_0=15^{o}$, $T_1=45^{o}$ F. 391

Figure 8.10c Harvest predator/prey mean population metrics as a function of introduction time with an uncontrolled predator. $T_0=15^o$, $T_1=50^o$ F. 392

Figure 8.11a Harvest predator/prey mean population metrics as a function of introduction time with an uncontrolled predator. $T_0=20^o$, $T_1=40^o$ F. 393

Figure 8.11b: Harvest predator/prey mean population metrics as a function of introduction time with an uncontrolled predator. $T_0=20^o$, $T_1=45^o$ F. 394

Figure 8.11c Harvest predator/prey mean population metrics as a function of introduction time with an uncontrolled predator. $T_0=20^o$, $T_1=50^o$ F. 395

Figure 8.12a Harvest predator/prey mean population metrics as a function of introduction time with an uncontrolled predator. $T_0=25^o$, $T_1=40^o$ F. 396

Figure 8.12b Harvest predator/prey mean population metrics as a function of introduction time with an uncontrolled predator. $T_0=25^o$, $T_1=45^o$ F. 397

Figure 8.12c Harvest predator/prey mean population metrics as a function of introduction time with an uncontrolled predator. $T_0=25^o$, $T_1=50^o$ F. 398

Figure 8.13a Comparison predator/prey mean population metrics summarizing Figures 8.10. $T_{10}=15^o$ F. 399

Figure 8.13b Comparison predator/prey mean population metrics summarizing Figures 8.11. $T_{10}=20^o$ F. 400

Figure 8.13c Comparison predator/prey mean population metrics summarizing Figures 8.12. $T_{10}=25^o$ F. 401

Figure 8.14a Comparison predator/prey mean population metrics summarizing Figures 8.10a, 8.11a, 8.12a. $T_{11}=40^o$ F. 402

Figure 8.14b Comparison predator/prey mean population metrics summarizing Figures 8.10b, 8.11b, 8.12b. $T_{11}=45^o$ F. 403

Figure 8.14c Comparison predator/prey mean population metrics summarizing Figures 8.10c, 8.11c, 8.12c. $T_{11}=50^{o}$ F. 404

Figure 8.15a Harvest predator/prey variation in population metrics as a function of introduction time with an uncontrolled predator. $T_{10}=15^{o}$, $T_{11}=40^{o}$ F. 405

Figure 8.15b Harvest predator/prey variation in population metrics as a function of introduction time with an uncontrolled predator. $T_{10}=15^{o}$, $T_{11}=45^{o}$ F. 406

Figure 8.15c Harvest predator/prey variation in population metrics as a function of introduction time with an uncontrolled predator. $T_{10}=15^{o}$, $T_{11}=50^{o}$ F. 407

Figure 8.16a Harvest predator/prey variation in population metrics as a function of introduction time with an uncontrolled predator. $T_{10}=20^{o}$, $T_{11}=40^{o}$ F. 408

Figure 8.16b Harvest predator/prey variation in population metrics as a function of introduction time with an uncontrolled predator. $T_{10}=20^{o}$, $T_{11}=45^{o}$ F. 409

Figure 8.16c Harvest predator/prey variation in population metrics as a function of introduction time with an uncontrolled predator. $T_{10}=20^{o}$, $T_{11}=50^{o}$ F. 410

Figure 8.17a Harvest predator/prey variation in population metrics as a function of introduction time with an uncontrolled predator. $T_{10}=25^{o}$, $T_{11}=40^{o}$ F. 411

Figure 8.17b Harvest predator/prey variation in population metrics as a function of introduction time with an uncontrolled predator. $T_{10}=25^{o}$, $T_{11}=45^{o}$ F. 412

Figure 8.17c Harvest predator/prey variation in population metrics as a function of introduction time with an uncontrolled predator. $T_{10}=25^{o}$, $T_{11}=50^{o}$ F. 413

Figure 8.18 Population metric over time with controlled predator, uncontrolled prey. 418

Figure 8.19 Population metric over time with uncontrolled predator, controlled prey. 419

Figure 8.20 Population metric over time with controlled predator, controlled prey. . 420

Figure 8.21a Harvest predator/prey mean population metrics as a function of introduction time with a controlled predator. $T_{10}=15^{o}$, $T_{11}=40^{o}$ F. 421

Figure 8.21b Harvest predator/prey mean population metrics as a function of introduction time with a controlled predator. $T_{10}=15^{o}$, $T_{11}=45^{o}$ F. 422

Figure 8.21c Harvest predator/prey mean population metrics as a function of introduction time with a controlled predator. $T_{10}=15^{o}$, $T_{11}=50^{o}$ F. 423

Figure 8.22a Harvest predator/prey mean population metrics as a function of introduction time with a controlled predator. $T_{10}=20^{o}$, $T_{11}=40^{o}$ F. 424

Figure 8.22b Harvest predator/prey mean population metrics as a function of introduction time with a controlled predator. $T_{10}=20^{o}$, $T_{11}=45^{o}$ F. 425

Figure 8.22c Harvest predator/prey mean population metrics as a function of introduction time with a controlled predator. $T_{10}=20^{o}$, $T_{11}=50^{o}$ F. 426

Figure 8.23a Harvest predator/prey mean population metrics as a function of introduction time with a controlled predator. $T_{10}=25^{o}$, $T_{11}=40^{o}$ F. 427

Figure 8.23b Harvest predator/prey mean population metrics as a function of introduction time with a controlled predator. $T_{10}=25^{o}$, $T_{11}=45^{o}$ F. 428

Figure 8.23c Harvest predator/prey mean population metrics as a function of introduction time with a controlled predator. $T_{10}=25^{o}$, $T_{11}=50^{o}$ F. 429

Figure 8.24a Comparison predator/prey mean population metrics summarizing Figures 8.21. $T_{10}=15^{o}$ F. 430

Figure 8.24b Comparison predator/prey mean population metrics summarizing Figures 8.22. $T_{10}=20^{o}$ F. 431

Figure 8.24c Comparison predator/prey mean population metrics summarizing Figures 8.23. $T_{10}=25^{o}$ F. 432

Figure 8.25a Comparison predator/prey mean population metrics summarizing Figures 8.21a, 8.22a, 8.23a. $T_{11}=40^{o}$ F. 433

Figure 8.25b Comparison predator/prey mean population metrics summarizing Figures 8.21b, 8.22b, 8.23b. $T_{11}=45^{o}$ F. 434

Figure 8.25c Comparison predator/prey mean population metrics summarizing Figures 8.21c, 8.22c, 8.24c. $T_{11}=50^{o}$ F. 435

Figure 8.26a Harvest predator/prey variation in population metrics as a function of introduction time with a controlled predator. $T_{10}=15^{o}$, $T_{11}=40^{o}$ F. 436

Figure 8.26b Harvest predator/prey variation in population metrics as a function of introduction time with a controlled predator. $T_{10}=15^{o}$, $T_{11}=45^{o}$ F. 437

Figure 8.26c Harvest predator/prey variation in population metrics as a function of introduction time with a controlled predator. $T_{10}=15^{o}$, $T_{11}=50^{o}$ F. 438

Figure 8.27a Harvest predator/prey variation in population metrics as a function of introduction time with a controlled predator. $T_{10}=20^{o}$, $T_{11}=40^{o}$ F. 439

Figure 8.27b Harvest predator/prey variation in population metrics as a function of introduction time with a controlled predator. $T_{10}=20^{o}$, $T_{11}=45^{o}$ F. 440

Figure 8.27c Harvest predator/prey variation in population metrics as a function of introduction time with a controlled predator. $T_{10}=20^{o}$, $T_{11}=50^{o}$ F. 441

Figure 8.28a Harvest predator/prey variation in population metrics as a function of introduction time with a controlled predator. $T_{10}=25^{o}$, $T_{11}=40^{o}$ F. 442

Figur

Figur

Figure 8.28b Harvest predator/prey variation in population metrics as a function of introduction time with a controlled predator. $T_{10}=25^{o}$, $T_{11}=45^{o}$ F. 443
Figure 8.28c Harvest predator/prey variation in population metrics as a function of introduction time with a controlled predator. $T_{10}=25^{o}$, $T_{11}=50^{o}$ F. 444

LIST OF ALGORITHMS

Algorithm 3.1 Computing the physiological time assuming a Type I model with a given instantaneous temperature profile T(t). 108

Algorithm 3.2: Computing the physiological time assuming a Type I model with daily mean temperature M(t) and daily temperature range R(t) profiles. 109

Algorithm 3.3: Computing the physiological time assuming a Type III model with a given instantaneous temperature profile T(t). 110

Algorithm 3.4 Computing the physiological time assuming a Type III model with daily mean temperature M(t) and daily temperature range R(t) profiles. . . . 111

Algorithm 4.1 Computing the critical points for a cubic temperature profile. 125

Algorithm 4.2 Computing the physiological time from a general cubic temperature profile given the critical points (as determined by Algorithm 4.1). 126

Algorithm 4.3 Numerically integrating the logistic population model using an Eulerian approximation. 149

Algorithm 4.4 Numerically integrating a generic population model (expressed as equation [4.16]) using the Runge-Kutta approximation. 155

Algorithm 5.1 Computing the threshold T_0 for a Type I model using the data independent solution approximation. 219

Algorithm 6.1: Computing critical elements for the Type I model. 234

Algorithm 6.2: The Golden Section Algorithm 243

Algorithm 6.3 Computing the physiological time for a Type III model. 258

Algorithm 6.4 Computing physiological times for multi-species population Type I models. . . . 262

Algorithm 6.5 Calculating the Logistic population model and objective function for any exogenous model. 273

Algorithm 6.6 Computing physiological times for multi-species population Type I models. 277

Algorithm 6.7 Calculating the objective rms function for the multi-species population (exponential) model. 278

Algorithm 6.8 Calculating the objective rms function for any population (both endogenous and exogenous) model. 279

Algorithm 7.1: Generating Gaussian normalized variates z and rectified variates x. 296

Algorithm 7.2: Generating a gamma variate x. 296

Algorithm 7.3: The Empirical Distribution Fitting Algorithm. 300

Algorithm 8.1: Computing the physiological time for the multiple growth rate system of equation [8.16]. 381

INTRO

This
popul
histo
last
appli
manne
the
acces
throu
than
"do i
model

Biolo
subje
analo
momen
analo
force

CHAPTER ONE

OVERVIEW OF RESEARCH

INTRODUCTION

This dissertation deals with the problem of modeling population systems. While this problem has been of historical interest for many years, it has been only in the last decade that the modeling of these systems has been applied to practical situations in a somewhat systematic manner. This has only become possible since solutions to the more precise computational models have become accessible. However, this accessibility has largely been through the availability of personal computer power rather than theoretical insight. For this reason, many researchers "do it", but there is little uniformity of approach and the models are somewhat ad hoc in nature.

Biological systems, like their physical counterparts, are subject to fundamental "laws" and principles. There are analogies to the conservation principles of mass, energy and momentum for mechanical systems. Indeed, there are also analogies to force and velocity. Isaac Newton stated that force equals the time rate of change in momentum:

[1.1]

For popu

the net

[1.2]

Early p

Newton

a biol

glossed

However

that th

success

model i

problem

the ou

popula

concept

$$[1.1] \qquad F = \frac{d}{dt}\, p(t)$$

where $p(t) = m(t) \cdot v(t)$

$= \text{momentum},$

$m(t) = \text{mass at time } t,$

$v(t)$ is the velocity at time t.

For population systems we say the Biological Activity equals the net population rate of change:

$$[1.2] \qquad \beta = \frac{1}{x(t)} \cdot \frac{d}{dt}\, x(t)$$

Early physicists talked around the notion of "force" before Newton made the definition precise. Similarly, the idea of a biological activity function has traditionally been glossed over and misunderstood.

However, biological systems have several major complications that their physical counterparts either do not have or can successfully be ignored. First, the state of a population model is usually dynamic. While most elementary mechanics problems assume the state (solid, liquid or gas) is fixed at the outset, the characteristics of a given species' population will change significantly over time. From conception through maturity to death this individual will

eat, rep

its devel

is throu

1345), wh

models.

these mo

by eleme

varying a

since th

solutions

Therefor

problems

more adv

heteroger

individua

solve th

models,

decompo

elementa

models ca

is an exc

It should

is almos

seasona

Predator/

eat, reproduce and grow differently from stage to stage in its development. A traditional way to model this phenomena is through the use of a so-called Leslie model (Leslie 1945), which is a special case of the class of compartmental models. The Leslie approach is not entirely bad, since these models are inherently linear, and therefore solvable by elementry techniques. However, they are usually time varying and explicit solutions are exceedingly rare. Also, since the population is represented by a vector, numerical solutions are inevitable.

Therefore, while any physics student reduces all dynamics problems to "F=ma", the biologist is confronted with a more advanced complication. Only by assuming a uniformly heterogenous population, i.e. there is a fixed proportion of individuals in each stage over time, can the biology student solve the most elementary of population problems. These models, which consider the population as a whole without decomposing it into stages, are the source of many elementary calculus examples. The exponential and logistic models can be solved explicitly and the Lotka-Volterra model is an excellent source of nonlinear demonstrations.

It should be noted that the uniform heterogeneity assumption is almost always inappropriate. For instance, in modeling seasonal crop production, the crops as well as the predator/pest growth stage, is homogenous by design! Crops

are plan

progress

Therefore

Lotka-Vo

therefor

so, thes

followin

understoo

apply t

compart

systemati

The seco

they are

tradition

is the c

populati

were the

time. T

actually

life in

populatio

can cloud

work is

physiolog

chronolog

are planted at the same time in the same (seed) stage and progress in essentially lock-step fashion until harvest. Therefore, strictly speaking, the exponential, logistic and Lotka-Volterra models are macro representations and therefore applicable in only limited circumstances. Even so, these will be the models of demonstration in the following research because they are elementary and well understood. It is left as the future "next obvious step" to apply the results of this research to the precise compartmental models and validate the results in a systematic manor.

The second problem unique to biological systems is that they are more sensitive to environmental factors than traditional physical systems. Indeed, it has been shown and is the central thesis to this work that a large class of population systems behave as if the effective temperature were the independent variable rather that chronological time. This fact need not be taken on faith and will actually be inferred by more primitive principles. Even-so, life in general is not even this simple. Precipitation, population density and a host of other environmental factors can cloud the issue. However, the primary tenant of this work is that a given species' population has its own physiological time scale (based on temperature) rather than chronological (mechanically based) time.

GENERALITY

The notion
time base
Physical
too. Pro
systems is
the Lorem
speeds ap
"time dila

Every phy
from the

Explicit

[1.3]

This int
way time

GENERALITY OF NON CHRONOLOGICAL TIME

The notion that a system's state varies with an endogenous time base is not confined to only biological systems. Physical and mechanical systems exhibit these properties too. Probably the most famous of these indigenous time systems is that from relativistic kinematics. According to the Lorentz transformation, a mechanical system moving at speeds approaching that of the speed of light will exhibit "time dilation" and "length contraction".

Every physics student is taught that "system time" differs from the "external time" by the magic factor:

$$\sqrt{(1-v^2/c^2)}$$

where v = velocity of the system relative to the "stationary" observer,

c = speed of light.

Explicitly,

[1.3] $$t_{sys} = t_{ext} \cdot \sqrt{(1-v^2/c^2)}$$

This introduces a paradox. Time within the system acts the way time "should" act. That is, an observer as part of the

system sees

Newtonian r

physics cr

considerab

Special re

Special pr

this syste

quantity, t

course, le

General Rel

Christoffe

continuum

of populat

Let us con

two system

In both sy

mechanica

time in a

Newton or

world "see

There is

systems.

depends on

time depe

system sees physics acting non-relativistically: normal, Newtonian mechanics. However, to the external observer, physics on the moving frame of reference has changed considerably. It now behaves according to early Einstein (special relativity) rather than classical Newton.

Special problems occur when one introduces dynamics into this system. Since the velocity v itself is a dynamic quantity, the Lorentz formulation no longer holds. This, of course, led to Einstein's celebrated work on the theory of General Relativity. Fortunately for us, an understanding of Christoffel symbols, tensor calculus and the space-time continuum is not necessary for the relatively tame problem of population dynamics!

Let us consider the similarities and differences between the two systems: relativistic mechanics and population dynamics. In both systems the internal entity (moving observer in the mechanical, species population in the biological) "sees" time in a classical form. His (internal) world behaves as Newton or Volterra described. Externally, of course, the world "sees" the behavior in a non-classical sense.

There is also a significant difference between these two systems. In the case of special relativity, system time depends on speed. In the case of population, physiological time depends on temperature. The important distinction is

that of v

internal

action by

dynamics

(velocity)

this lea

temperatu

certainly

independer

to be exc

there exi

compare ag

that of velocity compared to temperature. Velocity is an internal entity. How fast the system moves depends on action by the system itself. Thus, in studying system dynamics it is necessary to relate an internal quantity (velocity) to internal (system) time. As mentioned before, this leads to general relativity. On the other hand, temperature in the population model is generally external, certainly for cold-blooded organisms. Temperature is independent of system performance and can safely be assumed to be exogenous. Thus, even though it might be desirable, there exists no "general" temperature modeling theory to compare against relativity.

THE HEAT UNI

The notion
temperatur
Generally
of a temp
dormant
proportio

This biol
temperatu
time t.
called the
cumulativ
written as

[1.4]

The set
interval(
constant
and is ca

THE HEAT UNIT HYPOTHESIS

The notion that there exists a physiological time based upon temperature will be referred to as the heat unit hypothesis. Generally this will be interpreted as implying the existence of a temperature threshold T_0 below which the species lies dormant and above which the biological activity is proportional to the temperature in excess of this threshold.

This **biological activity** is in general a function of the **temperature profile** $T(t)$ which varies with **chronological time** t. As described above, there exists a function $\beta(T)$ called the activity function which depends directly on the cumulative temperature. Mathematically, this will be written as:

$$\beta(t) = r \cdot \int_{\Gamma(t)} [T(\mu) - T_0] \, d\mu \qquad [1.4]$$

The set $\Gamma(t)$ is called the **action set** and is simply the interval(s) of time where the system is non-dormant. The constant r is the proportionality constant inferred above and is called the **growth rate.**

It is important to note that $\beta(T)$ can be decomposed into the product of two factors: growth rate r and the factor depending only on temperature $T(t)$ and chronological time t:

$$[1.5] \qquad \tau(t) = \int_{\Gamma(t)} [T(\mu) - T_0] \, d\mu$$

Thus, we refer to it as the **physiological time** $\tau(t)$. The units of $\tau(t)$ are often degree-days because it represents the cumulative number of degrees acquired over a period of (chronological) time. The growth rate then must represent a measure of growth activity per unit of physiological time. Again, this is physiological time because the system by hypothesis is assumed to "see" in the internal classical sense.

This basic formula will be generalized in later chapters. However, it will be helpful to note several approaches that will not be taken first. There is no real reason to define the physiological time as the linear excess temperature. It could be just as valid to define either a quadratic measure such as:

$$[1.6] \qquad \tau(t) = \int_{\Gamma(t)} [T(\mu) - T_0]^2 \, d\mu$$

or

$$[1.7] \qquad \tau(t) = \left[\int_{\Gamma(t)} [T(\mu) - T_0]^2 \, d\mu \right]^{\frac{1}{2}}$$

The quadratic approach has been used in several ad hoc models in order to disproportionally weigh features of a logistic model at the mature growth stages. Indeed, this non-linear approach has met with moderate success for certain specific problems (Wolf, et al. 1986). Even so, the "correct" way to do this would be to compartmentalize the model.

As a general methodology, several authors have hypothesized multiple thresholds (Dintenfass, Bartell and Scott 1987), (Raworth 1984). That is to say, there can exist lethal thresholds below which the species dies as well as the dormancy threshold described earlier. A possible layered approach to this problem is illustrated in Figure 1.1. This is a truly valid approach and will be exploited in this dissertation. However, to attain a higher degree of realism using logistic models several authors (Lund and Herne 1980), (Wang 1960) have assumed these thresholds to vary with time. This is illustrated in Figure 1.2.

Using time varying thresholds, some computational algorithms can be implemented. Even so, there are significant problems with this approach. The first is theoretical. There is simply no basis from which to draw these inferences other than that they "work". Rather than do it by compartmentalizing the models, this variable threshold model simply assumes "dynamic variable constants" (or some such jargon) to reflect the staging effect of species population.

Time varying thresholds are inferior mechanisms for pragmatic reasons as well. When using them, there is no way to infer the best threshold curves. They simply reflect best guesses. For this reason, when introducing controlling or predator populations, the parameters used may no longer be appropriate or even feasible. Thus, in similar studies by different researchers, totally contradictory conclusions will invariably result! This is simply because each threshold will be defined differently in each study.

For these reasons, this study will focus on a "straight" definition of heat units (degree-days) using constant temperature thresholds in a linear excess temperature. It is hoped that this will lead to a systematic approach to temperature based biological models generally and to far fewer ad hoc type models. In this way, not only will a consistent set of assumptions be used, but also that a comparable set of results be established.

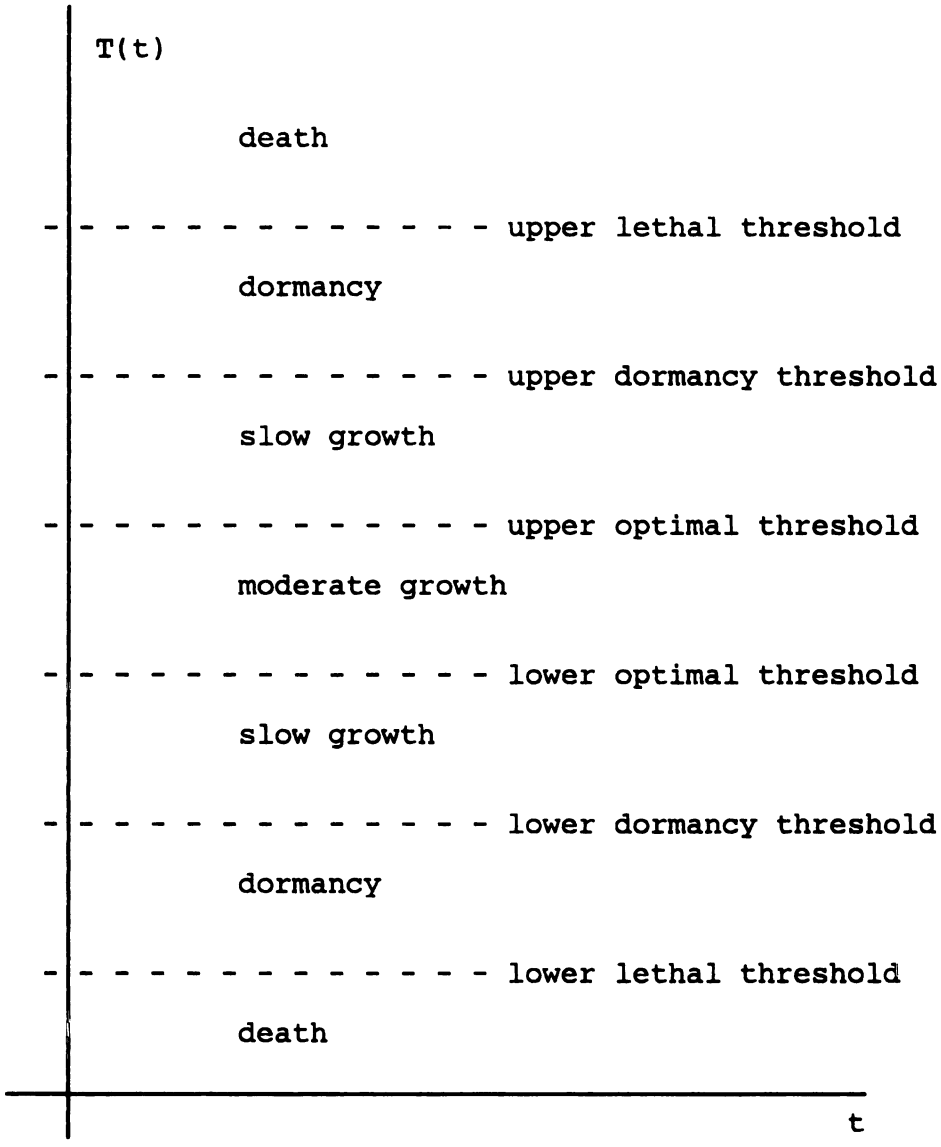


Figure 1.1: A possible multiple threshold model under the heat unit hypothesis using constant thresholds.

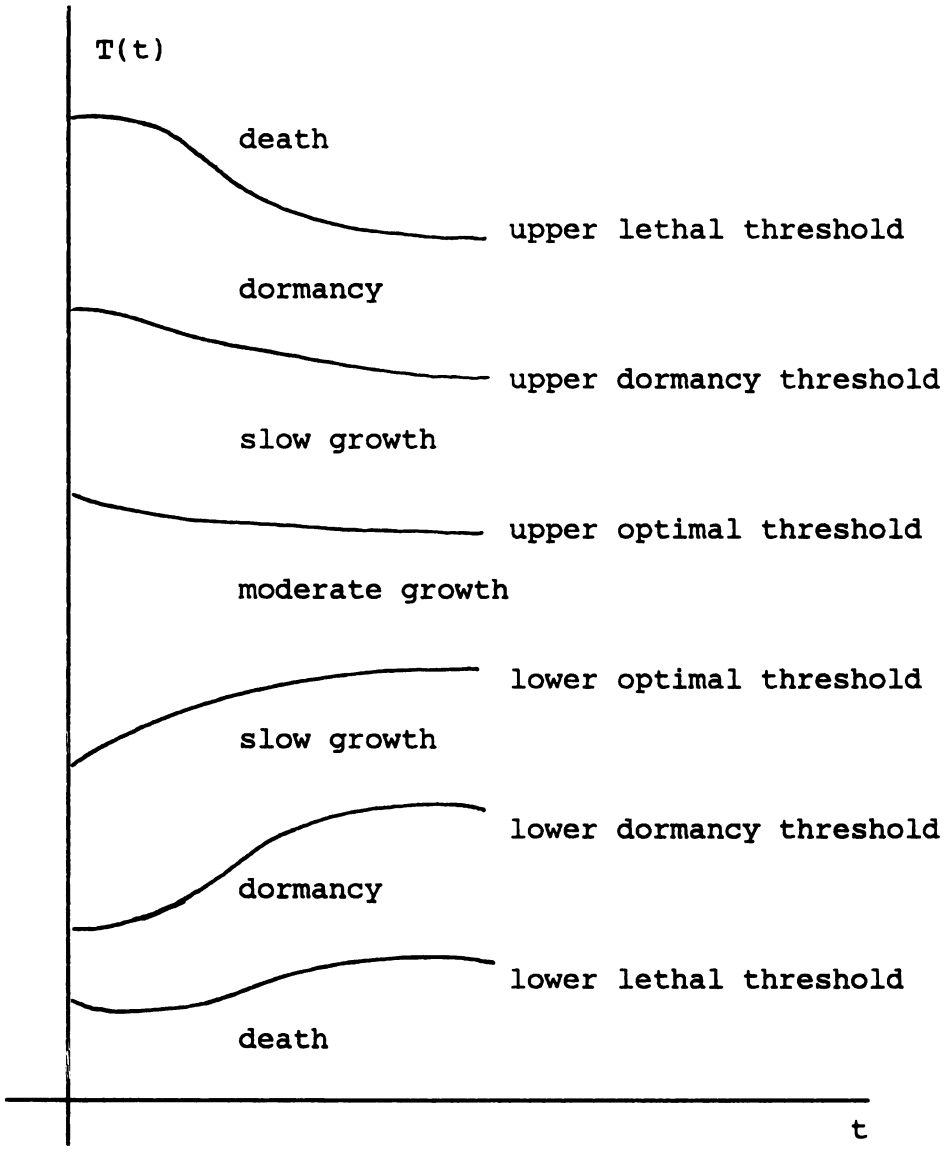


Figure 1.2: A possible multiple threshold model under the heat unit hypothesis using time varying thresholds.

APPLICATION PROBLEMS

The reason for developing accurate models of systems goes far beyond simple academic interest. Just as in physical systems where engineers have a long and proud tradition of system modeling, biological systems are studied primarily to facilitate better design and control. In population studies, this usually takes on one or the other of two forms. The first of these is system ecology and the second is that of agricultural production.

The key word in system ecological studies is "stability". For biological systems this is most usefully defined with a broader meaning than traditional physical systems. For instance, consider the following example:

> Given an existing system, supposedly modelled to perfection, a foreign species is introduced at a given time. The following questions arise:
>
> i. Does one population take over? That is, does it continually grow?
>
> ii. Does any population ever become extinct? Unlike physical systems, once a population is zero, it never returns!

It might be best to think of either of these eventualities as unfavorable, and thus "unstable".

iii. If either of these unfavorable circumstances arise, can a suitable protocol be found that will stabilize the situation?

iv. Can a balance be struck between system responsiveness and stability?

v. Is there a systematic means of measuring relative stability?

These are certainly exciting problems that deserve considerable attention because environmental impact studies are a very timely topic.

Agricultural production is not so concerned with stability. It would be unusual indeed for a farmer to object to driving a pest population to extinction on his farm! On the other hand, he would certainly be interested in maximizing the yield of his crop or perhaps optimizing his profits. Thus, his problem becomes one of seeking optimal control.

The modern agricultural planner has a multitude of possible controls at her disposal. They may be passive, such as simply varying the introduction time of the crop, or they

might be active where pesticides, herbicides and fertilizers might be applied. In either case, there is a cost associated with each option, a protocol for its application and a gain to be realized by its use. An objective function can be found and traditional optimization approaches applied.

A multitude of other possibilities also exists, but problems seem to fall into either the stability class, the optimal control class or both. Unfortunately, it is impossible to satisfactorily handle these problems until the modeling problem is well defined and solved. The intent of this dissertation is to systematically approach the modeling problem, and the ultimate goal is toward the solution of basic control problems. Needless to say, there is much work yet to be done here.

The classical approach to interacting populations is by using the Lotka-Volterra model. For non heat-unit systems this is well understood. Although it has no closed form solution, its nonlinear nature gives rise to the beauty of limit cycles, notions of asymptotic stability and system equilibrium. Also, solutions are obtainable only by numerical methods. For these reasons, upon introducing compartmental methods and heat-units, these models become oppressively complex.

The literature becomes rife with ad hoc examples at this point. Many observers feel that it is possible to begin with any single biological hypothesis, and with a variety of model interpretations attain almost any desired conclusion. Again, the need is for a systematic approach to the control as well as to the modelling problems.

There are two special requirements for accurate control systems. The first is that each interacting species' endogenous parameters be independently tuned. An **endogenous parameter** is one which is taken as species determined and independent of the environment or competing populations. For instance, the temperature threshold should depend only upon the particular species and should be independent of any other entity within the model. This disallows thresholds such as those illustrated in Figure 1.2.

The second obvious requirement for control problems is that they accurately reflect the temperature characteristics of the hypothesized environment. Useful models will have statistically accurate pictures of the temperature profile $T(t)$. While a simple history might be sufficient for estimates, control results will more likely be in the form of expected values (means) and confidence limits(standard deviations) about those values. The temperature will drive these variances.

Ideally the temperature profile can be represented by a **basis**, which would generate representative profiles from an ensemble of candidates of a stochastic process. In practice, this basis will need to be approximated from a finite but large historical database. Even so, the important point is that results will necessarily be statistical in nature and the temperature will drive these results.

The essence of this research is not to study the various types of control problems, or even attack a subclass of them. Even so, a control example will hopefully point the way toward further explorations in this most important area of study.

SYSTEM DEFINITION

In any system study there are of necessity certain basic underlying assumptions. Even though a sincere attempt will be made to keep the results as general as possible, there will need to be a distinct orientation in modeling perspective. Let us begin by describing the class of problems that we wish to consider.

Recall the sample problems defined earlier: environmental stability and agricultural production. It is obvious that time is on a more or less human scale and that neither the extremes of either micro-organisms with time horizons measured in hours or civilizations over centuries' is a concern. Also, it is apparent that we are discussing ordinary weather as the source of temperature fluxuation.

In the case of the environmental stability, the question is basically: "What happens when a foreign species is introduced?" It can further be made into a control problem if one asks: "What controls can be put on certain species within the system to force stability?" Any input to the system is strictly of a control nature and not overt or exogenous.

In the case of agricultural production one simply asks "what controls can lead to the best possible production results?"

If desired, a richer class of problems might require this to be done over a one year time horizon, or if a degree of stability is required here too, over an extended period of time.

TIME SCALE

Since thi

hypothesis

important

immediate

i.

ii.

iii.

iv.

v.

vi.

TIME SCALE AND HORIZONS

Since this is a study of models that invoke the heat-unit hypothesis, it should be evident that temperature is a most important factor. While this may be obvious, there are immediate questions that need be asked at the outset.

i. Does temperature drive the system or is it more akin to background noise that simply clouds the more significant phenomena?

ii. Is the temperature user-controlled or is it strictly an external contribution?

iii. Is the temperature deterministic or stochastic in nature?

iv. What is the frequency spectrum of the temperature like?

v. How is temperature data retained and how can data reduction be best achieved?

vi. Over what time horizon is the simulation to be operated?

From the

either ca

sense.

temperatur

stochastic

(daily HIG

Terrestri

dominated

annual (p

cycles toc

also disp

time horiz

or years,

The earth'

is not the

exhibited

The obv

temperatur

For instar

in daily

illustrat

years: 19

The varian

pronounce

From the simple system description it is apparent that in either case there are no overt inputs except in a control sense. The only source of variance and energy is temperature. We therefore consider the model as being both stochastic and temperature driven. A graph of the effective (daily HIGH-LOW average) is shown in Figure 1.3.

Terrestrial weather generally and temperature certainly is dominated by two cycles: the daily (period = 1 day) and the annual (period = 365 days). To be sure, there are other cycles too. Climatic changes occurring over centuries' time also display other lower frequencies. Even so, since the time horizons in this discussion are on the order of months or years, there is little need to consider these cases.

The earth's weather can also be considered stochastic. This is not the same as random since there is a definite pattern exhibited in conjunction with the afore-mentioned cycles. The obvious pattern is seasonal changes in mean temperature, but there are not so obvious patterns as well. For instance, there is also a significant seasonal variation in daily temperature ranges. These two effects are illustrated in Figures 1.3 and 1.4 for averages over 25 years: 1961 to 1985 for Lansing, Michigan.

The variance of the seasonal means and ranges is also quite pronounced and significant. This variance is reflected by

the standard deviation and is graphed for Lansing (1985) in Figures 1.5 and 1.6.

It would be distinctly advantageous to be able to use existing weather data as recorded by the National Weather Bureau. This data has been traditionally used by agronomists and ecologists in their modeling efforts for years and it is almost imperative that this source of data be sufficient in population studies.

The weather bureau keeps temperature records in primarily two formats. For certain stations, temperature is recorded at three hour intervals for eight data points per day. However, for the majority of stations, temperature readings are only kept for the daily highs and lows, without recording the corresponding times. Thus, there are effectively less than two data points recorded. This is a very real potential problem. It is well known by Nyquist's sampling theorem (Press, et al. 1986) that two complete points per [daily] cycle are necessary in order to recover an arbitrary function. Even so, it will be shown that this so-called HIGH-LOW data is also sufficient for our needs.

In both daily and annual temperature cycles there is evidenced a close to apparent sinusoidal periodicity. This is not truly sinusoid since atmospheric ambient response to the earth's rotation (diurnal) and revolution (annual) is

well k

quest

suffic

both d

functi

[1.9]

The te

define

stocha

the t

needs

quest

quest

confi

of th

variar

well known to exhibit a significant lag. Even so, the question begs to be asked: Can a sinusoidal model sufficiently describe the temperature profile? That is, for both diurnal and annual cycles, how well does the following function model the actual profile:

[1.9] $$T(t) = A + B \cdot \cos(2\pi t/p + \theta)$$

where A, B, θ are modelling parameters yet to be determined,

p = 1 for diurnal

p = 365 for annual

t (chronological time) in days.

The temperature might best be thought of as a random process defined by a basis profile. This can be implemented using a stochastic generator with an historical database to define the time slice distributions. Since a stochastic input needs to drive the system, answers to control and modelling questions will necessarily be statistical. Inferential questions can be asked and effectively answered by using confidence intervals. It should be possible to answer any of these question with only a knowledge of the means and variances of the output.

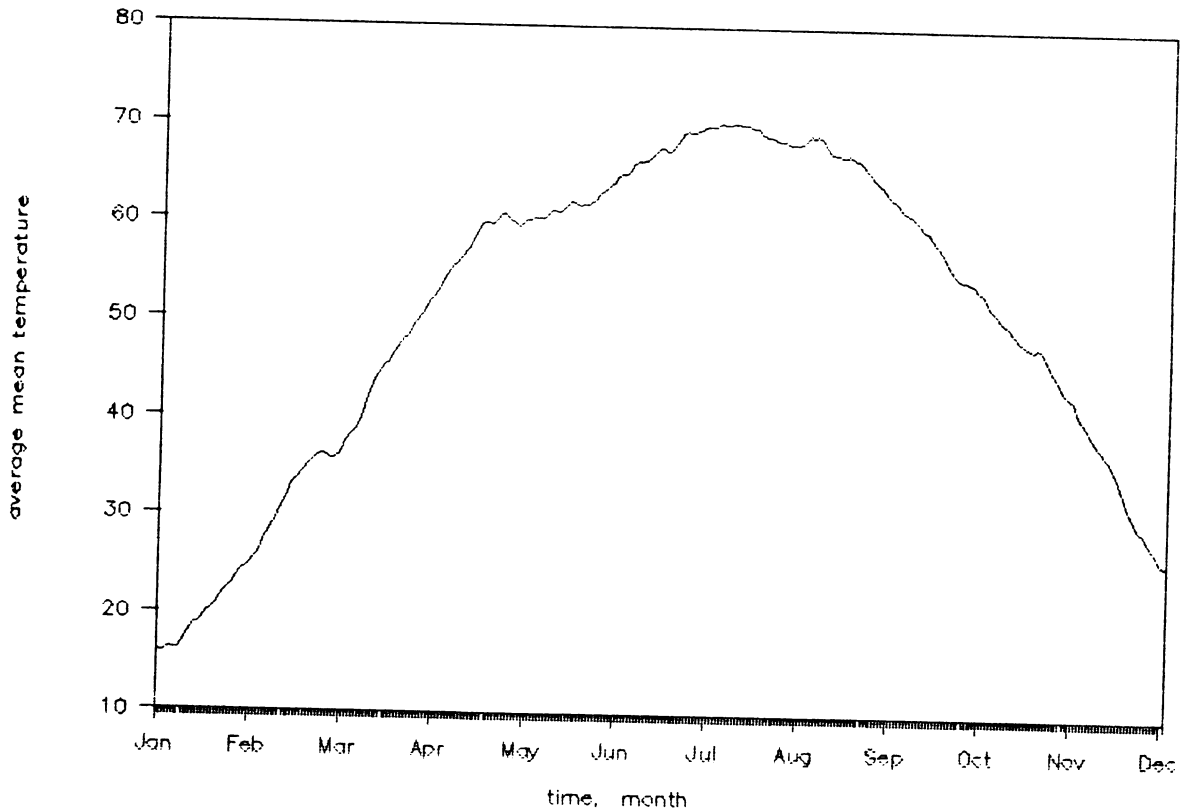


Figure 1.3 **The average mean temperature for Lansing, Michigan (1961 to 1985)**



Figure 1.4 **The average temperature range for Lansing, Michigan (1961 to 1985)**

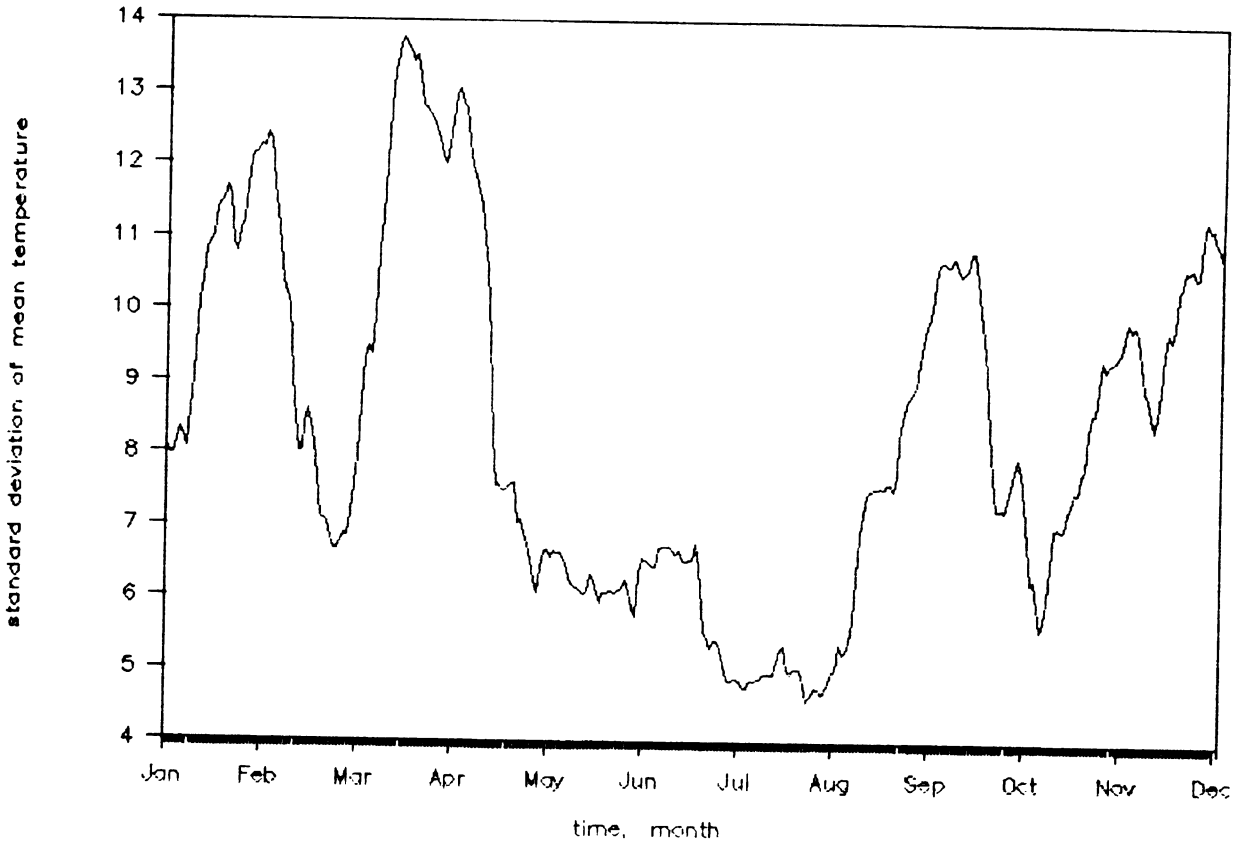


Figure 1.5 **The standard deviation of mean temperatures for Lansing, Michigan (1961 to 1985).**

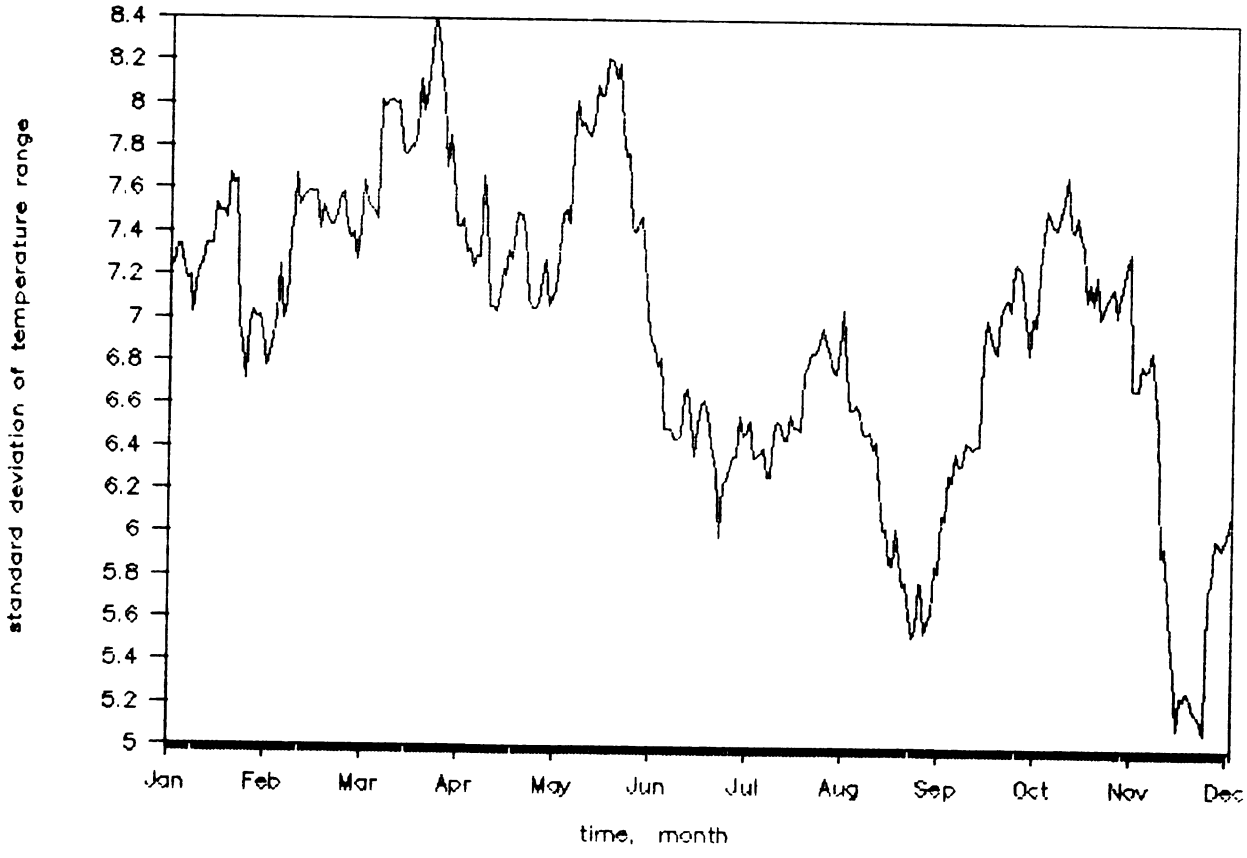


Figure 1.6 The standard deviation of temperature ranges for Lansing, Michigan (1961 to 1985).

THE S

A re
proj
adep
Even
mode

The
temp
crit
spec
For
both
rath
for
temp
usin
term

The
mech
Sinc
envi
valu
valu
deper

THE SYSTEM METHODOLOGY

A real design methodology is well outside the scope of this project. Presumably analysts and designers are sufficiently adept at their art and can perform this task per usual. Even so, there are several unique features of a heat-unit model that require special consideration.

The time horizon of a model also determines an effective temperature range for the model. This temperature range is critical since the biological activity function for each species will need to be accurate over the entire interval. For instance, if crop start-up time is an important entity, both death and dormancy threshold models are essential rather than a single threshold model. Presumably in order for any systematic treatment there will need to be temperature thresholds compiled for all species of interest using a variety of model Types (Types being a technical term indicating the number and character of thresholds).

The methodological question is that of establishing a mechanical procedure by which to compute these thresholds. Since the proposed model deals with a non-controlled environment, it will be necessary to acquire parametric values from nature's stochastic laboratory. Basically, values for the endogenous parameters (those that do not depend on the environment) must be determined using field,

rathe

to me

thre

popu

befor

Since

be.

tempe

requi

to ke

rather than laboratory data. The critical step is somehow to measure these activity function parameters (temperature thresholds and growth rates) by introducing a species' population to the system and obtaining endogenous data before it is contaminated by the environment.

Since this data is nature's, the weather history must also be. Therefore, the data must be historical and the temperature profile must be in a deterministic format. This requires another data reduction approach and it is important to keep these separate:

Deterministic profile: T(t) = effective daily temperature

Stochastic profile: M(t) = expected daily temperature

R(t) = daily temperature range

(t measured in days)

From each of profiles it is necessary to compute the physiological time and biological activity functions.

The execution phase of the model basically consists of a series of runs, each using a temperature sequence generated by the stochastic generator. It is the statistics gathered at the conclusion of this series of runs that constitute results.

THE RESEARCH OBJECTIVES

There have been a number of disclaimers up to this point on what the following research will not do. It is time to enunciate some of the things that will in fact be accomplished. The focus of this dissertation will be to accurately and systematically define and solve the significant modelling problems associated with biological systems invoking the heat-unit hypothesis. In addition to a significant degree of mathematical rigor, a very real attempt will be made to be as prescriptive as possible with suitable algorithms aimed toward the practitioner. The examples are intended to be purely illustrative and no attempt has been made to use "real" data to validate results. It is hoped that this study will provide a definitive approach to the modeling of heat-unit based biological systems.

i. Determination of the temperature profile

The heat-unit hypothesis presupposes the existence and knowledge of the temperature profile. This knowledge must be at two levels. First, one must be able to represent a temperature history, because only by using the actual history can one determine appropriate values of the system

parameters. This is not as trivial a problem as might be expected. Time is a continuous entity and temperature is defined at each point. However, in order that measurements be physically realizable, it is necessary that a sampling technique be employed. This is also more than simply a signal processing problem. For the vast history of weather data available through the National Weather Bureau to be useful, a data reduction technique must be derived that will be sufficiently accurate.

In order to maintain a maximum amount of data in a minimum amount of space, the National Weather Bureau traditionally keeps so-called HIGH-LOW data. That is to say, only the day's high and low temperatures are recorded and retained. Researchers have devised crude methods for recovering information from this data, but the justification is weak and results minimal. This research will justify and produce appropriate formulae for accurate data reduction.

In addition to reducing the essential temperature history, it must be possible to establish essential temperature statistics. Rather than simply deducing a temperature profile T(t),

recovery of the historical mean M(T) and range R(t) is sufficient to evolve very realistic temperature generators for the control problems outlined earlier.

ii. Parameter determination

Any mathematical model will employ system parameters in the form of constants. The values of these constants need to be determined. This identification problem is somewhat unusual in heat-unit models. There are really two types of parameters. Endogenous parameters as explained earlier, are those whose values are independent of the environment. Even when an accurate history profile is known, care must be taken so as to eliminate environmental effects when evaluating these parameters.

Once the endogenous parameters have been determined, there are yet **exogenous parameters.** These are simply parameters whose values do depend on the environment of other species populations within the model. Since this contrasts significantly with the endogenous variety, a new

set of techniques is appropriate here. Not surprisingly, the procedures derived for the computation of this set of parameters will involve both analytical and numerical techniques.

Methods currently employed in finding endogenous parameters are extremely rough. Most involve a crude graphical technique or simply guessing. In fact, the literature has a number of examples where the latter is unashamedly employed (Wang 1958, 1960)! This will lead to the methodological question of how to determine what constitutes endogenous data. Unfortunately, there is no fixed answer yet. This, too, will be a very useful topic for future study.

iii. Foundations and solutions

In all this work, models are expressible in the form of a system of first order differential equations (see equation [1.2]). Although this has a basic simplicity, complications arise quickly and it is necessary to apply numerical methods

practically from the outset. A study of different integration methods applied to the heat-unit problem is required in both a how-to-do-it fashion as well as error analysis. As was promised earlier, the emphasis will be on prescriptive approaches to general problems. This must be justified and based on the practitioner's needs.

iv. Stochastic modelling

In order to be useful, a system model must be more than simply a well defined, parameter determined set of differential equations. It must produce reasonable results in a variety of circumstances. As mentioned earlier, this requires the use of stochastically generated weather data. While the data need not be real, it must still accurately reflect the statistical properties of real data.

The (mathematical) device that performs this task is called a **stochastic data generator.** It generates a sequence of data that statistically

resembles a **basis** profile. These generators can be classified on the types of statistics that they keep. For instance, a simple but important property would be that the generated data have the same mean as that of the basis. Suitable generators will be described in detail.

While there will be no exhaustive work done on applications, illustrative examples on stochastic control problems will be applied showing that the approach is both reasonable and useful.

SUMMARY

The use of heat-unit models is population dynamics in not a new concept. However, it has been found that most of these models in practice have been ad hoc in nature and based upon tenuous theoretical underpinnings at best. This presents both aesthetic and pragmatic problems. A systematic approach is necessary to achieve further understanding.

This dissertation will not even come close to answering all questions. However, it is hoped that this research will provide a more definitive basis for describing and applying heat-unit models and their application in a stochastic environment. A genuine attempt will be made to be rigorous in the definitions and mathematics, but decidedly practical in the implementation of algorithms. Future analysts should be able to apply the results and procedures immediately with consistent and satisfactory results.

As in every dissertation, more questions, interesting problems and tempting paths surface. These will be acknowledged as future research possibilities, but the straight and narrow path will be followed to the goal of not only accurate, but valid modeling of population dynamic systems invoking the heat-unit hypothesis.

CHAPTER TWO

LITERATURE SURVEY

INTRODUCTION

The idea of heat units is not new. Much of the current practice is based on ideas that have been in existence over the past 250 years. Therefore, this literature review will be partially a historical survey, partially a look at current applications and a review of current theory.

Historically the terms heat-units and degree-days have been considered somewhat synonymous. However, using the terminology of this dissertation, **degree-days** are units of **physiological time** under the **heat-unit hypothesis.** The originators of the degree-day unit idea thought of these as time indices where certain events would, in principle, occur at the same [physiological] time during a species' existence. In other words, they were thought of as calendar events.

Thus, we note several of the traditional biases of these degree-day models. First, they were primarily agricultural and there were known start-up times that were annual. Time

horiz

say t

para

take

inde

grow

off

Thus

all

appr

anal

cons

The

pers

usi

fra

of

mod

ide

Thu

spr

etc

the

inf

and

horizon was taken to be the annual harvest. This is not to say that only crops were involved, but most animal life was parasitic or pest-like in nature. The degree-day unit was taken as an attempt to measure the species growth on an independent index. Most users tried to incorporate all growth effects on this single index, and then simply mark off events as a function of this new (generalized) index. Thus, an ideal index would be a composite with weights for all significant factors. This is to be contrasted with the approach taken in this research where the view is one of analysis, and where only the temperature component is considered.

The conceptual leap is important. Classical physical perspective saw only one [mechanical] time unit, but by using heat-units, there are now several independent time frames. However, these early researchers did not conceive of applying the degree-day index to system simulation models. The species was considered in isolation and the idea was simply to make a calendar of significant events. Thus, from this view one might find tabulations such as sprouting: 450 degree-days; flowering: 800 degree-days, etc. The idea that species interaction might well change the event calendar, and that population size, etc. would be influenced by competition was considered another problem and not part of the heat unit approach.

HIST

In t
temp
disc
foll
for
temp
plan
publ
ther
for

In
temp
esta
gene
(he
Bous
noti
stag
temp
esta

Othe
Sunl

HISTORICAL PERSPECTIVE

In the 1730's the Frenchman Rene' Reaumur (famous for the temperature scale used in the manufacture of cheese) discovered that the development of his plants very closely followed the cumulative mean temperatures in his locality for the months of April, May and June. By assuming the temperature expressed the amount of heat required for a plant to achieve each stage of maturity, he reasoned and published this temperature sum (Reaumur 1735) as "Reaumur's thermal constant of phenology". This provided for the basis for the heat-unit system used today.

In 1750 Adamson used Reaumur's idea, but only used temperatures in excess of the freezing point of water. This established the so-called "remainder index method", whereby generally only temperatures in excess of a given threshold (here: 0^{o} C) are considered in the accumulation. Boussingault (1837) published results using this threshold noting that the period from germination to any other growth stage was inversely proportional to the sum of the daily temperatures over that period. Edwards and Colin (1834) established similar results by using an upper threshold too.

Others realized that more was at work than simply heat. Sunlight, reflecting solar radiation, factors were added by

Humbol

Fritsc

[2.1]

In 1

diff

temp

forw

The

186

con

that

cum

conc

of

exce

Humbolt, Gasparin and Kabsch (1863). From these ideas, Fritsch (1861) published a popular thermal "constant":

$$[2.1] \qquad x = t + \frac{h \cdot t \cdot c}{12} - C$$

where
- x = thermal "constant"
- C = total heat from planting to germination
- t = total number of days
- c = sum of mean daily temperatures
- h = average number of sunlight hours over the period

In 1861 Fritsch determined the thermal constants of 889 different plant species. This was done using the Reaumur temperature scale and a threshold of 0^{o} C from January 1 forward.

The first significant analytical treatment was produced in 1869 by Linssar. Using some elegant mathematical contortions and a purely heat-unit argument, he concluded that plants achieve successive growth stages at the time the cumulative mean temperature is reached. He further concluded this was true using temperatures in linear excess of freezing, as well as those of the squares of these excesses.

The

pure

19th

the

popu

belo

[2.2

[2.

The idea of an actionametric constant (as opposed to a purely thermal constant) was developed further in the late 19th century and continued into the early 20th century. At the same time, the remainder index method also became more popular. Several of the more notable attempts are shown below:

i. Lehenhauser (1914) defined a "physiological summation index":

[2.2] $$\log_2(U) = \frac{T-40}{18} \qquad \text{for } T>40$$

where T = temperature (° F)

U = summation index

ii. Nuttason (1948) created a "photo-thermal-unit" (PTU) defined as follows:

[2.3] $$PTU = \sum_{i=D_p}^{D_m} [T - T_c] \cdot t_s \qquad \text{for } T>T_c$$

where D_p = planting date

D_m = maturity date

T = daily mean temperature

T_c = critical temperature

t_s = day length (hours)

[2.4]

[2.5]

iii. Lindsey, Newman (1956) attempted to measure the durations of different temperature levels by the so-called "effective diurnal spread" in units of "degree hours" (DH):

[2.4] $$DH = \frac{12 \cdot (h-t)^2}{h \cdot w}$$

where h = daily maximum temperature

w = daily minimum temperature

t = temperature threshold

iv. Holmes and Robertson (1959) introduced a means for computing "weekly degree-days"(D) as:

[2.5] $$D = N \cdot \left[(t-b) + \frac{L \cdot \sigma}{M} \right] \qquad \text{for } t>b$$

where D = degree-days accumulated for the month concerned

N = number of days in the month

M = mean number of days per month

The
of
exc
Upo
spe
cor
est
19t
deg

In
usi
a s
to
tem
a s
app
tem
new
leg

L = proportionality constant as tabulated by the author

t = monthly mean temperature

σ = standard deviation of the monthly mean temperature

b = temperature threshold

The first half of the 20th century also saw a large number of studies applying these method. Wang (1960) presents an excellent overview of the history of these pre-1955 studies. Upon review of this early work it is apparent that the species considered were viewed on a very high level (wheat, corn, etc.) and that the thresholds were only crudely estimated. Even so, this was a vast improvement over the 19th century where almost all thresholds were taken to be 0 degrees Celsius.

In a study by C. Y. Arnold (1954) which analyzed results using a linear temperature excess (recall equation 1.] found a significant upward bias in early estimates. He proceeded to give the first systematic method to determine the temperature threshold of a single threshold model. This is a statistical procedure in which linear regression is applied comparing growth results against threshold temperature. This study might be considered the start of a new era since the "degree-day" model was considered a legitimate measure in its own right (without photo-units,

precipita

that a

threshold

Because t

met with

existed a

below:

i.

ii.

iii

precipitation measures, etc.) and there was the recognition that a well defined procedure for establishing the threshold was necessary.

Because the heat-unit approach had a good intuitive feel, it met with considerable use and approval. Even so, there existed apparent shortcomings. Several of these are listed below:

i. Wang (1960) points out that almost any given species will invariably respond differently to the same environmental factor during different growth stages. It seemed that the temperature threshold constant was, in fact, a variable that changed over the various growth stages.

ii. Thornthwaite (1952) determined that any threshold between 32^{o} and 50^{o} F gave essentially identical results. This was confirmed by Wang (1957).

iii. Daubenmire (1947) noted that a peculiar problem existed at temperatures close to the threshold. Translating Daubenmire's jargon to that of this dissertation, the biological activity was expected to have a discontinuous derivative at the

threshold. However, Daubenmire found there to be a ill-defined scattered data transition. In fact, there was a real problem in applying the HIGH-LOW data as retained by the weather service. These problems were finally pinned down in a paper by Baskerville and Emin (1969) in which a data reduction technique was applied to HIGH-LOW data.

iv. Multiple researchers, as mentioned above, were dissatisfied that the myriad of growth factors other than temperature were not taken into account. These include soil characteristics, precipitation, light, etc.

Wang (1960) reflected the feelings of the traditionalists with his appeal for time varying temperature thresholds, a differentiation between growth and development and a new index including a complete set of environmental factors in the formula. This essentially was the state of heat-unit affairs in the middle 1960's. The degree-day was largely treated as an index to measure the growth stage of a plant throughout its [annual] life. Populations were taken to be almost exclusively homogeneous.

MODERN APPLICATIONS

The 1970's saw the introduction of inexpensive and convenient computer power. No longer was this machine the luxury of only physical scientists and engineers., but biologists as well. It was no surprise that the way was pointed toward ecological systems modeling. Rather than simply try to correlate a plant's growth with its temperature experience, these machines allowed relatively easy access to attempt small ecological systems models. In particular, animals as well as plants were modelled. Predator-Prey models, especially in agricultural applications were developed. In fact, agriculture remains the largest application by far of these systems models.

During the 1980's computers were not only accessible, but common place. The IBM PC, introduced in 1981, allowed virtually every student the chance at a personal modelling experience. Thus, there was a corresponding increase in population simulation models.

An extensive literature review was conducted on relevant publications for the years 1975 through July 1, 1987. Table 2.1 shows the list of journals surveyed in this review. All relevant papers were subjected to a survey as to year published and type of material (application, methodological, stochastic model or control application).

The results of this survey are summarized in Tables 2.2 and 2.3.

While this review is blatantly subjective and in no way should be taken statistically significant, the effect is clearly shown that there is a decided increase in interest in the heat-unit area. The year 1986 was a banner year! It should also be noted that the vast majority of publications are applications based. Certainly this is admirable and demonstrates that the simulation activity must in some measure be worthwhile, but very few of these applications constituted what this author would consider control problems. Most are still attempts at descriptions of the growth curves. Also, a very small percentage treated the problem as stochastic in nature. Unfortunately, the need for more theoretical direction is indicated with the almost insignificant figure for methodological considerations (mathematics, parameter estimation and data analysis).

Ta

Table 2.1: List of publications reviewed in study

ACTA Horticulture
Agricultural Forest Meteorology
Agronomy Journal
Annual Entomological Society of America
Canadian Entomology
Canadian Journal of Botany
Canadian Journal of Plant Science
Citrograph
Crop Physiology
Crop Science
Ecology
Entomological Society of Ontario
Environmental Entomology
Forestry Science
Hortoscience
Journal of America Society of Horticultural Scientists
Journal of Applied Ecology
Journal of Economic Entomology
Journal of Mathematical Biology
Journal of Nematology
Journal of Theoretical Biology
Mathematical Bioscience
Phytopathology
Southwest Entomology

Tabl

Ta

Table 2.2: Growth of Relevant Articles

year	% relevant articles	
1975-9	11	
1980	7	
1981	2	
1982	5	
1983	14	
1984	14	
1985	11	
1986	25	
1987*	11	(first six months)

Table 2.3: Classification of publication by application

type publication	percentage
Application	95
Stochastic	12
Methodology	12
Controls	7

The sub
over er
not to
author
Intere
Howeve

There
proble
conti
degre
commo
Dale
unit
and
be m
thes
espe
Gut
mod
mod
(th
rad

Oth
re
te

The subjective nature of this literature review cannot be over emphasized. No doubt many publications were missed, not to mention the journals themselves. Even-so, this author feels these qualitative results show the trend: Interest in the heat-unit approach is on the increase. However, too high a proportion is in the application area.

There are two primary directions to population dynamics problems applying the heat unit approach. Some researchers continue to search for ever more powerful indices based on degree-days and non-temperature factors. Probably the most common of these other factors are of solar units (Linvill, Dale and Hodges 1948), (Schreiber, et al. 1978) and photo units (Bohne, Rock and Stinner 1985), (Gutierrez, Williams and Kido 1985). As might be expected, these units tend to be most powerful when applied to leaf development. Of these, one particular study stands out as not only being especially successful, but also seems mathematically sound. Gutierrez (Gutierrez, Williams and Kido 1985) devised a model to measure dry matter production of grapes. In this model he used the traditional HIGH-LOW temperature data (threshold: 10^{o} C) as a time basis, but used solar radiation in effect as an independent time scale.

Other workers have taken the approach as espoused in this research . They have confined their attention to only the temperature based time scale. Still there are those who

look for time varying thresholds, but a fair number are going the multi-stage route outlined in chapter one of this dissertation. The usual approach to the multi-stage problem is in the use of a Leslie matrix mode. Hogg, in an excellent study (Hogg 1985) used a Leslie matrix to map out the growth stages of a potato leafhopper and derived life Tables using several different temperature regimes. This method led to moderate success.

While Hogg's model demonstrated no spacial component, others (Flinn, Taylor and Howes 1986) applied the Leslie method to simulate the leafhopper population but included immigration. His model had a time horizon of several growth cycles and met with what has been considered to be a rather successful result. However, as a cautionary note, his model also proved to be very sensitive to small changes in leafhopper survivorship.

It is not necessary to express these models in the inherently discrete and linear Leslie model. The more general compartmental model which was applied by Raworth (1984) in the study of aphids is a moderately successful attempt. The interested reader should note the excellent discussion in this Hogg's paper on variable development time.

More s
used t
each g
enviro
Bailey
of the
Partia
temper

There
logis
al. 1
toma
compu
stag
stag
stag

By f
were
the
app
dev
succ
as
as
few

More successful models using the compartmental model have used the Logistic equation to characterize growth within each growth stage. However the effect of non-temperature environmental factors has given superior results. Lasack, Bailey and Pedigo (1947) successfully simulated development of the Stalk Borer using a multistage logistic process. Partial life tables were derived and a good analysis of non-temperature environmental factors was provided.

There have also been experiments. In a four-stage (non-logistic) modal, Wolf, Rudich, Marani and Rekah (Wolf, et al. 1986) developed a simulation for the growth stages of a tomato. He assumed linear temperature excesses in the computation of physiological time for the first and second stages, but quadratic excesses in the third and fourth stage. This was compared against linear excesses in all stages and was found superior.

By far and away (in this author's opinion) the best results were achieved by those experimenters who took advantage of the stochastic nature of their models. A very traditional approach by Lund and Herne (1980) to model red mite development over a multi-year time horizon was quite successful. Even though they treated the physiological time as an index rather than a time, their output was expressed as resulting from a random process. This is one of the very few such studies found by this author.

Not all random process experiments need be based only on temperature. Waggoner (1984) took egg hatching time to be a Poisson process and expressed this as a stochastic driver. It is interesting to note that he also defined the mean daily temperature somewhat differently in order to negate a statistical bias in the data. These two added "extras" were enough to show a significant improvement over the straight degree-day model.

Probably one of the most consistent users of the multistage logistic model has been Kemp. In three publications (Dennis, Kemp and Beckwith 1986), (Kemp, Dennis and Beckwith 1986) and (Kemp, Onsager 1986) he has very clearly demonstrated the approach advocated in this dissertation. In the later paper he illustrates the principle of a stochastic generator based on a historical database. He also gives an explicit procedure (based on the maximum likelihood estimator) to determine the [biological activity] parameters. These studies are excellent methodological examples.

In addition to the modeling problem per se is the closely related identification problem. Many workers have simply guessed at reasonable values. Others have attempted an estimate by reviewing the data in the analysis of growth curves. In a stochastic model of balsam fir, Osawa, Shoemaker and Stedinger (1983) show this so-called maximum

likeliho

Boisvert

both tem

assumed

variatio

The re

exhaust

results

variety

In clo

been a

too.

posit

and (M

desi

appli

likelihood method for parameter estimation. Castonguay, Boisvert and Dube (1984) illustrate techniques to estimate both temperature threshold and the effective start date (if assumed unknown). These are called the coefficient of variation and standard error of prediction methods.

The references cited in this section are not at all exhaustive or even necessarily the most effective model results. The list was simply meant to illustrate the wide variety of approaches to this problem.

In closing, it should be noted that the heat unit model has been also applied on a climatic scale (Shugart, et al. 1980) too. Also, not all researchers have assumed strictly positive physiological time (Coakley, Boyd and Line 1982) and (Mikkelsen 1981). Several have tackled the experimental design problem (Eskridge and Stevens 1987). Each application is unique.

THEORET

The pr
models
suffic
structu
the pa
distin
are th
paid t

There
only a
as we
the c
funct
U(a,x
follo

[2.6]

THEORETICAL AND MODELING QUESTIONS

The primary characteristic of agricultural population models is the homogeneity of the population. It is not sufficient to apply a simple logistic or Lotka-Volterra structure to the problem because each of these tend to treat the population en mass rather than as a sequence of distinct stages. Since the primary examples of these models are the classical, it is worth while that acknowledgement is paid the work done on more precise models.

There have been several significant means devised that not only accommodate the population homogeneity, but exploit it as well. McKendrick (1926) developed a generalization on the classical by defining the population density, now a function of both age a and physiological time τ as $U(a,\tau)$. The fundamental McKendrick relationship is as follows:

$$[2.6] \qquad [D_a + D_\tau]U(a,\tau) = -\mu(a,\tau)\cdot U(a,\tau)$$

$$\text{where} \qquad U(0,\tau) = \int_0^\infty \beta(a,\tau)\cdot\delta(a,\tau)\, da$$

$$U(a,0) = U_0(a)$$

A thorou

by Hoppe

In prac

numerica

Leslie m

nice ap

Chapter

part, t

reflect

This me

also ha

of many

A drawb

that al

not as

mention

to thi

stocha

with t

$\beta(a,\tau)$ = birth rate

$\delta(a,\tau)$ = death rate

$U_0(a)$ = initial age distribution

A thorough discussion of properties of this system is given by Hoppensteadt (1975).

In practice, of course, equation [2.6] must be solved numerically. Probably the most straight-forward is the Leslie model. Leslie (1945) proposed this model as a very nice application of matrices to the modelling problem. Chapter nine describes this in more detail, but for the most part, the Leslie matrix describes each growth stage as the reflection of a linear combination of predecessor stages. This method not only is a good approximation to [2.6] but it also has an excellent intuitive appeal as it is the source of many pedagogical examples in use today.

A drawback to the Leslie model is the implicit assumption that all individuals mature at the same rate. While this is not as serious a problem as many other pragmatic oversights mentioned earlier, it is simply not realistic. The solution to this problem is to model the delay between stages stochastic. This was done by Manetsch (1966, 1976, 1980) with the so-called distributed delay technique. This

approach also compartmentalizes the growth stages, but at the same time also forces a stochastic element into maturity rates. Van Sickle (1977) applied this distributed delay model to the physiological time.

Taking a cue from the distributed delay model, the Leslie model was generalized to represent populations whose individuals mature at different rates. This was first proposed by Slobodkin (1953) and applied quite successfully by Werner and Coswell (1977). Each of these models represents a variance in the maturation rate of a cohort of individuals describable in terms of a probability distribution. An excellent review of these distributed age structure models can be read in Plant and Wilson (1986).

The major thrust of this research has not been to describe the exogenous model, as in the McKendrick, Manetsch and Leslie approaches, but to consider the endogenous model. That is to say, the parameters that are used to describe the system and its interaction with the environment are the "usual" model descriptors. This dissertation will consider those parameters that describe the chronological-physiological times. These entities constitute the essence of this work.

Unfortunately, this author has found no systematic research in this area. On the question of treating the system input as a stochastically temperature driven, there is also a dearth of activity. Indeed, the temperature response problem, the systematic determination of endogenous model types and parameters of that model are problems that have been , if at all, handled on an ad hoc basis. Thus, it is the intent of this research to at least initiate this action.

CHAPTER THREE

THE COMPUTATION OF HEAT UNITS IN BIOLOGICAL SYSTEMS

INTRODUCTION

Classical biological models relate the state of an appropriate biological metric (mass, area, number of organisms, etc.) to time. Heat unit models also require an accurate knowledge of temperature. Typically these models describe this metric as a function of temperature and time. Clearly, this mandates an understanding of the time - temperature relationship.

This [time - temperature] relationship is called the temperature profile. While this function T(t) can be easily understood in principle, in practice it requires data which may not be sufficiently precise. In a study involving micro-organisms, this profile will likely require accuracy at the minute or even second level. In this case, a real-time temperature measuring device would likely be in order. At the other extreme, models using a geological time scale would require only course climatological estimates. Not surprisingly, a large proportion of biological models

are on a more human time scale. For instance, the world's agricultural system is based primarily on an annual cycle.

The time scale is important because in practice the temperature is sampled. There is no truly continuous measurement. In the instance of moderate (say annual) models this sampling rate is typically hourly or daily.

The National Weather Bureau records the temperature at three hour intervals for a number of stations (typically at airports). This EIGHT-POINT data is hardly a continuous measurement. However, it clearly shows the daily temperature cycle and can provide relatively good statistics. For the data shown in Table 3.1, the sample mean μ' and variance σ' are as defined below.

$$\mu' = \frac{1}{8} \cdot \sum_{k=0}^{7} T(t_k)$$

[3.1]

$$(\sigma')^2 = \frac{1}{8} \cdot \sum_{k=0}^{7} [T(t_k)-\mu']^2$$

The EIGHT-POINT data actually retains 16 values per day; the temperatures and the times at which they occur. This can lead to inefficient data storage, but for a model requiring high predictive power, it is necessary to retain this precision.

On the other hand, field researchers traditionally simply record the daily high and low temperatures. Times of occurrence are not kept. Thus, this HIGH-LOW data retains two values per day. Statistics gained by this method are the sample mean M and range R as defined in equation 3.2.

[3.2]
$$M = \tfrac{1}{2} \cdot (T_{max} + T_{min})$$
$$R = T_{max} - T_{min}$$

The EIGHT-POINT data require eight times the storage of the HIGH-LOW data. However, the loss of precision using the HIGH-LOW data and the [supposedly] incomparable statistics are a potential problem.

THE SINUSOID MODEL

Daily fluxuations in temperature are basically oscillatory. As rough approximation, we consider them to be sinusoidal; written as

[3.3] $$T(t) = A \cdot \cos(2\pi t + \theta) + B$$

where $T(t)$ is the temperature over t days, $0 \le t \le 1$.

A, B, θ are parameters yet to be determined.

It might be desirable to relax the single day restriction. However in this case, the parameters A, B and θ will slowly vary with time.

It is the goal of this chapter to apply this [continuous] model to the HIGH-LOW data. It will be shown that the statistics derived will effectively recover the information contained in the EIGHT-POINT data. This will be true not only for statistics but for the heat-unit measurement of physiological time in degree-days.

STATEMENT OF THE PROBLEM

This chapter will show the following:

i. Reliable statistics can in fact be inferred using HIGH-LOW data.

ii. A reliable heat-unit measure can also be inferred using HIGH-LOW data.

iii. Traditional interpretations of heat-unit measures using HIGH-LOW data are not sufficiently valid compared to EIGHT-POINT data.

iv. A new measure for heat-units using HIGH-LOW data will be derived which will be effective for all temperatures. This is to be contrasted with traditional measures which are effective only at temperatures not in the neighborhood of the threshold. Again, this measure will be compared against that of the EIGHT-POINT data.

Table 3.1 EIGHT-POINT temperature record for a single daily cycle.

time	t_0	t_1	t_2	t_3	t_4	t_5	t_6	t_7
clock time	1:00	4:00	7:00	10:00	13:00	16:00	19:00	22:00
temperature	$T(t_0)$	$T(t_1)$	$T(t_2)$	$T(t_3)$	$T(t_4)$	$T(t_5)$	$T(t_6)$	$T(t_7)$

Table 3.2 HIGH-LOW temperature record for a single daily cycle.

	High	Low
temperature	T_{max}	T_{min}

APPLYING HIGH-LOW DATA TO THE SINUSOID MODEL

In order to calculate the unknown parameters we use the maximum and minimum functional values T_{max}, T_{min}. Since we do not know the corresponding times, we assume the values simply occur at the sinusoid extremes, 1/2 day apart as illustrated in Figure 3.1.

It follows that...

[3.4]
$$A \cdot \cos(2\pi t_{max} + \theta) + B = T_{max}$$
$$A \cdot \cos(2\pi t_{min} + \theta) + B = T_{min}$$

$$\text{where} \quad t_{max} = -\frac{\theta}{2\pi}$$

$$t_{min} = \frac{\pi - \theta}{2\pi}$$

Constants A and B are easily evaluated. Also, by incorporating equations [3.2],

[3.5]
$$A = \tfrac{1}{2}(T_{max} - T_{min}) = \tfrac{1}{2}R$$
$$B = \tfrac{1}{2}(T_{max} + T_{min}) = M$$

The phase angle cannot be evaluated (two equations and three unknowns). Even so, it is not necessary for the results that follow.

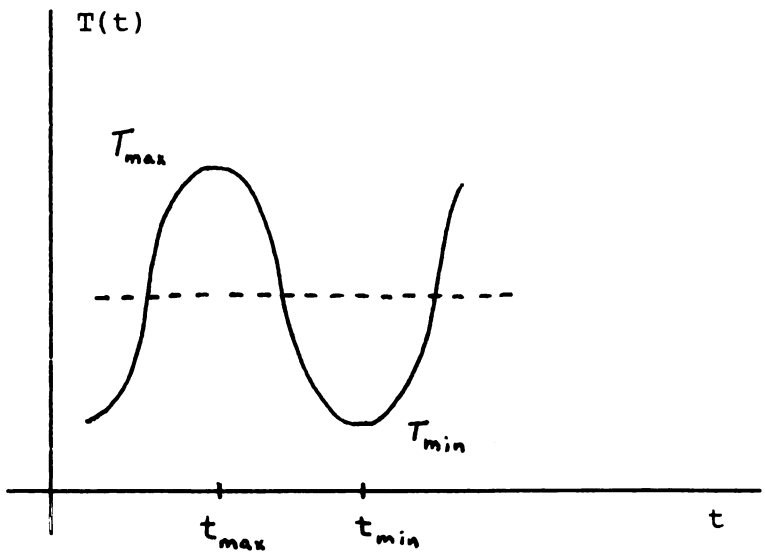


Figure 3.1 An idealized daily temperature profile.

THE STATISTICAL INTERPRETATION

Recall that the mean of a function T(t) defined on an interval [a,b] is given by

$$[3.6] \qquad \mu = \frac{1}{b-a} \cdot \int_a^b T(t)dt$$

while the variance is

$$[3.7] \qquad \sigma^2 = \frac{1}{b-a} \cdot \int_b^a [T(t)-\mu]^2 dt$$

where we take the density function to be uniform so as to weight each temperature measurement equally. In this application,

$$\mu = \int_0^1 [A \cdot \cos(2\pi t+\theta)+B]dt$$

$$[3.8] \qquad = \left[\frac{A}{2\pi} \cdot \sin(2\pi t+\theta)+B \cdot t \right]_0^1 = B$$

$$\sigma^2 = \int_0^1 [A\cdot\cos(2\pi t+\Theta)+B-B]^2 dt$$

[3.9]

$$= \left[\tfrac{1}{2}A^2 + \tfrac{1}{4}A^2\cdot\cos(4\pi t 2\Theta) \right]_0^1 = \tfrac{1}{2}A^2$$

Applying equations [3.5] yields...

$$\mu = \tfrac{1}{2}(T_{max}+T_{min}) = M$$

[3.10]

$$\sigma = (T_{max}-T_{min})/\sqrt{8} = R/\sqrt{8}$$

Thus, by using the continuous SINUSOID model, the HIGH-LOW data should reproduce the standard distribution statistics.

In order to verify the legitimacy of these results, we consider the EIGHT-POINT temperature data taken from the National Weather Bureau for Lansing, Michigan during the year 1985. See Tables 3.3 and 3.4. These Tables exhibit both daily EIGHT-POINT data (Table 3.3) and HIGH-LOW data (Table 3.4) for the first three months of 1985.

Using equations [3.1], means μ' and standard deviations σ' are computed using the EIGHT-POINT data. Equations [3.10] are used with the HIGH-LOW data to give corresponding

HIGH-LOW statistics (μ,σ) for the SINUSOID model. Results for each method are summarized in Tables 3.5 and 3.6. Table 3.5 relates results for the mean (μ and μ') while Table 3.6 relates the standard deviations (σ and σ'). In order to compare these two methods, ratios (HIGH-LOW/EIGHT-PT) are computed for both mean and standard deviations. Using the null hypothesis:

H_0: ratio = 1 (ie. both methods are identical.),

standard errors for these ratios are computed monthly below.

January...	Mean:	0.049
	Standard Deviation:	0.124
February..	Mean:	0.076
	Standard Deviation:	0.048
March.....	Mean:	0.030
	Standard Deviation:	0.062

These standard errors are extremely small. The alternative hypothesis, that the statistics are significantly different, is to be quickly rejected. There is no significant difference between the EIGHT-POINT statistics and the HIGH-LOW (assuming the SINUSOID model) statistics.

Table 3.3: EIGHT-POINT temperature data, °F
Lansing, Michigan (1985).

	January								February								March							
	01	04	07	10	13	16	19	22	01	04	07	10	13	16	19	22	01	04	07	10	13	16	19	22
1	31	32	31	29	26	23	19	22	3	2	2	5	8	8	2	-8	34	32	28	36	37	37	30	26
2	25	25	23	22	23	23	17	12	-1	-2	1	4	7	9	5	-7	24	24	25	26	34	32	28	27
3	11	10	8	14	21	24	13	9	-9	-6	-1	4	13	18	13	8	23	21	26	31	33	47	44	33
4	9	8	10	19	29	28	27	28	6	2	-6	0	14	17	8	6	30	27	25	24	27	24	22	19
5	30	30	26	24	27	30	29	30	10	12	13	14	18	19	16	16	15	12	12	21	28	30	28	25
6	30	30	28	31	34	34	31	31	10	14	15	16	20	20	17	18	23	24	26	27	32	35	37	37
7	31	31	31	31	31	31	30	29	15	10	10	15	17	17	11	0	38	39	38	37	43	44	39	35
8	22	15	9	9	12	15	13	12	-9	-9	-9	3	13	16	8	7	31	29	26	36	41	41	36	34
9	9	6	9	11	18	19	13	13	-4	-7	-9	4	18	20	9	9	34	31	30	44	55	59	53	51
10	15	15	16	17	20	20	18	18	0	13	18	20	21	21	20	21	48	47	45	42	45	43	43	45
11	17	13	8	10	18	20	20	20	21	22	23	25	30	33	34	32	41	31	34	34	43	43	38	31
12	16	21	20	22	27	27	22	20	33	32	32	29	28	26	23	22	28	26	27	33	37	36	35	34
13	15	14	15	17	21	24	21	26	21	21	20	20	23	24	21	19	34	34	32	35	40	42	37	28
14	28	30	27	28	27	23	19	16	20	18	17	18	20	19	18	17	27	25	22	29	35	39	33	28
15	15	13	10	13	15	16	14	13	17	10	13	16	20	21	15	13	30	30	32	41	50	40	36	34
16	6	9	14	18	20	20	18	17	9	6	7	12	20	25	26	24	30	27	25	29	29	31	29	24
17	17	18	21	23	24	23	22	22	24	26	23	29	32	29	25	23	19	17	15	25	33	37	35	32
18	14	13	15	17	19	20	17	21	22	20	22	28	36	34	34	32	30	30	33	40	48	57	47	37
19	18	16	14	10	10	6	2	-1	22	18	16	20	24	28	25	25	31	30	29	38	42	43	34	28
20	-2	-3	-5	-4	0	0	-9	-9	22	19	17	24	34	37	34	30	24	20	19	32	40	44	39	29
21	-9	-6	1	10	11	15	18	25	29	28	32	38	43	43	40	39	26	23	21	39	49	51	47	42
22	23	22	20	21	24	23	20	23	40	39	40	40	41	41	41	41	37	35	34	35	39	37	37	36
23	22	22	23	25	24	23	23	23	39	41	39	46	46	46	43	39	36	36	36	36	36	35	34	33
24	24	24	22	23	26	26	26	24	37	34	32	33	34	34	34	34	32	29	26	35	42	44	40	31
25	21	24	23	22	21	21	14	14	33	33	33	35	38	38	37	35	29	29	29	44	55	59	55	50
26	14	13	14	16	19	16	13	10	35	34	34	37	44	45	38	34	51	53	52	57	51	56	57	51
27	12	15	18	20	24	25	27	27	32	28	21	21	24	25	21	20	48	55	50	57	43	42	42	43
28	24	21	19	18	22	22	21	20	19	21	25	30	39	43	38	35	46	46	44	42	51	55	50	39
29	19	16	18	20	20	20	19	19	35	35	36	39	43	38	36	36	36	33	32	32	34	34	33	31
30	19	19	15	19	24	27	24	22									33	33	32	33	36	49	39	33
31	21	20	20	16	23	23	17	9									31	30	30	31	34	32	30	29

Table 3.4: HIGH-LOW Temperature data, °F
Lansing, Michigan (1985).

Date	January		February		March	
	High	Low	High	Low	High	Low
1	32	19	8	-8	37	26
2	25	12	9	-7	34	24
3	24	8	18	-9	47	21
4	29	8	17	-6	30	19
5	30	24	19	10	30	12
6	34	28	20	10	37	23
7	31	29	17	0	44	35
8	22	9	16	-9	41	26
9	19	6	20	-9	59	30
10	20	15	21	0	48	42
11	20	8	34	21	43	31
12	27	16	33	22	37	26
13	26	14	24	19	42	28
14	30	16	20	17	39	22
15	16	10	21	10	50	30
16	20	6	26	6	31	24
17	24	17	32	23	37	15
18	21	13	36	20	57	30
19	18	-1	28	16	43	28
20	0	-9	37	17	44	19
21	25	-9	43	28	51	21
22	24	20	41	39	39	34
23	25	22	46	39	36	33
24	26	22	37	32	44	26
25	24	14	38	33	59	29
26	19	10	45	34	57	51
27	27	12	32	20	57	42
28	24	18	43	19	55	39
29	20	16	43	35	36	31
30	27	15			49	32
31	23	9			34	29

Table 3.5: MEAN Calculations using EIGHT-POINT data, HIGH-LOW data (with the SINUSOID model), °F Lansing, Michigan (1985).

	EIGHT POINT	HIGH LOW	Ratio	EIGHT POINT	HIGH LOW	Ratio	EIGHT POINT	HIGH LOW	Ratio
1	26.63	25.50	0.958	2.75	0.00	0.000	32.50	31.50	0.969
2	21.25	18.50	0.871	2.00	1.00	0.500	27.50	29.00	1.055
3	13.75	16.00	1.164	5.00	4.50	0.900	32.25	34.00	1.054
4	19.75	18.50	0.937	5.88	5.50	0.935	24.75	24.50	0.990
5	28.25	27.00	0.956	14.75	14.50	0.983	21.38	21.00	0.982
6	31.13	31.00	0.996	16.25	15.00	0.923	30.13	30.00	0.996
7	30.63	30.00	0.979	11.88	8.50	0.715	39.13	39.50	1.009
8	13.38	15.50	1.158	2.50	3.50	1.400	34.25	33.50	0.978
9	12.25	12.50	1.020	5.00	5.50	1.100	44.63	44.50	0.997
10	17.38	17.50	1.007	16.75	10.50	0.627	44.75	45.00	1.006
11	15.75	14.00	0.889	27.50	27.50	1.000	36.88	37.00	1.003
12	21.88	21.50	0.983	28.13	27.50	0.978	32.00	31.50	0.984
13	19.13	20.00	1.045	21.13	21.50	1.018	35.25	35.00	0.993
14	24.75	23.00	0.929	18.38	18.50	1.007	29.75	30.50	1.025
15	13.63	13.00	0.954	15.63	15.50	0.992	36.63	40.00	1.092
16	15.25	13.00	0.852	16.13	16.00	0.992	28.00	27.50	0.982
17	21.25	20.50	0.965	26.38	27.50	1.042	26.63	26.00	0.976
18	17.00	17.00	1.000	28.50	28.00	0.982	40.25	43.50	1.081
19	9.38	8.50	0.906	22.25	22.00	0.989	34.38	35.50	1.033
20	-4.00	-4.50	1.125	27.13	27.00	0.995	30.88	31.50	1.020
21	8.13	8.00	0.984	36.50	35.50	0.973	37.25	36.00	0.966
22	22.00	22.00	1.000	40.38	40.00	0.991	36.25	36.50	1.007
23	23.13	23.50	1.016	42.38	42.50	1.003	35.25	34.50	0.979
24	24.38	24.00	0.984	34.00	34.50	1.015	34.88	35.00	1.003
25	20.00	19.00	0.950	35.25	35.50	1.007	43.75	44.00	1.006
26	14.38	14.50	1.008	37.63	39.50	1.050	53.50	54.00	1.009
27	21.00	19.50	0.929	24.00	26.00	1.083	47.50	49.50	1.042
28	20.88	21.00	1.006	31.25	31.00	0.992	46.63	47.00	1.008
29	18.88	18.00	0.953	37.25	39.00	1.047	33.13	33.50	1.011
30	21.13	21.00	0.994				36.00	40.50	1.125
31	18.63	16.00	0.859				30.88	31.50	1.020
Average			0.980			0.939			1.013

Table 3.6: STANDARD DEVIATION Calculations using EIGHT-POINT data HIGH-LOW data (with the SINUSOID model), °F Lansing, Michigan (1985).

	EIGHT POINT	HIGH LOW	Ratio	EIGHT POINT	HIGH LOW	Ratio	EIGHT POINT	HIGH LOW	Ratio
1	4.55	4.60	1.011	4.71	5.66	1.202	3.94	3.89	0.987
2	4.21	4.60	1.093	4.92	5.66	1.150	3.46	3.54	1.023
3	5.43	5.66	1.042	9.08	9.55	1.052	8.73	9.19	1.053
4	8.83	7.42	0.840	6.94	8.13	1.171	3.15	3.89	1.235
5	2.17	2.12	0.977	2.86	3.18	1.112	7.00	6.36	0.909
6	1.90	2.12	1.116	3.11	3.54	1.138	5.44	4.95	0.910
7	0.70	0.71	1.014	5.25	6.01	1.145	2.80	3.18	1.136
8	3.90	4.60	1.179	9.62	8.84	0.919	5.04	5.30	1.052
9	4.21	4.60	1.093	10.30	10.25	0.995	10.83	10.25	0.946
10	1.87	1.77	0.947	6.81	7.42	1.090	1.92	2.12	1.104
11	4.49	4.24	0.944	4.97	4.60	0.926	4.73	4.24	0.896
12	3.44	3.89	1.131	3.92	3.89	0.992	4.06	3.89	0.958
13	4.23	4.24	1.002	1.54	1.77	1.149	4.15	4.95	1.193
14	4.63	4.95	1.069	1.11	1.06	0.955	5.21	6.01	1.154
15	1.73	2.12	1.225	3.46	3.89	1.124	6.38	7.07	1.108
16	4.87	4.95	1.016	7.96	7.07	0.888	2.29	2.47	1.079
17	2.33	2.47	1.060	3.08	3.18	1.032	8.18	7.78	0.951
18	2.69	2.83	1.052	5.98	5.66	0.946	9.08	9.55	1.052
19	6.26	6.72	1.073	3.77	4.24	1.125	5.54	5.30	0.957
20	3.32	3.18	0.958	7.11	7.07	0.994	8.89	8.84	0.994
21	11.10	12.02	1.083	5.63	5.30	0.941	11.41	10.61	0.930
22	1.41	1.41	1.000	0.70	0.71	1.014	1.48	1.77	1.196
23	0.93	1.06	1.140	3.08	2.47	0.802	1.09	1.06	0.972
24	1.41	1.41	1.000	1.32	1.77	1.341	6.09	6.36	1.044
25	3.61	3.54	0.981	2.05	1.77	0.863	12.13	10.61	0.875
26	2.50	3.18	1.272	4.21	3.89	0.924	2.55	2.12	0.831
27	5.29	5.30	1.002	3.94	4.24	1.076	5.63	5.30	0.941
28	1.76	2.12	1.205	8.32	8.49	1.020	4.85	5.66	1.167
29	1.27	1.41	1.110	2.54	2.83	1.114	1.45	1.77	1.221
30	3.59	4.24	1.181				5.36	6.01	1.121
31	4.33	4.95	1.143				1.45	1.77	1.221
Average			1.063			1.041			1.039

THE NATURE OF THE HEAT UNIT MODEL

The justification for the heat unit model is based on the observation that many species essentially lie dormant while the temperature is below a certain critical temperature T_0. Above this threshold the biological activity is roughly proportional to the effective temperature (i.e. the temperature in excess of T_0). Thus, the biological activity is a function $\beta(T)$ of temperature T. This is illustrated in Figure 3.2.

$$[3.11] \qquad \beta(T) = \begin{cases} 0, & T(t) \leq T_0 \\ r \cdot [T(t) - T_0], & T(t) > T_0 \end{cases}$$

$$= r \cdot \begin{cases} 0, & T(t) \leq T_0 \\ T(t) - T_0, & T(t) > T_0 \end{cases}$$

The biological activity has two species explicit parameters, r and T_0. Indeed, the initial identification problem (to be discussed in chapter seven), is simply evaluating these two [endogenous] parameters.

Suppose the temperature profile $T(t)$ is a continuous function of chronological time t. We define the

physiological time $\tau(t)$ to be the cumulative heat units in excess of the species specific temperature threshold T_0. This is illustrated in Figure 3.3 where only the cumulative temperature above $T(t)=T_0$ contributes to the physiological time. Geometrically the physiological time is represented by the area above the threshold T_0. Therefore it can be evaluated as the following integral:

$$[3.12] \qquad \tau = \int_{t_0}^{t_1} [T(t)-T_0]dt + \int_{t_2}^{t_3} [T(t)-T_0]dt$$

In addition to the threshold T_0, τ is a function of the initial and final times t_0, t_3. For an arbitrary time interval $[t_0,t]$ this is simply

$$[3.13] \qquad \tau = \int_{t_0}^{t} \begin{bmatrix} 0, & T(t)\leq T_0 \\ T(\mu)-T_0, & T(t)>T_0 \end{bmatrix} d\mu$$

This can also be written as

$$[3.14] \qquad \frac{d}{dt}\tau(t) = \begin{cases} 0, & T(t)\leq T_0 \\ T(\mu)-T_0, & T(t)>T_0 \end{cases}$$

$$\equiv \frac{1}{r} \cdot \beta(T(t))$$

Thus, the fundamental relationship between physiological time $T(t)$ and biological activity $\beta(T)$ is:

$$[3.15] \qquad \beta(T(t)) = r \cdot \frac{d}{dt} \tau(t)$$

While typical units for chronological time are days, those for heat units are degree-days.

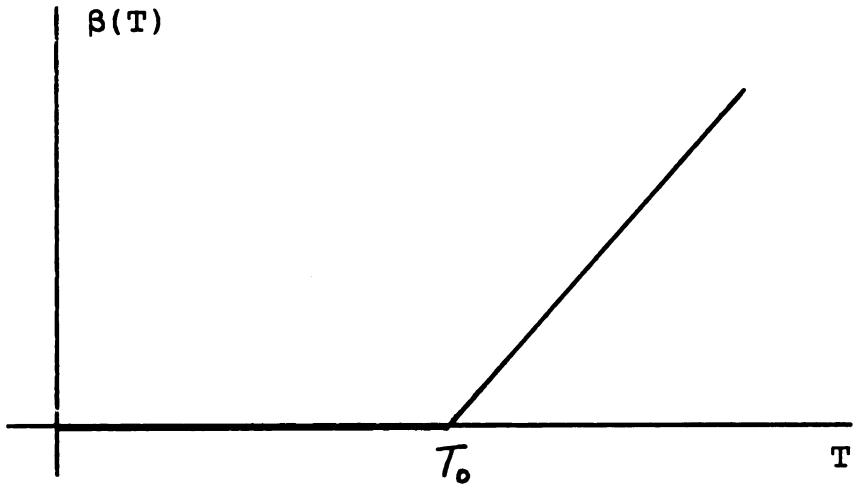


Figure 3.2 The Biological Activity β(T).

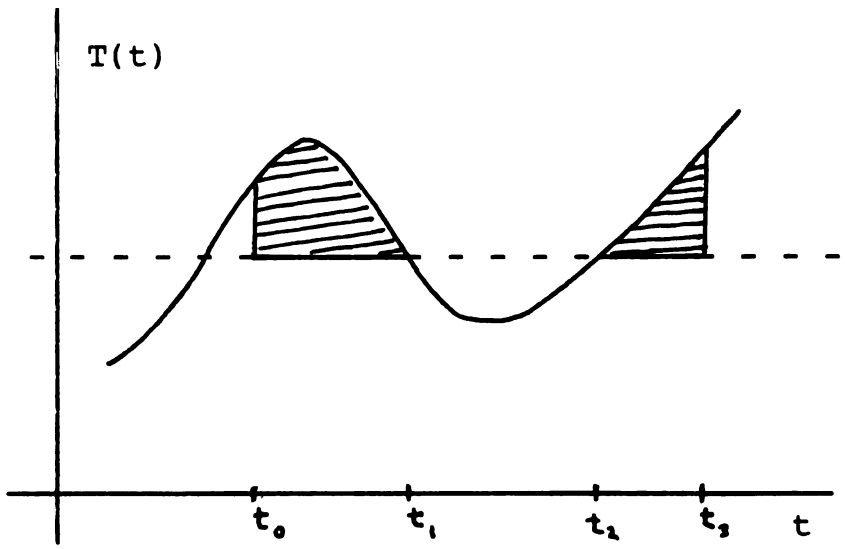


Figure 3.3 Relationship of physiological time to the temperature profile T(t).

THE CASE OF DISCRETE DATA

As a practical matter, data is available in only a discrete format, usually at fixed chronological time intervals:

$$[3.16] \qquad t_k = t_0 + k\cdot\delta, \qquad k=0,1,\ldots,n-1$$

The fixed step size is δ, which represents the sampling interval.

If we assume the temperature profile to be a sufficiently smooth function, a simple rectangular approximation to this integral is quite reasonable. Even though higher order integration techniques give a more refined accuracy. this will be shown to be quite sufficient. This will be discussed further in chapter four.

Thus, the discrete approximation is

$$[3.17] \qquad \tau \approx \delta\cdot\sum_{k=0}^{n-1} \begin{cases} 0, & T(t_k)\le T_0 \\ T(t_k)-T_0, & T(t_k)>T_0 \end{cases}$$

We consider the two particular cases of interest for daily heat unit measurements.

i. EIGHT-POINT data.

$$\delta = 1/8 \quad \text{day}$$

[3.18] $$t_k = (1+3t)/24 \quad \text{days}, \qquad k=0,1,\ldots,7$$

[3.19] $$\tau = \frac{1}{8} \cdot \sum_{k=0}^{7} \max \{T(t_k)-T_0, 0\}$$

ii. HIGH-LOW data.

$$\delta = 1 \quad \text{day}$$

$$t_k = \tfrac{1}{2} \quad \text{day}$$

[3.20] $$\tau = \max \{T(\tfrac{1}{2})-T_0, 0\}$$

$$= \begin{cases} 0, & M \le T_0 \\ M - T_0, & M > T_0 \end{cases}$$

$$= \begin{cases} 0, & T_{max}+T_{min} \le 2 \cdot T_0 \\ \tfrac{1}{2}(T_{max}+T_{min})-T_0, & T_{max}+T_{min} > 2 \cdot T_0 \end{cases}$$

It is very desirable that both the EIGHT-POINT data and the HIGH-LOW data achieve similar results for all temperatures. Indeed they do, with one notable exception.

If all the EIGHT-POINT sampled data are below the threshold, $T(t_k)<T_0$ for $k=0,1,...,7$, then both the EIGHT-POINT and HIGH-LOW data sets will yield zero degree-days. Similarly, if all the sampled data are above the threshold, $T(t_k)>T_0$ for $k=0,1,...,7$, the data sets will be nearly identical:

$$\tau(\text{EIGHT-POINT}) = \frac{1}{8} \cdot \sum_{k=0}^{7} [T(t_k)-T_0]$$

$$= \frac{1}{8} \cdot \sum_{k=0}^{7} T(t_k) - T_0 \qquad [3.21]$$

$$= \mu' - T_0$$

$$\approx M - T_0$$

$$= \tau(\text{HIGH-LOW})$$

On the other hand, if $T(t_j)>T_0$ for some j but $T(t_k)<T_0$ for some k, the EIGHT-POINT data will yield a non-zero degree-day but the HIGH-LOW data likely will not. Therefore, at temperatures close to threshold, there will be a significant bias toward higher physiological times with the EIGHT-POINT data! This can be the source of a significant potential problem since this situation

invariably occurs in the spring. Thus, during the critical germination stages, the modeling process is most susceptible to error.

These facts are exhibited in Table 3.7. This table compares the number of degree-days produced during the months of January, February and March of 1985 in Lansing, Michigan using a threshold of $T_0=25^{o}$. A detailed look at the daily output bears out the above observations. Also notice that the monthly ratio (HIGH-LOW/EIGHT-POINT) improves as the average temperature increases.

January:	0.730
February:	0.881
March:	0.946

Table 3.7: DEGREE-DAY Calculations using Straight HIGH-LOW data (Compared against EIGHT-POINT data), °F Lansing, Michigan (1985).

	EIGHT POINT	HIGH LOW	Ratio	EIGHT POINT	HIGH LOW	Ratio	EIGHT POINT	HIGH LOW	Ratio
1	3.00	0.50	0.167	0.00	0.00	1.000	7.50	6.50	0.867
2	0.00	0.00	1.000	0.00	0.00	1.000	2.75	4.00	1.455
3	0.00	0.00	1.000	0.00	0.00	1.000	8.00	9.00	1.125
4	1.50	0.00	0.000	0.00	0.00	1.000	1.13	0.00	0.000
5	3.38	2.00	0.592	0.00	0.00	1.000	1.38	0.00	0.000
6	6.13	6.00	0.979	0.00	0.00	1.000	5.50	5.00	0.909
7	5.63	5.00	0.888	0.00	0.00	1.000	14.13	14.50	1.026
8	0.00	0.00	1.000	0.00	0.00	1.000	9.25	8.50	0.919
9	0.00	0.00	1.000	0.00	0.00	1.000	19.63	19.50	0.993
10	0.00	0.00	1.000	0.00	0.00	1.000	19.75	20.00	1.013
11	0.00	0.00	1.000	3.63	2.50	0.689	11.88	12.00	1.010
12	0.50	0.00	0.000	3.75	2.50	0.667	7.00	6.50	0.929
13	0.13	0.00	0.000	0.00	0.00	1.000	10.25	10.00	0.976
14	1.88	0.00	0.000	0.00	0.00	1.000	5.13	5.50	1.072
15	0.00	0.00	1.000	0.00	0.00	1.000	11.63	15.00	1.290
16	0.00	0.00	1.000	0.13	0.00	0.000	3.13	2.50	0.799
17	0.00	0.00	1.000	2.00	2.50	1.250	4.63	1.00	0.216
18	0.00	0.00	1.000	4.88	3.00	0.615	15.25	18.50	1.213
19	0.00	0.00	1.000	0.38	0.00	0.000	9.38	10.50	1.119
20	0.00	0.00	1.000	4.38	2.00	0.457	7.38	6.50	0.881
21	0.00	0.00	1.000	11.50	10.50	0.913	13.00	11.00	0.846
22	0.00	0.00	1.000	15.38	15.00	0.975	11.25	11.50	1.022
23	0.00	0.00	1.000	17.38	17.50	1.007	10.25	9.50	0.927
24	0.38	0.00	0.000	9.00	9.50	1.056	9.88	10.00	1.012
25	0.00	0.00	1.000	10.25	10.50	1.024	18.75	19.00	1.013
26	0.00	0.00	1.000	12.63	14.50	1.148	28.50	29.00	1.018
27	0.50	0.00	0.000	1.25	1.00	0.800	22.50	24.50	1.089
28	0.00	0.00	1.000	7.50	6.00	0.800	21.63	22.00	1.017
29	0.00	0.00	1.000	12.25	14.00	1.143	8.13	8.50	1.046
30	0.25	0.00	0.000				11.00	15.50	1.409
31	0.00	0.00	1.000				5.88	6.50	1.105
Average			0.730			0.881			0.946

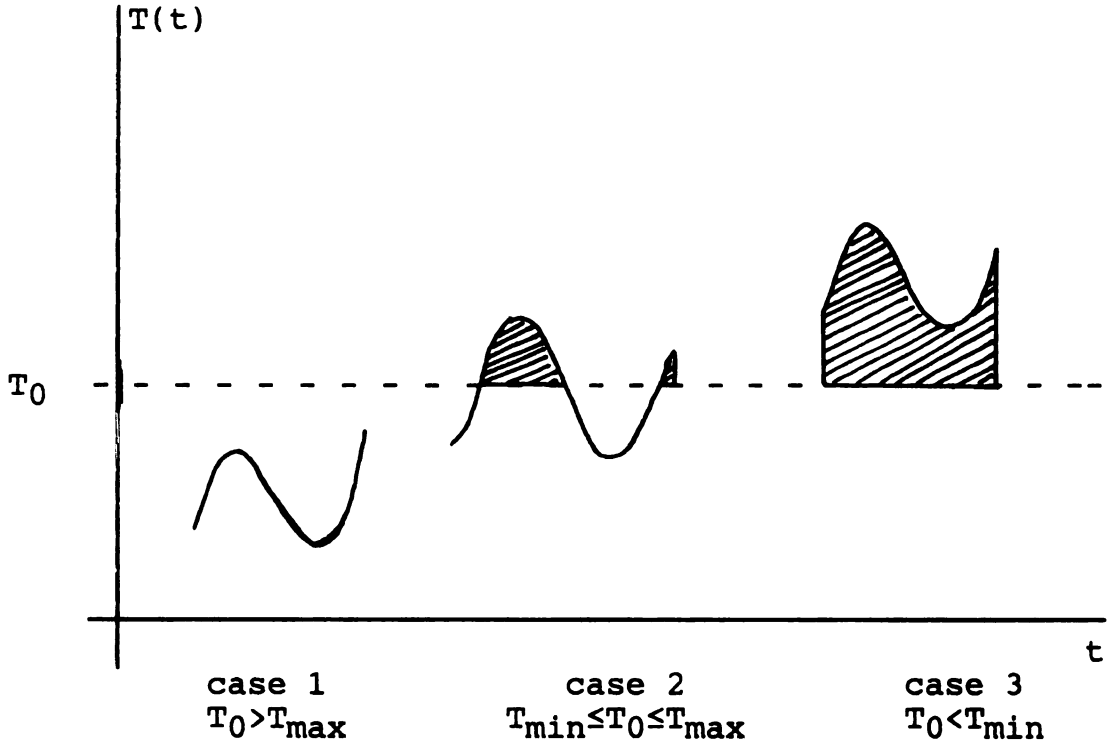


Figure 3.4 All possible cases for the sinusoidal heat unit model - TYPE I.

DEGREE-DAYS IN THE SINUSOID MODEL

In order to improve on the degree-days produced using HIGH-LOW data, we again consider the SINUSOID model. Using Figure 3.4 as a reference we note that this computation depends on where the threshold is located relative to the temperature profile.

There are three cases...

[3.22] Case 1. If $T_0 > T_{max}$, then $\tau = 0$.

Case 2. If $T_{min} < T_0 < T_{max}$ then:

$$\tau = \int_0^{t_1} [A \cdot \cos(2\pi t + \theta) + (B - T_0)]\, dt$$

$$+ \int_{t_2}^{1} [A \cdot \cos(2\pi t + \theta) + (B - T_0)]\, dt$$

$$= \left[\frac{A}{2\pi} \cdot \sin(2\pi t + \theta) + (B - T_0)t \right]_0^{t_1}$$

$$+ \left[\frac{A}{2\pi} \cdot \sin(2\pi t+\theta) + (B-T_0)t \right]_{t_2}^{1}$$

where $A \cdot \cos(2\pi t_1+\theta)+B = T_0$

[3.23] $$t_1 = \frac{\cos^{-1}((T_0-B)/A)-\theta)}{2\pi}$$

$$= \frac{\cos^{-1}(z)-\theta}{2\pi}$$

defining $z = (T_0-B)/A$

[3.24] $$= \frac{2 \cdot T_0 - T_{max} - T_{min}}{T_{max} - T_{min}}$$

and $A \cdot \cos(2\pi t_2+\theta) + B = T_0$

[3.25] $$t_2 = 1 - \frac{\cos^{-1}((T_0-B)/A)+\theta)}{2\pi}$$

$$= 1 - \frac{\cos^{-1}(z) + \theta}{2\pi}$$

also, $\delta = t_2-t_1$

[3.26] $$= 1 - \frac{1}{\pi} \cdot \cos^{-1}(z)$$

Thus,

$$\tau = \frac{A}{2\pi}\cdot\sin(2\pi t_1+\theta) + (B-T_0)\cdot t_1 - \frac{A}{2\pi}\cdot\sin(\theta)$$

$$+ \frac{A}{2\pi}\cdot\sin(2\pi+\theta) + (B-T_0) - \frac{A}{2\pi}\cdot\sin(2\pi t_2+\theta)$$

$$- (B-T_0)t_2$$

$$= \frac{A}{2\pi}\cdot[\sin(2\pi t_1\theta)-\sin(2\pi t_2+\theta)] + (B+T_0)\cdot(1-\delta)$$

[3.27]
$$= \frac{A}{\pi}\cdot[\sqrt{(1-z^2)} - z\cdot\cos^{-1}(z)]$$

Case 3. If $T_2<T_{min}$, simple geometry gives:

$$\tau = B - T_2$$

[3.28]
$$= \tfrac{1}{2}(T_{max}+T_{min})-T_2$$

$$= T(t) - T_0$$

This is an important result. It says that we may disregard the effect of daily fluxuations in the [instantaneous] temperature profile T(t) **if we instead consider two discrete [daily] profiles:**

i. Mean temperature profile: M(t), t=1,...,p

ii. Temperature range profile: R(t), t=1,...,p

In each case t is to be measured in days.

Using equations [3.5] we express T_{max} and T_{min} in terms of these profiles:

$$T_{max} = M(t) + \tfrac{1}{2}R(t)$$

[3.29]

$$T_{min} = M(t) - \tfrac{1}{2}R(t)$$

Thus, assuming only a lower threshold (in what will later be termed a **Type I model**, equations [3.22], [3.27] and [3.28] are summarized in equations [3.30] and [3.31].

$$[3.30] \qquad \tau(t) = \begin{cases} 0\ , & T_0 > T(t)+\tfrac{1}{2}R(t) \\ \dfrac{T(t)}{\pi}\cdot[\sqrt{(1-z^2)} - z\cdot\cos^{-1}(z)], & T(t)-\tfrac{1}{2}R(t) < T_0 < T(t)+\tfrac{1}{2}R(t) \\ T(t)-T_0\ , & T_0 < T(t)-\tfrac{1}{2}R(t) \end{cases}$$

$$[3.30] \qquad \text{where} \qquad z = \frac{2\cdot(T_0-T(t)}{R(t)}$$

In chapter seven this fact will be exploited in that for annual distributions, no daily fluxuations will be noted, but by equation [3.30] they will still be accounted for!

THE DOUBLE THRESHOLD MODEL

It is possible to extend this basic single (lower) threshold model to one in which there is also an upper threshold above which no activity takes place. Later this will be formally defined to be a **Type III model**. It has the effect of increasing the species physiological time only with an active band (see Figure 3.5):

$$T_0<T(t)<T_1$$

Again we assume a sinusoidal model. However, this time there are six possible cases (see Figures 3.6). It is necessary to find a formulation for the physiological time $\tau(t)$. The physiological time is simply the integral of the sinusoid over the active region:

$$T(t) = A\cdot\cos(2\pi t+\theta) + B$$

Figures 3.6 show this area shaded for each case. These cases are considered separately below.

Case 1: $T_{max}<T_0$

[3.32] $\tau = 0$

Case 2: $T_0 < T_{max} < T_1$, $T_{min} < T_0$

[3.33]
$$\tau = \int_{t_1}^{t_2} [A \cdot \cos(2\pi+\theta)+B-T_0]dt$$

$$= \left[\frac{A}{2\pi} \cdot \sin(2\pi+\theta)+(B-T_0)t \right]_{t_1}^{t_2}$$

$$= \frac{M(t)}{\pi} \cdot \left[\sqrt{(1-z_0{}^2)} + z \cdot [\pi-\cos^{-1}(z_0)]\right]$$

$$\text{where } z_0 = \frac{2 \cdot [T_0-M(t)]}{R(t)}$$

Case 3: $T_1 > T_{max}$, $T_0 < T_{min}$

[3.34]
$$\tau = \int_{0}^{1} [A \cdot \cos(2\pi t+\theta)+B-T_0]dt$$

$$= \left[\frac{A}{2\pi} \cdot \sin(2\pi+\theta)+(B-T_0)t \right]_{0}^{1}$$

$$= M(t) - T_0$$

Case 4: $T_1<T_{max}$, $T_0<T_{min}<T_1$

$$\tau = \int_0^{t_3} [A\cdot\cos(2\pi t+\theta)+B-T_0]dt + (T_1-T_0)\cdot(t_4-t_3) + \int_{t_4}^{1} [A\cdot\cos(2\pi t+\theta)+B-T_0]dt$$

$$= \left[\frac{A}{2\pi}\cdot\sin(2\pi+\theta)+(B-T_0)t \right]_0^{t_3} + (T_1-T_0)\cdot(t_4-t_3) + \left[\frac{A}{2\pi}\cdot\sin(2\pi+\theta)+(B-T_0)t \right]_{t_4}^{1}$$

[3.35] where $t_3 = \dfrac{\cos^{-1}(z_1)-\theta}{2\pi}$

[3.36] $$z_1 = \frac{T_1-B}{A} = \frac{2\cdot T_1-T_{max}-T_{min}}{T_{max}-T_{min}} = \frac{T_1-M(t)}{\frac{1}{2}R(t)}$$

[3.37] and $t_4 = \dfrac{\cos^{-1}(z_0)-\theta}{2\pi}$

[3.38]
$$z_0 = \frac{T_0-B}{A} = \frac{2\cdot T_0-T_{max}-T_{min}}{T_{max}-T_{min}}$$

$$= \frac{T_0-M(t)}{\frac{1}{2}R(t)}$$

Thus,
$$\tau = \frac{A}{2\pi}\cdot\sin(2\pi t_3+\theta)+(B-T_0)\cdot t_3 - \frac{A}{2\pi}\cdot\sin(\theta)$$

$$+ (T_1-T_0)\cdot(t_4-t_3) + \frac{A}{2\pi}\cdot\sin(2\pi+\theta)$$

$$+ (B-T_0) - \frac{A}{2\pi}\cdot\sin(2\pi t_4+\theta) - (B-T_0)\cdot t_4$$

[3.39]
$$= \frac{M(t)}{2\pi}\cdot\left[\sqrt{(1-z_1^2)} + z_1\cdot\cos^{-1}(z_1)\right]$$

$$+ (T_1-T_0)\cdot\left[1 - \frac{1}{\pi}\cdot\cos^{-1}(z_1)\right]$$

Case 5: $T_1 < T_{max}$, $T_0 > T_{min}$

$$\tau = \int_0^{t_1} [A \cdot \cos(2\pi t + \theta) + B - T_0]\,dt$$

$$+ \int_{t_2}^{t_3} [A \cdot \cos(2\pi t + \theta) + B - T_0]\,dt$$

$$+ (T_1 - T_0) \cdot (t_4 - t_3)$$

$$+ \int_{t_4}^{1} [A \cdot \cos(2\pi t + \theta) + B - T_0]\,dt$$

$$= \left[\frac{A}{2\pi} \cdot \sin(2\pi + \theta) + (B - T_0)t \right]_0^{t_1}$$

$$+ \left[\frac{A}{2\pi} \cdot \sin(2\pi + \theta) + (B - T_0)t \right]_{t_2}^{t_3}$$

$$+ (T_1 - T_0) \cdot (t_4 - t_3)$$

$$+ \left[\frac{A}{2\pi} \cdot \sin(2\pi + \theta) + (B - T_0)t \right]_{t_4}^{1}$$

$$= \frac{A}{2\pi} \cdot \left[\sin(2\pi t_1 + \theta) - \sin(\theta) \right]$$

$$+ \frac{A}{2\pi} \cdot \left[\sin(2\pi t_3 + \theta) - \sin(2\pi t_2 + \theta) \right]$$

$$+ (T_1 - T_0) \cdot (t_4 - t_3)$$

$$+ \frac{A}{2\pi} \cdot \left[\sin(2\pi + \theta) - \sin(2\pi t_4 + \theta) \right]$$

[3.40]

$$= \frac{M(t)}{\pi} \cdot \left[\sqrt{(1 - z_0^2)} + \sqrt{(1 - z_1^2)} \right]$$

$$+ (T_1 - T_0) \cdot \left[1 - \frac{1}{\pi} \cdot \cos^{-1}(z_1) \right]$$

Case 6: $T_1 < T_{min}$

[3.41] $\tau = 0$

The physiological time $\tau(t)$ for the day t in question is summarized for each case by noting that

[3.42]
$$T_{max} = T(t) + \tfrac{1}{2}R(t)$$
$$T_{min} = T(t) - \tfrac{1}{2}R(t)$$

in equation [3.43].

This result permits the computation of physiological time given the **daily mean temperature profile** and the **daily temperature range profile.** The diurnal fluxuations are not retained, but their effect is still present in the actual heat unit (degree-day) calculation. Both the daily mean temperature and daily temperature range profiles may be immediately calculated from the standard HIGH-LOW data as maintained by the National Weather Bureau. Examples follow.

[3.43]

$$\tau(t) = \begin{cases} 0, & M(t)+\tfrac{1}{2}R(t)<T_0 \\ \dfrac{M(t)}{\pi}\cdot\left[\sqrt{(1-z_0^2)}+z_0\cdot[\pi-\cos^{-1}(z_0)]\right], & \\ \qquad M(t)-\tfrac{1}{2}R(t)<T_0<M(t)+\tfrac{1}{2}R(t)<T_1 & \\ M(t)-T_0, \quad T_1>M(t)+\tfrac{1}{2}R(t),\ T_0<M(t)-\tfrac{1}{2}R(t) & \\ \dfrac{M(t)}{2\pi}\cdot\left[\sqrt{(1-z_1^2)}+z_1\cdot\cos^{-1}(z_1)\right] & \\ \qquad +(T_1-T_0)\cdot\left[1-\dfrac{1}{\pi}\cdot\cos^{-1}(z_1)\right], & \\ \qquad T_1<M(t)+\tfrac{1}{2}R(t),\ T_0<T(t)-\tfrac{1}{2}R(t)<T_1 & \\ \dfrac{M(t)}{\pi}\cdot\left[\sqrt{(1-z_0^2)}+\sqrt{(1-z_1^2)}\right] & \\ \qquad +(T_1-T_0)\cdot\left[1-\dfrac{1}{\pi}\cdot\cos^{-1}(z_1)\right], & \\ \qquad T_1<M(t)+\tfrac{1}{2}R(t),\ T_0>M(t)-\tfrac{1}{2}R(t) & \\ 0, & M(t)-\tfrac{1}{2}R(t)>T_1 \end{cases}$$

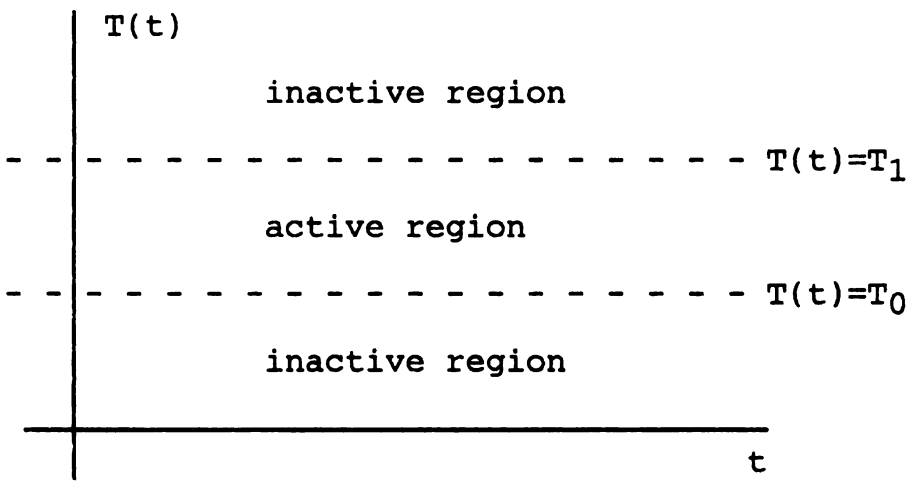


Figure 3.5 The active region for the TYPE III model.

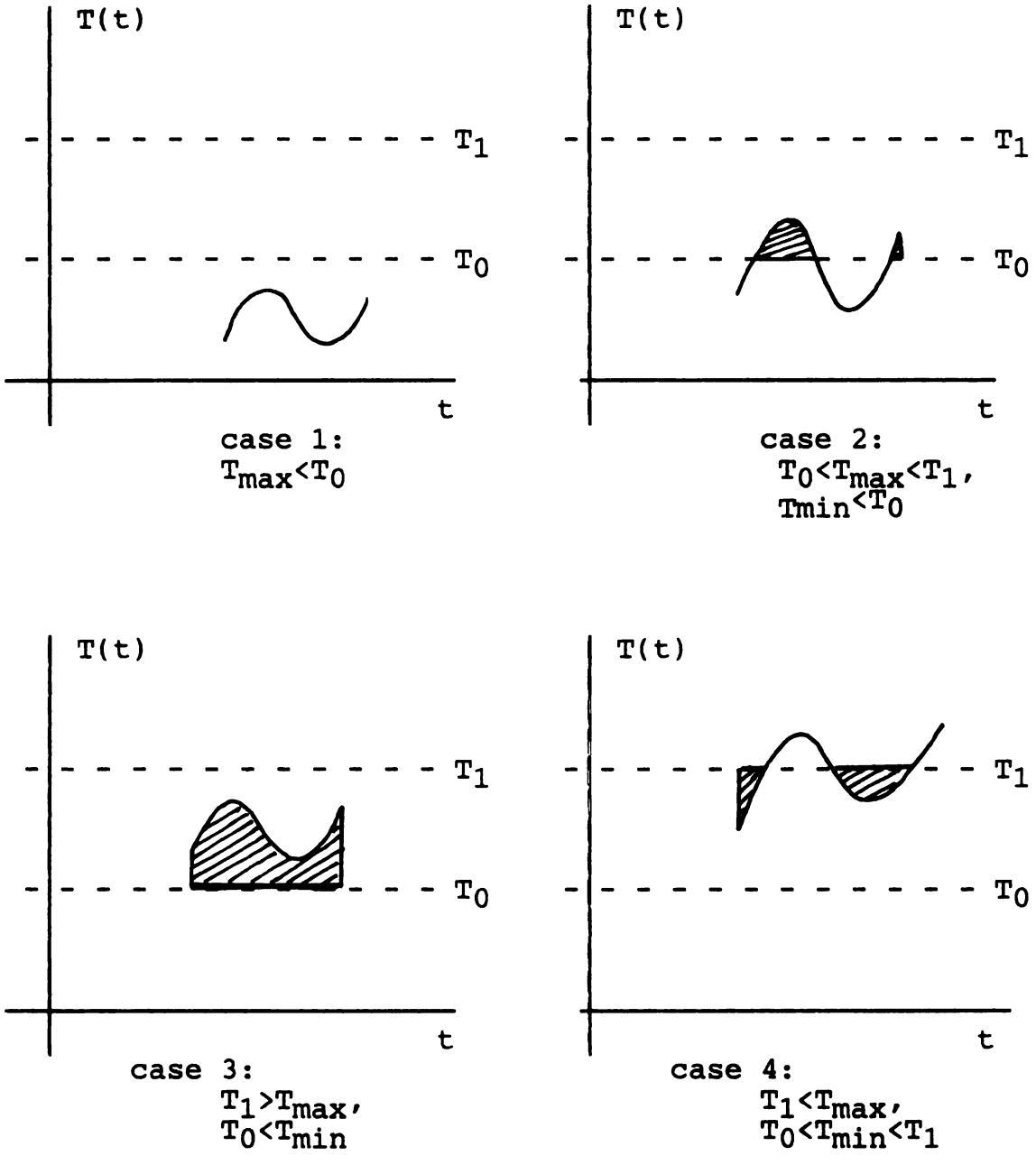


Figure 3.6a: The first four cases of a TYPE III model using the sinusoidal approximation.

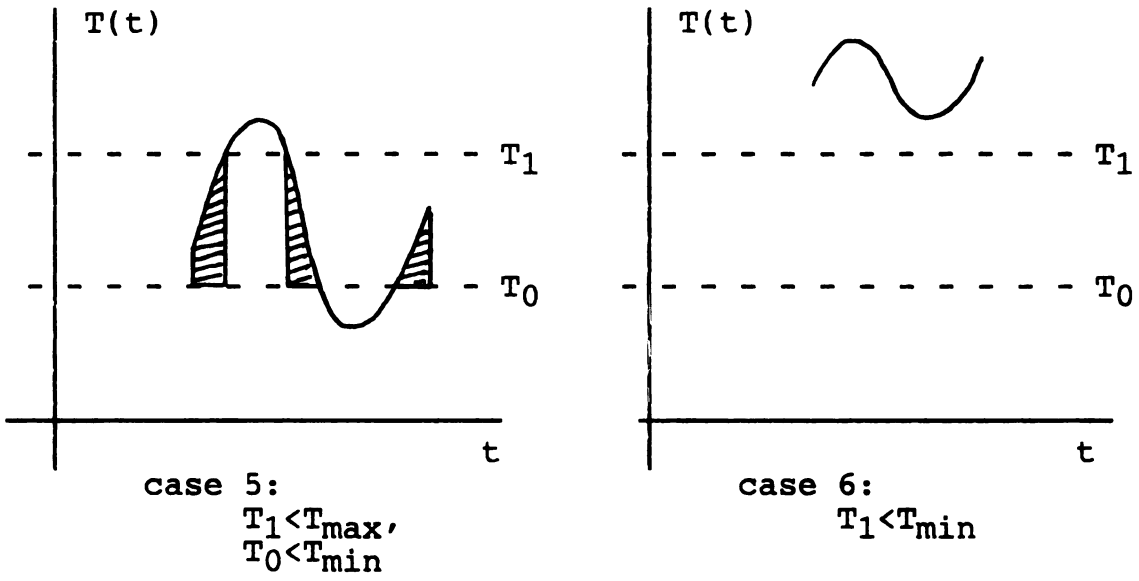


Figure 3.6b: Cases 5 and 6 for the TYPE III model using a sinusoidal approximation.

APPLICATION OF THE SINUSOID MODEL

Equation [3.30] provides an improved means for computing degree-day estimates using only HIGH-LOW data. Table 3.8 shows where January, February and March degree-day results are computed using EIGHT-POINT data and HIGH-LOW data in the SINUSOID mode. Note the following facts:

i. The EIGHT-POINT data produces zero degree-days when and only when the SINUSOID model produces zero degree-days. This eliminates the downward bias of the straight HIGH-LOW calculation

ii. There is a reasonable correlation between degree-days as computed by the SINUSOID model.

iii. At higher temperatures (compared to the threshold), both the "straight" HIGH-LOW data and the HIGH-LOW data using the SINUSOID model produce identical results.

Again we compare progressive monthly average ratios...

January: 0.928
February: 1.061
March: 1.055

These ratios are consistently (and significantly) close to unity while those in Table 3.7 are not. Clearly the SINUSOID model produces a superior estimate of the degree-day measure.

Table 3.8: DEGREE-DAY Calculations using HIGH-LOW with SINUSOID Model (Compared against EIGHT-POINT data), °F Lansing, Michigan (1985).

	EIGHT POINT	HIGH LOW	Ratio	EIGHT POINT	HIGH LOW	Ratio	EIGHT POINT	HIGH LOW	Ratio
1	3.00	2.33	0.777	0.00	0.00	1.000	7.50	6.50	0.867
2	0.00	0.00	1.000	0.00	0.00	1.000	2.75	4.14	1.505
3	0.00	0.00	1.000	0.00	0.00	1.000	8.00	9.68	1.210
4	1.50	0.76	0.507	0.00	0.00	1.000	1.13	1.51	1.336
5	3.38	2.18	0.645	0.00	0.00	1.000	1.38	1.15	0.833
6	6.13	6.00	0.979	0.00	0.00	1.000	5.50	5.33	0.969
7	5.63	5.00	0.888	0.00	0.00	1.000	14.13	14.50	1.026
8	0.00	0.00	1.000	0.00	0.00	1.000	9.25	8.50	0.919
9	0.00	0.00	1.000	0.00	0.00	1.000	19.63	19.50	0.993
10	0.00	0.00	1.000	0.00	0.00	1.000	19.75	20.00	1.013
11	0.00	0.00	1.000	3.63	3.47	0.956	11.88	12.00	1.010
12	0.50	0.37	0.740	3.75	3.18	0.848	7.00	6.50	0.929
13	0.13	0.12	0.923	0.00	0.00	1.000	10.25	10.00	0.976
14	1.88	1.32	0.702	0.00	0.00	1.000	5.13	6.04	1.177
15	0.00	0.00	1.000	0.00	0.00	1.000	11.63	15.00	1.290
16	0.00	0.00	1.000	0.13	0.10	0.769	3.13	2.66	0.850
17	0.00	0.00	1.000	2.00	2.91	1.455	4.63	4.02	0.868
18	0.00	0.00	1.000	4.88	4.23	0.867	15.25	18.50	1.213
19	0.00	0.00	1.000	0.38	0.65	1.711	9.38	10.50	1.119
20	0.00	0.00	1.000	4.38	4.25	0.970	7.38	7.78	1.054
21	0.00	0.00	1.000	11.50	10.50	0.913	13.00	11.63	0.895
22	0.00	0.00	1.000	15.38	15.00	0.975	11.25	11.50	1.022
23	0.00	0.00	1.000	17.38	17.50	1.007	10.25	9.50	0.927
24	0.38	0.22	0.579	9.00	9.50	1.056	9.88	10.00	1.012
25	0.00	0.00	1.000	10.25	10.50	1.024	18.75	19.00	1.013
26	0.00	0.00	1.000	12.63	14.50	1.148	28.50	29.00	1.018
27	0.50	0.31	0.620	1.25	2.44	1.952	22.50	24.50	1.089
28	0.00	0.00	1.000	7.50	7.31	0.975	21.63	22.00	1.017
29	0.00	0.00	1.000	12.25	14.00	1.143	8.13	8.50	1.046
30	0.25	0.35	1.400				11.00	15.50	1.409
31	0.00	0.00	1.000				5.88	6.50	1.105
Average			0.928			1.061			1.055

NUMERICAL COMPUTATION OF THE PHYSIOLOGICAL TIME

The value of this formulation of physiological time lies not so much in calculating the degree-days through out the course of a given day, but in obtaining an accurate cumulative total through the entire year. We therefore proceed to restate these results in the form of algorithms by which a physiological time scale can be calculated for each day of the year.

In order to calculate the physiological time, it is necessary that the following parameters be given:

t_0 = the initial chronological time of the species introduction.

T_0 = the lower temperature threshold of the model which is indigenous to the species.

T_1 = the upper threshold (if a two threshold model is used) of the model.

The cases and their corresponding algorithms are considered below.

i. Type I model given the Instantaneous temperature profile. (That is, the model assumes only the lower threshold and the temperature profile T(t) including diurnal fluxuations). See Algorithm 3.1.

ii. Type I model given the Daily mean temperature and daily temperature range profiles. (as above, but with the two discrete distribution profiles, M(t) and R(t)). See Algorithm 3.2.

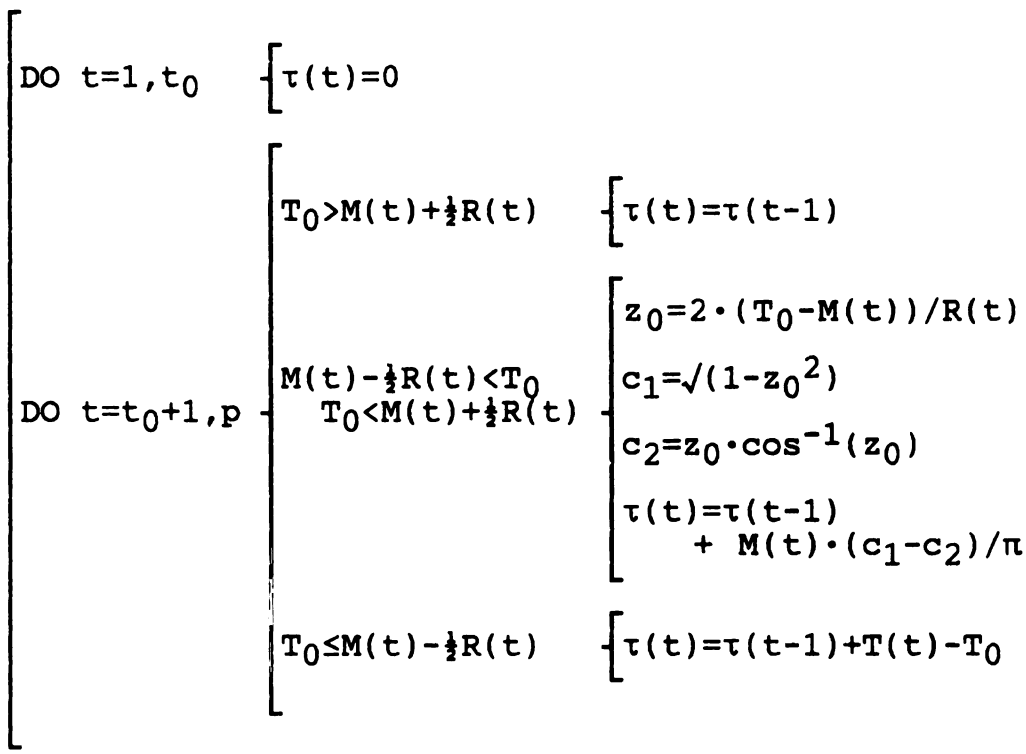
iii. Type III model given the Instantaneous temperature profile. (two thresholds and one profile T(t)). See Algorithm 3.3.

iv. Type III model given the Daily mean temperature and daily temperature range profiles. (two thresholds and two distribution profiles, M(t) and R(t)). See Algorithm 3.4.

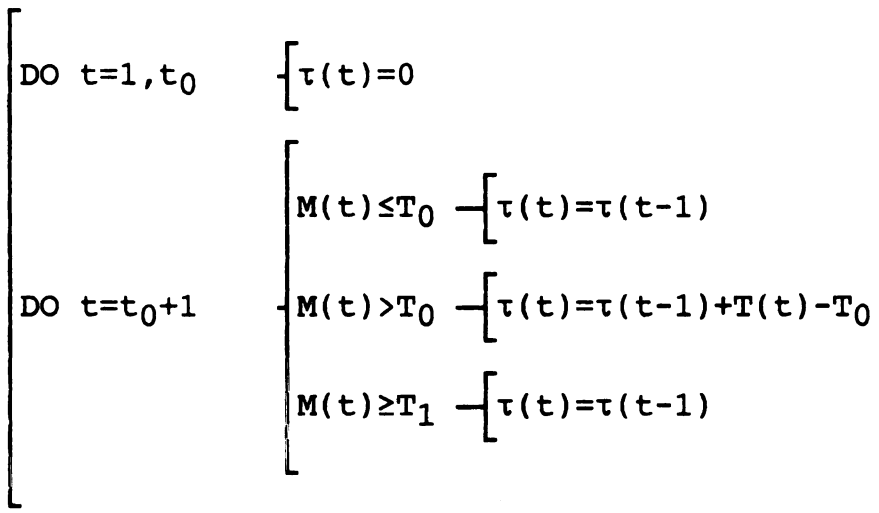
The implementation of these algorithms is straight forward and will be demonstrated in the chapters that follow.

$$
\begin{cases}
\text{DO } t=1,t_0 \;\rightarrow\; \{\tau(t)=0 \\
\text{DO } t=t_0+1,p \;\rightarrow\; \begin{cases} M(t)\le T_0 \;\rightarrow\; \{\tau(t)=\tau(t-1) \\ M(t)>T_0 \;\rightarrow\; \{\tau(t)=\tau(t-1)+T(t)-T_0 \end{cases}
\end{cases}
$$

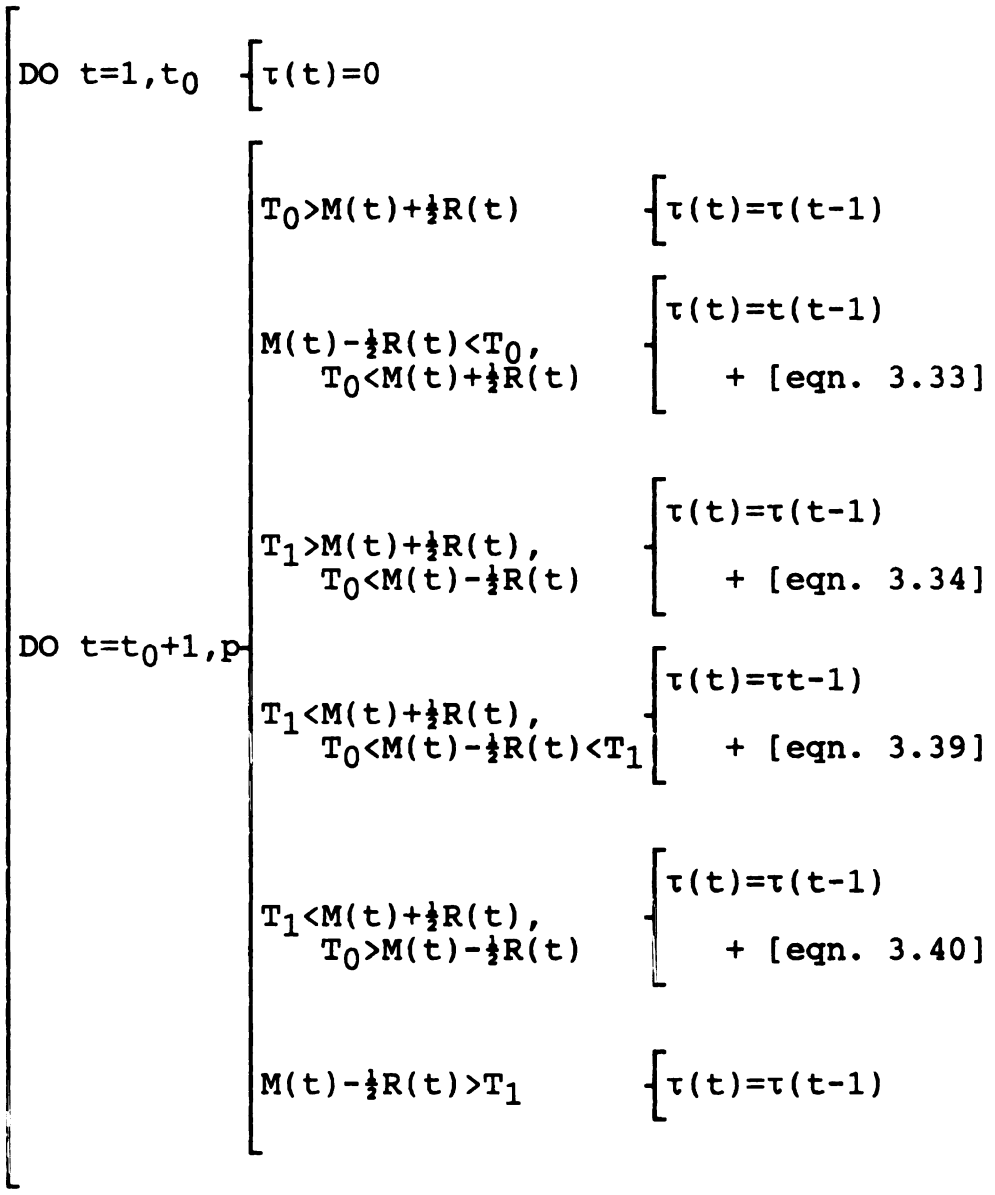
Algorithm 3.1 Computing the physiological time assuming a Type I model with a given instantaneous temperature profile T(t).



Algorithm 3.2: Computing the physiological time assuming a Type I model with daily mean temperature $M(t)$ and daily temperature range $R(t)$ profiles.



Algorithm 3.3: Computing the physiological time assuming a Type III model with a given instantaneous temperature profile T(t).



Algorithm 3.4 Computing the physiological time assuming a Type III model with daily mean temperature M(t) and daily temperature range R(t) profiles.

CONCLUSION

Degree-day calculations are important in many biological, primarily annual population, models. Due to the nature of weather data and the need for computational simplicity, the traditional statistics kept for temperature calculations are daily minimums and maximums (HIGH-LOW data). By implementing a sinusoid to model the oscillatory behavior of hourly temperature, we are able to improve on these traditional methods.

It has been shown that the HIGH-LOW statistics (mean and range) are related to the distribution statistics (mean and standard deviation) in a simple manner. This legitimizes their use in practical computations. Use of the SINUSOID model also produces formulae that are quite useful for temperatures close to the degree-day model threshold. This significantly improves the performance over the HIGH-LOW model.

The results are validated by comparing against EIGHT-POINT temperature data (see Table 3.8). In each case statistical significance was achieved.

CHAPTER FOUR

SOLVING SYSTEMS USING THE HEAT UNIT HYPOTHESIS

INTRODUCTION

The essential feature of a heat unit model is the assumption that **biological activity** $\beta(T)$ is a function of temperature. This is further simplified by assuming the existence of a threshold temperature T_0 where the system lies dormant for temperatures $T \leq T_0$ and linear for temperatures in excess of T_0. Mathematically this can be stated as follows:

$$[4.1] \qquad \beta(T) = \begin{cases} 0, & T < T_0 \\ r \cdot [T(t) - T_0], & T \geq T_0 \end{cases}$$

where r is a proportionality constant.

This is displayed graphically in Figure 3.2.

We need to state this **heat unit hypothesis** more formally while at the same time showing how it applies to elementary biological modeling situations.

The objectives of this chapter are the following:

i. Formally state the heat unit hypothesis.

ii. Interpret the heat unit hypothesis intuitively.

iii. Apply this hypothesis to single species models.

iv. By means of examples, demonstrate both exact and approximate solution techniques.

v. Extend this formulation to multiple species models.

vi. Illustrate these solution techniques.

It will also be necessary to compare these heat unit results against the classical. To do this it is first necessary to calculate the critical parameters of both system descriptions. This will be done in a later chapter.

AN ELEMENTARY INTERPRETATION

Consider a single species model with initial population x_0. If $x(t)$ represents the population at **chronological time** t, we think of the **net growth rate** as being proportional to the **biological activity.** By equation [4.1],

$$[4.2] \qquad \frac{dx(t)/dt}{x(t)} = \beta(T(t))$$

Since the left-hand-side of this is easily integrable,

$$[4.3] \qquad \int_{x_0}^{x(t)} dx/x = r \cdot \int_{t_0}^{t} \beta(T(\mu))d\mu$$

$$[4.4] \qquad \ln [x(t)/x_0] = r \cdot \int_{t_0}^{t} \beta(T(\mu))d\mu$$

It is convenient to define a so-called **action set:**

$$[4.5] \qquad \Gamma_{t0,T0}(t) = \{t>t_0: T(t)\geq T_0\}$$

This set is a function of the three parameters t, t_0, T_0. However, it is often the case that the context of our problem is such that t_0 and T_0 are fixed. Thus, we write:

$$[4.6] \qquad \Gamma(t) = \Gamma_{t0,T0}(t)$$

We also define the **norm** of $\Gamma(t)$ as the measure of the action set:

$$[4.7] \qquad \|\Gamma(t)\| = \text{meas}\{t \geq t_0 : T(t) > T_0\} \equiv \int_{\Gamma(t)} d\mu$$

where $\text{meas}\{\cdot\}$ is the measure of the set.

Using this notation, the solution can be written especially conveniently:

$$[4.6] \qquad \ln [x(t)/x_0] = r \cdot \int_{\Gamma(t)} [T(\mu)-T_0]d\mu$$

$$[4.8] \qquad x(t) = x_0 \cdot \exp \left[r \cdot \int_{\Gamma(t)} (T(\mu)-T_0)d\mu \right]$$

Thus, in principle, once the temperature profile is known, equation [4.8] gives an explicit formula for x(t). However, since this profile T(t) is generally an empirical function, numerical methods must generally be employed from the outset.

It is worth comparing this new model against the classical (exponential) model with solution as follows:

[4.9] $$x(t) = x_0 \cdot e^{r' \cdot t}$$

where r' is a constant, not necessarily the same as r.

Comparison of equations [4.8] and [4.9] clearly show the effect of employing the heat-unit hypothesis: the characteristic coefficient is now time varying.

THE HEAT UNIT HYPOTHESIS

From equation [4.2] it is apparent that the integral

$$[4.10] \qquad \tau(t) = \int_{\Gamma(t)} [T(\mu)-T_0]d\mu$$

plays an important role. Thus we name it: the **physiological time**. Before showing the importance of this definition, we list several useful properties.

THEOREM 4.1. Let the physiological time $\tau(t)$ be defined as in equation [4.10]. The following properties hold for all $t \geq t_0$.

[4.11]

i. $\tau(t_0) = 0$

ii. $\tau(t) \geq 0$

iii. $\tau(t_2) \geq \tau(t_1)$ for $t_2 \geq t_1$

iv. $\tau(t) = \text{constant}$ for $t \in \Gamma'(t)$

v. $$\frac{d\tau(t)}{dt} = \begin{cases} 0, & T(t)<T_0 \\ T(t)-T_0, & T(t)\geq T_0 \end{cases}$$

$$= \frac{1}{r}\cdot\beta(T(t))$$

The proof of this theorem follows immediately from the definition of physiological time (equation [4.10]).

The intuitive interpretation of these properties is that the physiological time τ is a non-negative, monotonic increasing function that begins at the point $\tau(t_0)=0$. It is strictly monotonic increasing for points in the action set. It is constant for points not in the action set. Figure 4.2 illustrates these relationships for an arbitrary profile $T(t)$.

We now formally state the HEAT UNIT HYPOTHESIS:

> The population rate of change in physiological time is proportional to the effective temperature experienced by the species in question (ie. the temperature in excess of the threshold for a Type I model). This rate of change is proportional to the biological activity of this species (in what ever way this activity is defined).

We further justify the term **physiological time**.

Equation [4.2] may be restated in derivative form as:

$$[4.12] \qquad \frac{dx(t)/dt}{x(t)} = r \cdot \frac{d\tau(t)}{dt}$$

or

$$[4.13] \qquad \frac{dx(t)}{d\tau} = r \cdot x(t)$$

By changing variables, from chronological to physiological time (i.e. using equation [4.10]), it follows that:

$$[4.14] \qquad \frac{d\underline{x}(\tau)}{\underline{x}(\tau)} = r \cdot d\tau, \qquad \underline{x}(0) = x_0$$

with $\underline{x}(\tau) = \underline{x}(\tau(t)) = x(t)$ for all t

The functions x and $\underline{x}$ are not the same. However, since we work only in one of the two time systems in a given problem, no confusion should arise if we simply refer to either of them as x. Thus, we will generally write:

$$[4.15] \qquad \frac{d\,x(\tau)}{x(\tau)} = r \cdot d\tau, \qquad x(0) = x_0$$

An interesting result! By using the physiological time scale, we have the classical exponential model. In other words, the species "sees" the usual classical statement in its own time base.

This most important property is true for a more general class of systems as well. Suppose we have n species, each

with population $x_i(t)$, $i=1...,n$. Each will behave according to its own initial time t_{i0} and threshold T_i. The temperature thresholds, in turn, will determine characteristic action sets and physiological time bases.

While there are differing techniques for modeling species interactions, many relate a species rate of change only to the **current population state variables.** Such a system's state equations relate to the biological activity, $d\tau_i(t)/d\tau$ as the system of equations [4.16].

$$
\begin{aligned}
&\frac{dx_1(t)/dt}{g(x_1(t),\ldots,x_n(t))} = \frac{d\tau_1(t)}{dt} \quad , \quad x_1(t_{10})=x_{10} \\
&\frac{dx_2(t)/dt}{g(x_1(t),\ldots,x_n(t))} = \frac{d\tau_2(t)}{dt} \quad , \quad x_2(t_{20})=x_{20} \qquad [4.16] \\
&\qquad\qquad \vdots \qquad\qquad\qquad\qquad \vdots \\
&\frac{dx_n(t)/dt}{g(x_1(t),\ldots,x_n(t))} = \frac{d\tau_n(t)}{dt} \quad , \quad x_n(t_{n0})=x_{n0}
\end{aligned}
$$

Notice that the functions $g_i(x_1(t),\ldots,x_n(t))$ have no discrete delays or derivatives. Since this model occurs so frequently, we call such a system **current state sensing.** By writing system [4.16] in terms of the physiological times rather than the chronological time t, we now have n different time scales rather than one. However, the system

is easier to describe (and solve) as the system of equations [4.17].

$$
\begin{aligned}
&\frac{dx_1(\tau_1)}{d\tau_1} = g(x_1(t),\ldots,x_n(t)) \quad , \quad x_1(0)=0 \\
&\frac{dx_{21}(\tau_2)}{d\tau_2} = g(x_1(t),\ldots,x_n(t)) \quad , \quad x_2(0)=0 \\
&\qquad\qquad \vdots \qquad\qquad\qquad\qquad \vdots \\
&\frac{dx_n(\tau_n)}{d\tau_n} = g(x_1(t),\ldots,x_n(t)) \quad , \quad x_n(0)=0
\end{aligned}
\tag{4.17}
$$

These results are summarized in the form of a theorem.

> THEOREM 4.2: Let S be an n^{th} order current state sensing system with each species having a physiological time base defined as in equation [4.10]. Then S can be written in the form of equation [4.17], ie. in the classical (non heat unit hypothesis) system description by using the [species specific] physiological time base.

Application of this result follows in succeeding chapters.

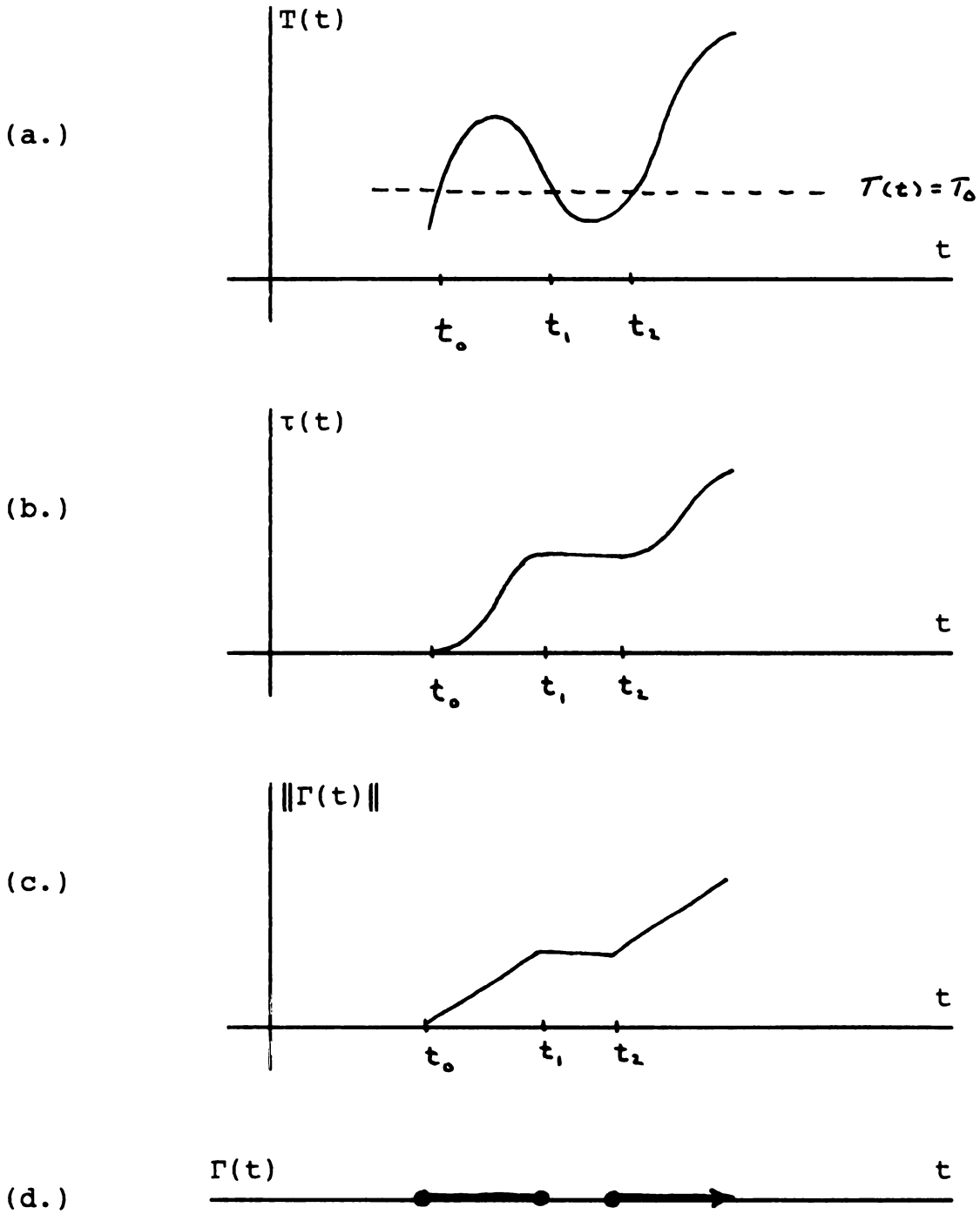


Figure 4.2 The relationship between fundamental heat-unit functions.

a. Temperature Profile
b. Physiological time
c. Action set norm
d. Action set

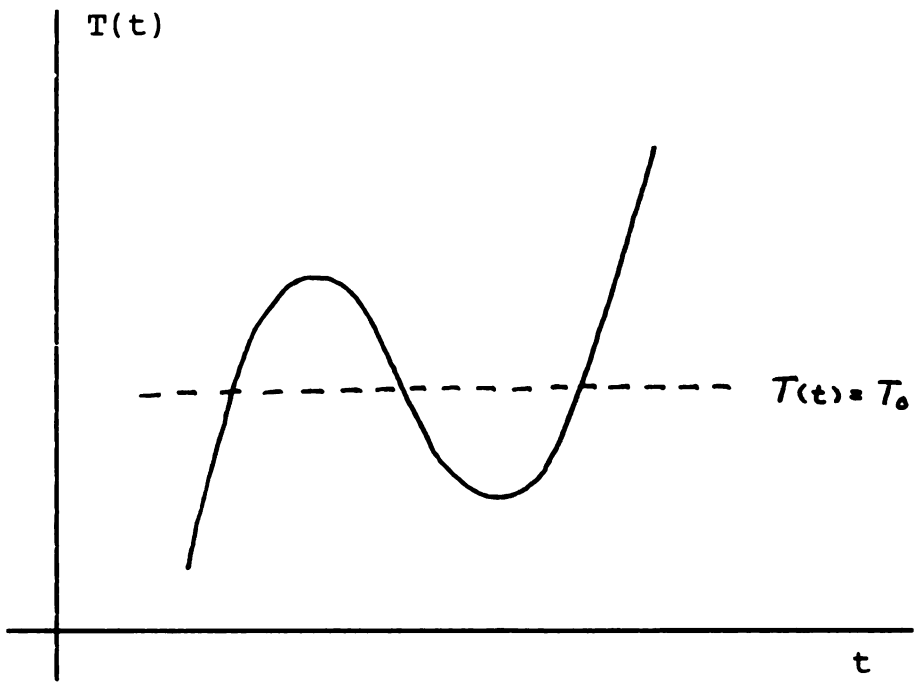
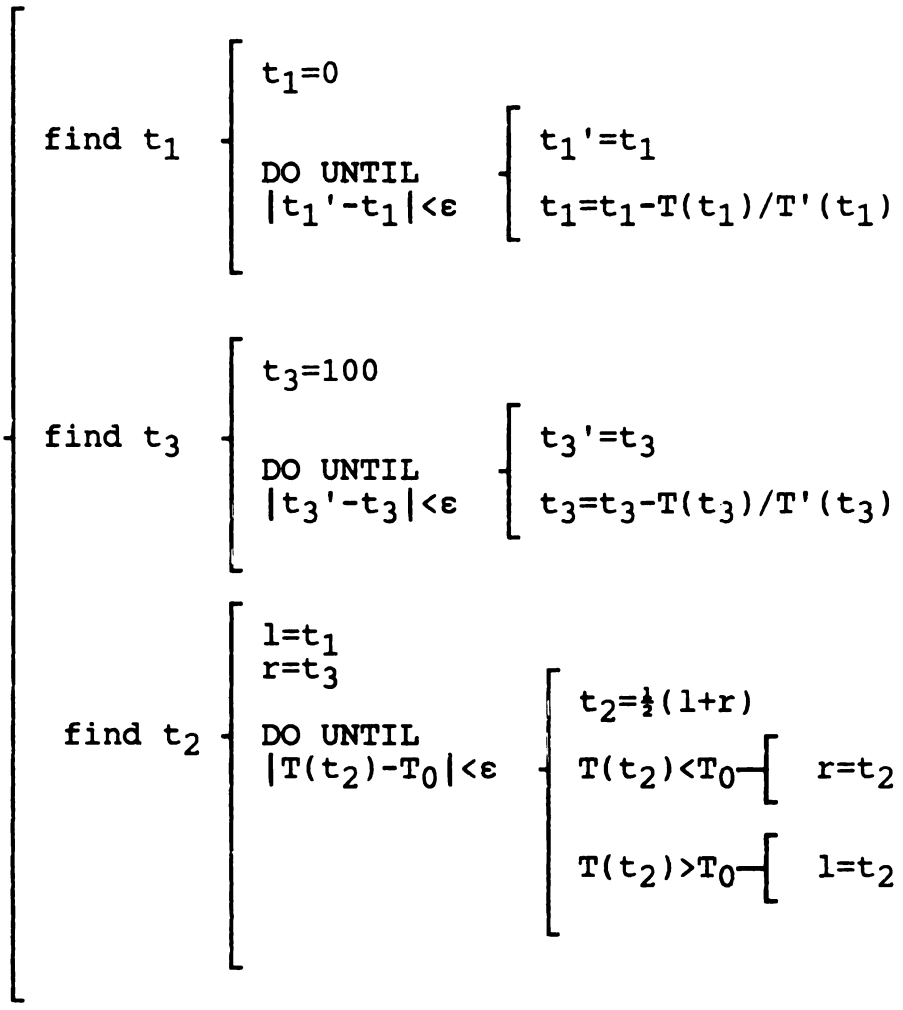


Figure 4.3: A cubic temperature profile.

$T(t) = a \cdot t^3 + b \cdot t^2 + c \cdot t + d$



where T(t) is a defined in equation [4.18],

ε is an accuracy factor: $\varepsilon \approx .01$

Algorithm 4.1 Computing the critical points for a cubic temperature profile.

$$
\begin{cases}
t_0<t_1 & \begin{cases} t<t_1 & \tau(t)=0 \\ t_1\le t<t_2 & \tau(t)=J(t_1,t) \\ t_2\le t<t_3 & \tau(t)=J(t_1,t_2) \\ t\ge t_3 & \tau(t)=J(t_1,t_2)+J(t_3,t) \end{cases} \\
t_1\le t_0<t_2 & \begin{cases} t<t_2 & \tau(t)=J(t_0,t) \\ t_2\le t<t_3 & \tau(t)=J(t_0,t_2) \\ t\ge t_3 & \tau(t)=J(t_0,t_2)+J(t_3,t) \end{cases} \\
t_2\le t_0<t_3 & \begin{cases} t<t_3 & \tau(t)=0 \\ t\ge t_3 & \tau(t)=J(t_3,t) \end{cases} \\
t_0\ge t_3 & \{\ \tau(t)=J(t_0,t)
\end{cases}
$$

where it is assumed that $t>t_0$ in all cases,

[4.22] $$J(\mu_1,\mu_2) = \int_{\mu_1}^{\mu_2} [T(\mu)-T_0]d\mu$$

Algorithm 4.2 Computing the physiological time from a general cubic temperature profile given the critical points (as determined by Algorithm 4.1).

EXAMPLE ONE

Suppose the daily mean temperature can be described using the following temperature profile:

[4.18] $$T(t) = a \cdot t^3 + b \cdot t^2 + c \cdot t + d$$

where $a = 0.00096$, $b = -0.114$, $c = 5.4$, $d = 20$

on the interval [0,100]. Let us also consider thresholds T_0 on the interval [20,80]. Figure 4.3 illustrates this situation where $t_0<t_1<t_2<t_3<t$ and t_0, t are the initial and final times while t_1, t_2, t_3 are the end points of the action set $\Gamma(t)$.

[4.19] $$T(t_1) = T_0$$
$$T(t_2) = T_0$$
$$T(t_3) = T_0$$

also, $$\Gamma(t) = \{t>t_0: t_1 \le t \le t_2 \text{ or } t \ge t_3\}$$

[4.20] $$= \{t>t_0: t_1 \le t \le t_2\} \cup \{t>t_0: t \ge t_3\}$$

$$= \Gamma_1(t) + \Gamma_2(t)$$

$$[4.21] \qquad \tau(t) = \int_{t_1}^{t_2} [a\mu^3+b\mu^2+c\mu+d-T_0]d\mu$$

$$+ \int_{t_3}^{t} [a\mu^3+b\mu^2+c\mu+d-T_0]d\mu$$

The specific values of t_1, t_2, t_3 are critical to evaluation of this integral and this is not easy analytically. Therefore, we use numerical methods. Algorithm 4.1 performs this task by first finding t_1, t_3 via the Newton-Raphson method and then computes the intermediate point t_2 using Bijection.

Generally speaking, t_0 and t can be anywhere on the interval [0,100]. Thus, the diagram in Figure 4.3 is illustrative of but one particular case. Each of these cases is handled in Algorithm 4.2.

Using these two Algorithms (4.1 and 4.2), it is possible to find the physiological time $\tau(t)$ from a general cubic temperature profile. Figure 4.4 shows a typical situation: given a fixed initial time t_0=20, a family of physiological time curves are plotted for various thresholds T_0. As in Figure 4.2, this family will have a common initial point $\tau(t_0)$=0 and proceed in a somewhat

par

r'(t

Ind

phy

thes

Aga

defi

Ther

time

time

t_0

same

t_f=

prop

parallel manner being horizontal over the inaction set $\Gamma'(t)$.

Indeed, two further properties relating members of the physiological time family are now evident. We also state these in the form of a theorem.

THEOREM 4.3 For any two thresholds T_{10} and T_{20} with corresponding physiological times $\tau_{10}(t)$ and $\tau_{20}(t)$ such that $T_{10}<T_{20}$,

i. The physiological times are weakly ordered:

$$\tau_{10}(t) \leq \tau_{20}(t) \qquad \text{for } t>t_0$$

ii. The action sets are weakly nested:

$$\Gamma_{10}(t) \leq \Gamma_{20}(t)$$

Again, these properties follow immediately from the definition of physiological time (equation 4.7).

There are also other illustrative views of the physiological time. Figure 4.5 plots the physiological time at the time horizon $\tau(t_f)$, where $t_f=80$, for various initial times t_0 and a typical threshold of $T_0=40$. Figure 4.6 does the same for various values of T_0 and a typical final time of $t_f=80$. Each of these views illustrate the important properties stated in Theorems 4.1 and 4.3.

These effects on the physiological time are important, but even more important is the manner in which the heat-unit hypothesis effects the population metric itself. Toward this end we consider several population models as illustrative cases in point. We begin by recalling two classical descriptions: Exponential and Logistic growth.

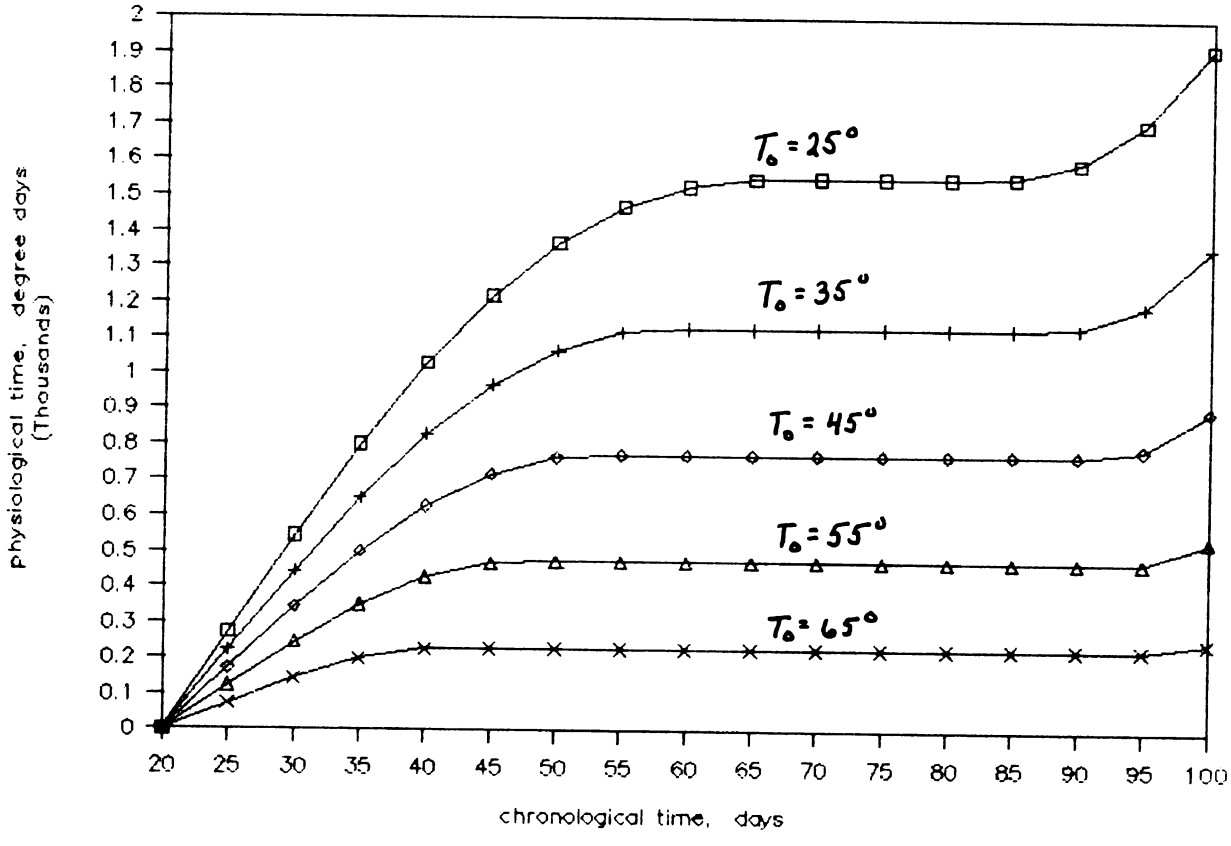


Figure 4.4: A family of physiological time curves for various thresholds.

$t_0 = 20$
$T_0 = 25, 35, 45, 55, 65.$

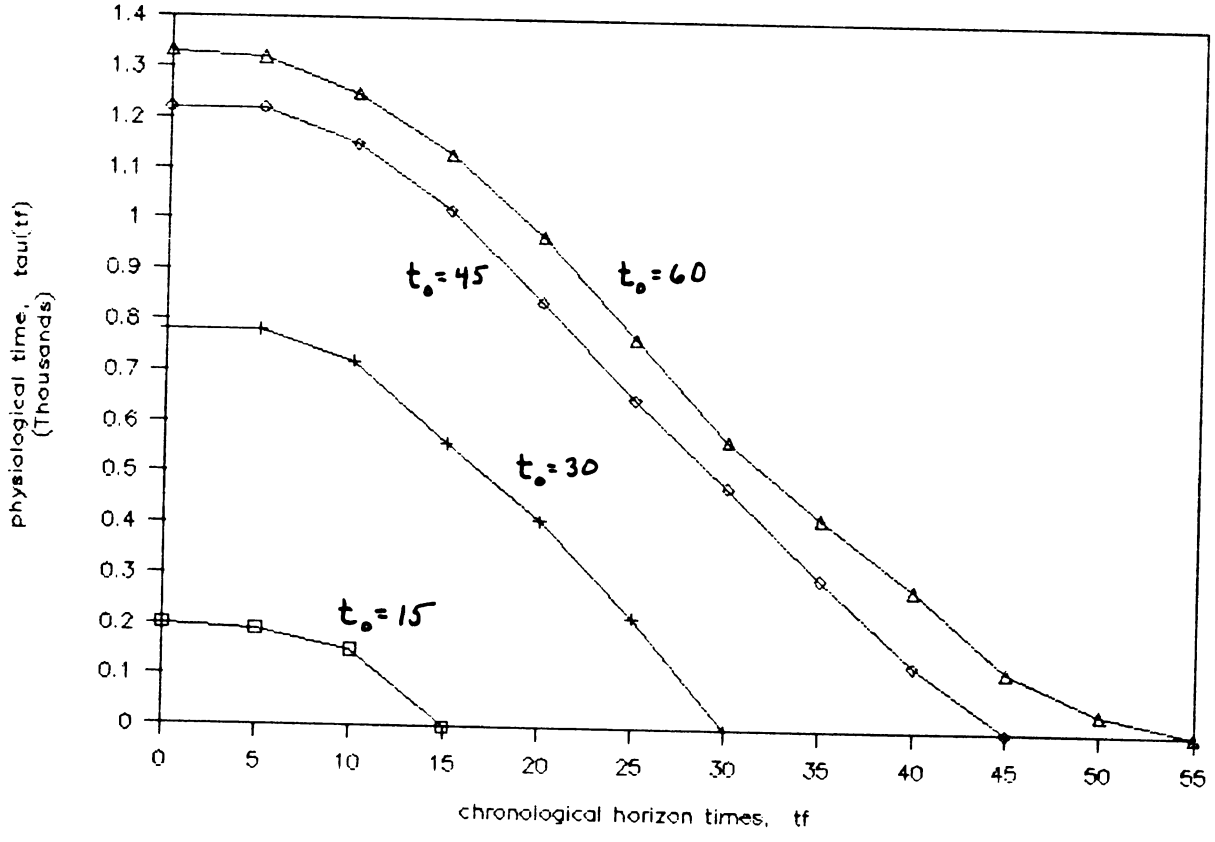


Figure 4.5: **A family of physiological horizon times for various initial times.**

$t_f = 80$
$t_0 = 15, 30, 45, 60.$

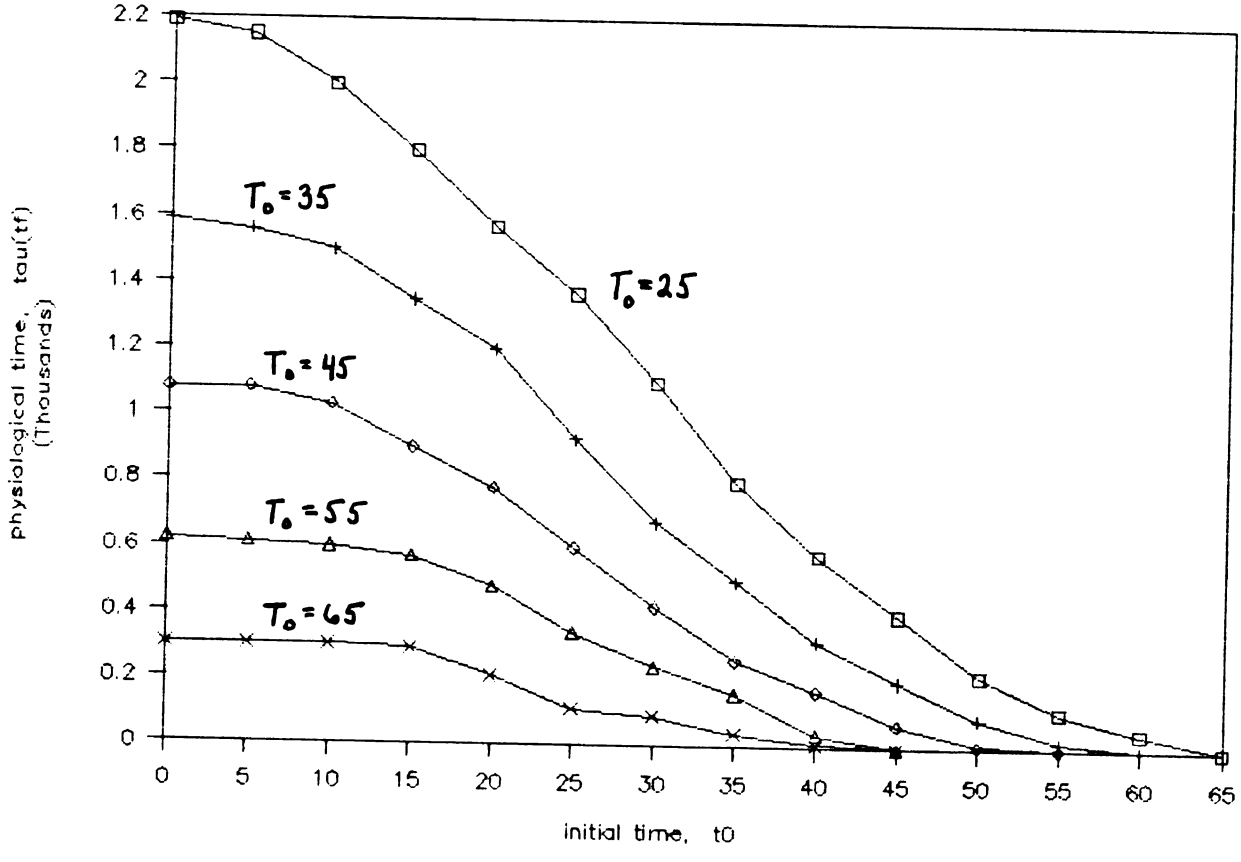


Figure 4.6: A family of physiological horizon times for various thresholds.

$t_f = 80$
$T_0 = 25, 35, 45, 55, 65.$

i. Exponential Growth

Equation [4.9] describes the exponential system in the classical sense. However, by using the physiological time $\tau(t)$ rather than the chronological time t, we now apply equation [4.8]. Therefore, rather than uninhibited exponential growth, there is only growth over the action set $\Gamma(t)$. Figure 4.7 exhibits a family of such "exponential" growth curves.

ii. Logistic Equation

A natural extension of the exponential model is the logistic equation. This is normally applied to accommodate the reality of population density effects, ie. as the density increases, the growth rate declines. Mathematically this is written as...

[4.23]
$$\frac{dx(t)}{dt} = r \cdot x(t) \cdot [1 - x(t)/x_m]$$

$$\text{with} \quad x(t_0) = x_0$$

It is easy to show that the constant x_m (called the **system capacity**) is an equilibrium point and that the curve is a classical sigmoid with $x(t)=x_m$ as an asymptote for $t\to\infty$. The population metric $x(t)$ is bounded on the interval $[x_0,x_m]$. Figure 4.8 illustrates a family of such sigmoids in a classical system.

It is also an easy matter to apply the heat unit hypothesis to a Logistic system. Using Theorem 4.2,

$$[4.24] \qquad \frac{dx(\tau)/d\tau}{x(\tau)\cdot[1-x(\tau)/x_m]} = r, \qquad x(0)=x_0$$

This is easily integrated:

$$[4.25] \qquad \int_{x_0}^{x} \frac{dx}{x(1-x/x_m)} = r\cdot\int_{0}^{\tau} d\tau$$

$$[4.26] \qquad \ln\left(\frac{x}{1-x/x_m}\right) = r\cdot\tau$$

$$[4.27] \qquad x(\tau) = \frac{x_m}{1-(1-x_m/x_0)\cdot e^{-r\tau}}$$

Figures 4.7 and 4.8 illustrate the results of applying equation [4.27]. Using x_0=10, x_m=1000, Figure 4.9a plots x(t) using a threshold of T_0=25^o (unrealistic, but illustrative). Figure 4.9b plots x(t) using a threshold of T_0=45^o (quite typical). Notice that the only growth rate common to both figures, r=0.006, results in a considerably different solution.

It will be useful to compare the classical results against the heat unit results. However, a means for computing the critical biological activity parameters T_0 and r is needed first. This will be done in succeeding chapters.

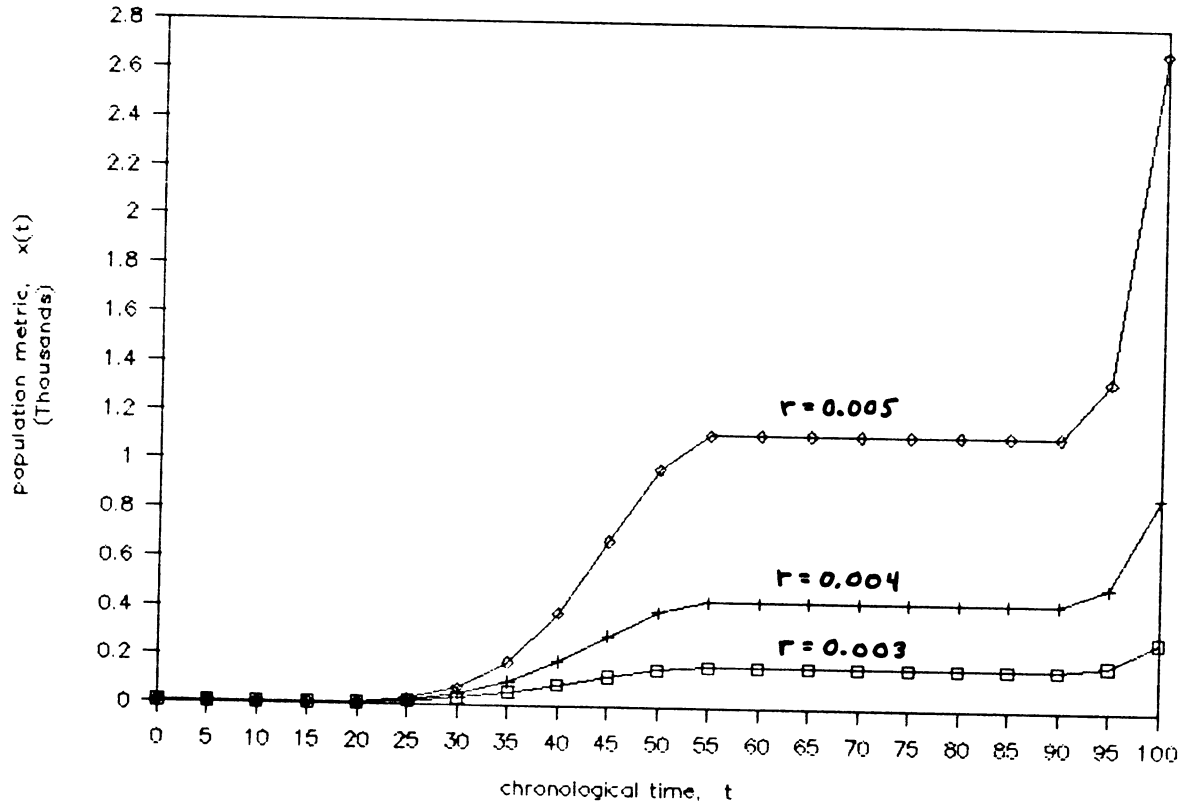


Figure 4.7: **A family of exponential model population curves for various growth rates.**

$T_0 = 45^{\circ}$ F
$r = 0.003, 0.004, 0.005$

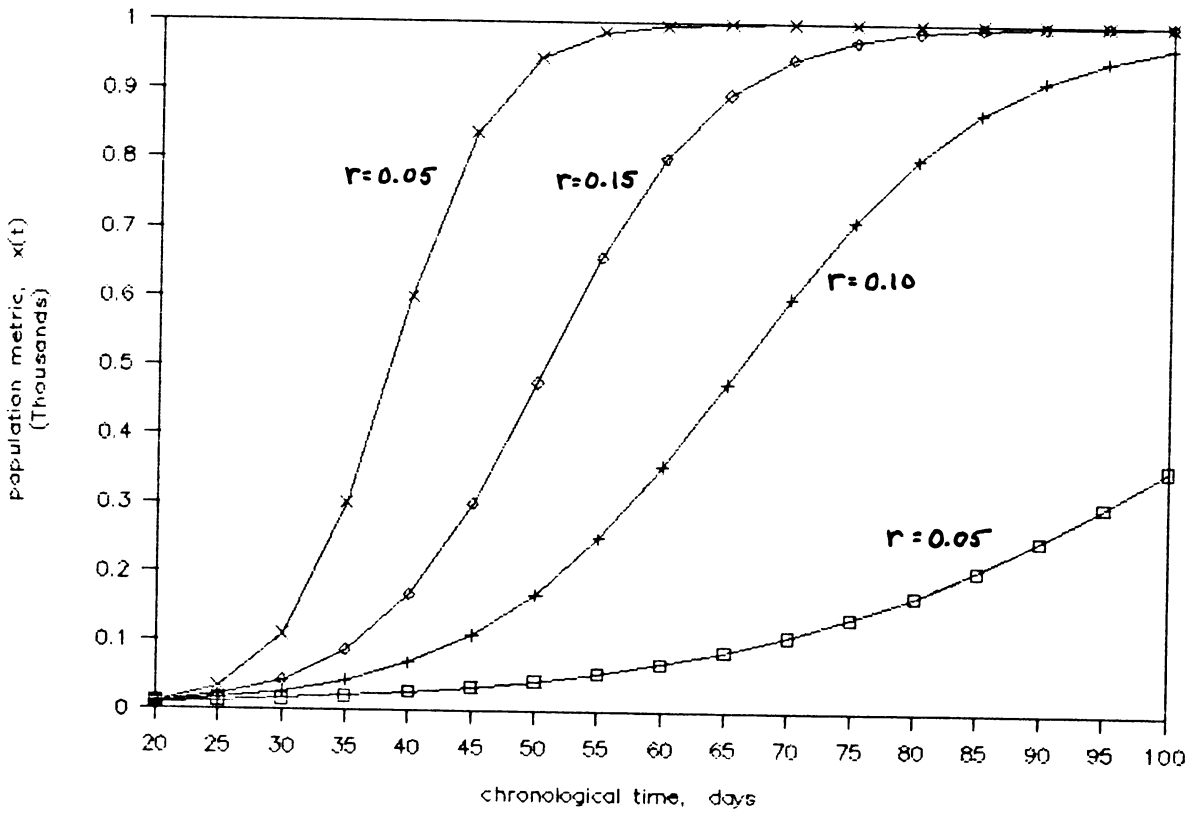


Figure 4.8: A family of classical logistic model population curves for various growth rates.

t_0 = 20 days
r = 0.05, 0.10, 0.15, 0.25

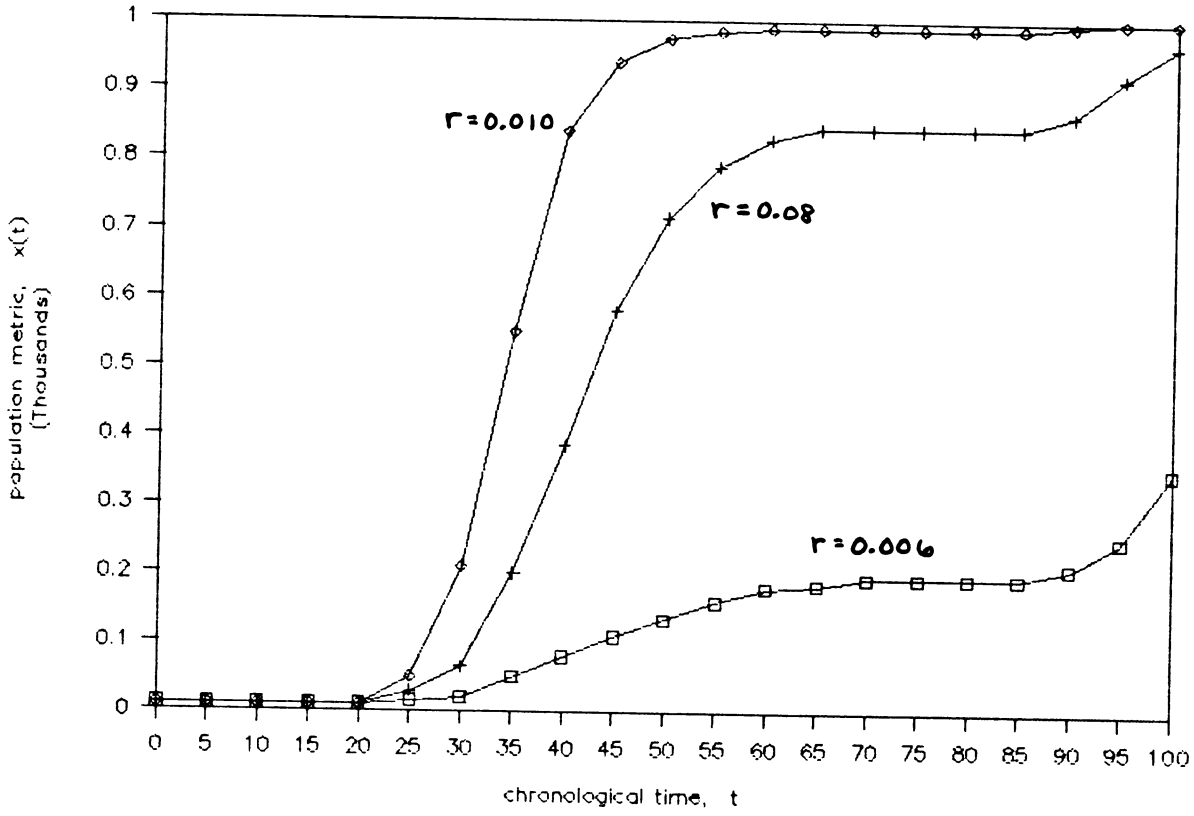


Figure 4.9a: A family of heat-unit based logistic model population curves for various growth rates.

$T_0 = 25^{\circ}$ F
r = 0.006, 0.008, 0.010

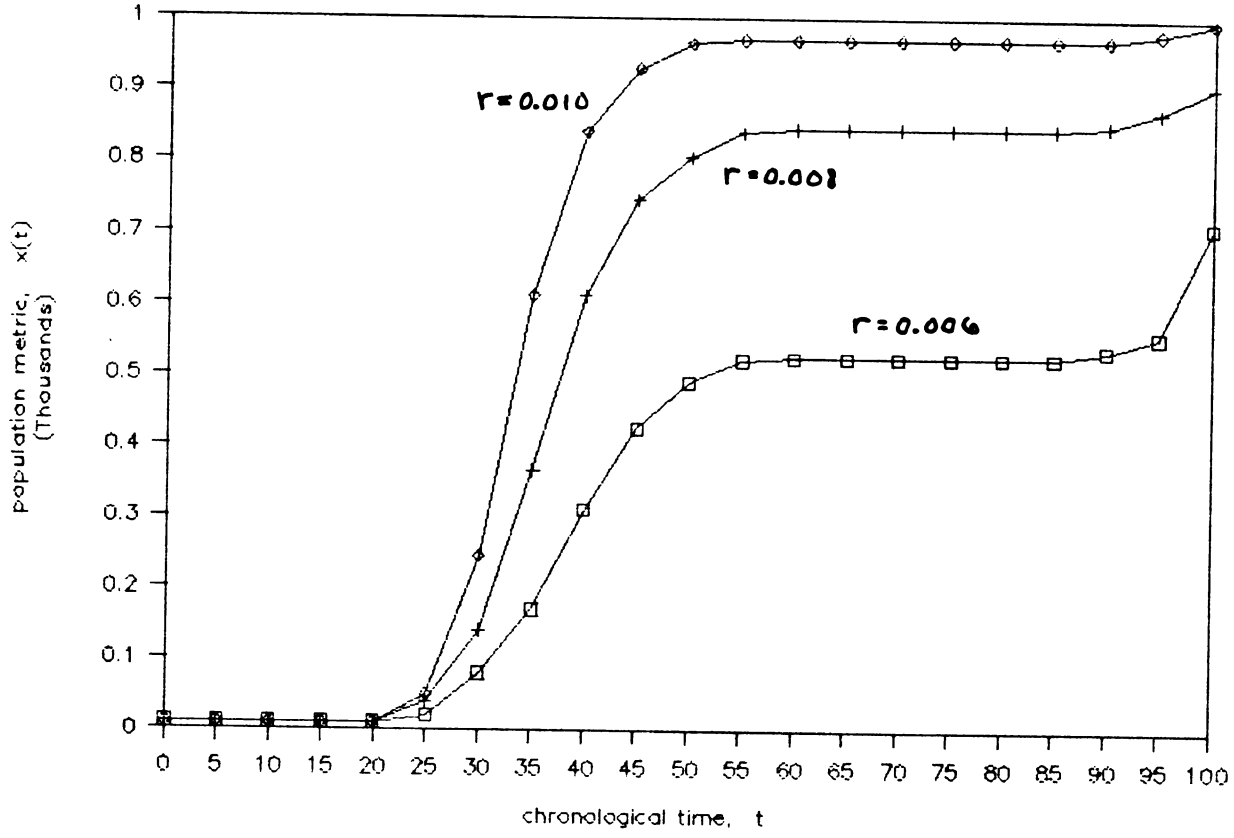


Figure 4.9b: A family of heat-unit based logistic model population curves for various growth rates.

$T_0 = 45^{\circ}$ F
r = 0.006, 0.008, 0.010

NUMERICAL METHODS

In applying the heat unit hypothesis it is first necessary to accurately model the physiological time base. As explained above, this depends totally on the temperature profile. Unless the model is mathematically idealized, the actual application will require numerical methods applied to the empirical data.

There are two basic approaches to the use of empirical data: numerical integration and curve fitting. Since physiological time is defined by an integral (equation [4.10]) and is then applied to a differential equation (equation [4.15]), the standard numerical integration techniques can be applied. On the other hand, curve fitting can lead to an analytical solution. However, using this method first requires parameter estimation.

In the best of worlds, these two approaches would lead to the same results. Unfortunately this is not the case.

EULERIAN METHODS

Consider a current state sensing system of a single species population $x(t)$; t being the chronological time.

$$[4.28] \qquad \frac{d}{dt} x(t) = r \cdot g(x(t)), \qquad x(t_0)=x_0$$

By hypothesis we can rewrite this description using physiological time.

$$[4.29] \qquad \frac{dx(\tau)}{d\tau} = r \cdot g(x(\tau)), \qquad x(0)=x_0$$

$$\text{where} \quad \tau = \int_{\Gamma(t)} [T(\mu)-T_0]d\mu$$

By using a first order approximation for the derivative, we derive an Eulerian type formula with **chronological step-size** δ.

$$[4.30] \qquad \frac{x(\tau(t+\delta)) - x(\tau(t))}{\tau(t+\delta) - \tau(t)} \approx r \cdot g(x(\tau(t)))$$

[4.31] $$x(\tau(t+\delta)) \approx x(\tau(t)) + h(t,\delta)\cdot r\cdot g(\tau(t))$$

$$\text{where} \quad h(t,\delta) = \tau(t+\delta) - \tau(t)$$

Unlike the traditional Euler formulas, the **physiological step size** $h(t,\delta)$ is a function of the **chronological step size** δ and t! This is because we are actually applying our method to the composite function, $x(\tau(t))$.

Let us consider step size δ to be a predetermined constant. This, in turn, determines $h(t,\delta)$. Since $x(\tau(t))$ is then a composite function, the nature of $h(t,\delta)$ must be understood in order that convergence of any numerical procedure can be guaranteed. This is done as follows:

Let us define a difference equation analogous to equation [4.15]. Let $\{y_k\}$ be such a sequence. Then by discretizing the chronological time,

[4.32]
$$y_k = x(\tau(t_k))$$
$$t_k = t_0+\delta\cdot k, \qquad k=0,1,\ldots,n$$

Equation [4.31] reduces to

$$[4.33] \qquad y_{k+1} = y_k + r \cdot h(t,\delta) \cdot g(y_k)$$

with discretization error:

$$[4.34] \qquad e_k = x(\tau(t_k)) - y_k$$

Expanding $x(\tau(t_{k+1}))$ about $t=t_k$ by Taylor's theorem,

$$[4.35] \qquad x(\tau(t_{k+1})) = x(\tau(t_k)) + h(t,\delta) \cdot Dx(\tau(t_k)) + \tfrac{1}{2}h^2(t,\delta) \cdot D^2x(\tau(\theta_k))$$

$$\text{where} \quad \tau(t_k) < \tau(\theta_k) < \tau(t_{k+1})$$

Since the quantity $\tfrac{1}{2}h^2(t,\delta) \cdot D^2x(\tau(\theta_k))$ is the error encountered at a single step (from t_k to t_{k+1}), it is referred to as the **local error.**

Combine [4.17] and [4.18]:

$$[4.36] \qquad e_{k+1} = x(\tau(t_{k+1})) - x(\tau(t_k))$$

$$= e_k + r \cdot h(t,\delta) \cdot [g(\tau(t_k)) - g(y_k)] + \tfrac{1}{2}h^2(t,\delta) \cdot D^2x(\tau(\theta_k))$$

By the Mean Value Theorem,

$$[4.37] \qquad g(\tau(t_k)) - g(y_k) = D_\tau g(\tau(\underline{\theta}_k))\cdot[x(\tau(t_k)) - y_k]$$

$$= D_\tau g(\tau(\underline{\theta}_k))\cdot e_k$$

where $\underline{\theta}_k$ is such that $x(\tau(t_k))<\tau(\underline{\theta}_k))<y_k$

Therefore,

$$[4.38] \qquad e_{k+1} = e_k\cdot\left[1 + h(t,\delta)\cdot D_\tau(\tau(\underline{\theta}_k)) + \tfrac{1}{2}h^2(t,d)\cdot D^2x(\tau(\theta_k))\right]$$

We consider that our systems are fairly well behaved and that the physiological step size is bounded. Thus,

$$[4.39] \qquad \begin{aligned} &|D^2x(\theta_k)|<\alpha \\ &|D_\tau g(\tau(t))|<\beta \\ &|h(t,\delta)|<h \end{aligned}$$

over the interval $t_0 \le t \le t_n$.

Applying the triangle inequality,

$$[4.40] \qquad |e_{k+1}| \le (1+h\cdot\beta)\cdot|e_k| + \tfrac{1}{2}h^2\cdot\alpha$$

It follows that the difference equation

$$[4.41] \qquad \Theta_{k+1} = (1+h\cdot\beta)\cdot\Theta_k + \tfrac{1}{2}h^2.\alpha, \qquad \Theta_0=0$$

dominates [4.31], ie. $\Theta_k \ge |e_k|$ for all k. The proof of this follows by induction...

Assuming $e_0 = \Theta_0 = 0$;

further, $\Theta_k \ge |e_k|, \qquad k=0,1,...$

$$\Theta_{k+1} = (1+h\cdot\beta)\cdot\Theta_k + \tfrac{1}{2}h^2\cdot\alpha$$

$$[4.42] \qquad \ge \Theta_k + \tfrac{1}{2}h^2\cdot\alpha$$

$$\ge |e_{k+1}| \qquad \text{as required.}$$

An explicit solution of the difference equation [4.24] is given by:

$$[4.43] \qquad \Theta_k = (1+h\cdot\beta)^k + \tfrac{1}{2}h\cdot\alpha/\beta$$

$$= (1+h\cdot\beta)^{(tk-t0)/\delta} + \tfrac{1}{2}h\cdot\alpha/\beta$$

Thus, we have the shown that the local error is bounded:

$$[4.44] \qquad |e_k| \le (1+h\cdot\beta)^{(tk-t0)/\delta} + \tfrac{1}{2}h\cdot\alpha/\beta$$

By this result, we are guaranteed that the Eulerian method converges to the solution of equation [4.33] as desired. Unfortunately this bound is rather crude. While it does suffice to establish the convergence it does not provide a useful error estimate.

EXAMPLE (cont'd)

We will now numerically integrate the logistic equation using the physiological time base calculated in the earlier part of this example. Let us use a growth rate of $r=0.008$ and an equilibrium value of $x_m=1000$.

Numerically this can be done according to Algorithm 4.3. This has been implemented for differing chronological step sizes $\delta=1, 2, 4, 5, 8, 10$ and thresholds $T_0=35, 40, 50, 55$. The results are summarized in Tables 4.1 and 4.2.

Note that the error increases with step size as expected. This is also true for the relative error. Even using a step size of $\delta=1$ day, this error can be considerable. It is also interesting that the error varies with threshold T_0. In each case we see the error rise and fall as T_0 increases. This is illustrated in more detail in Figures 4.10 and 4.15 where both absolute and relative errors are plotted against threshold T_0.

Clearly the Eulerian approach gives only a rough solution estimate. This is not surprising since the local error of Euler is of the first order in step size. Since the physiological step size varies with t and δ this can be highly significant during certain critical times during the species' growth cycle.

$$
\begin{bmatrix}
Y_0 = x_0 \\
\\
\text{DO } k=0,-1+360/\delta \begin{bmatrix}
t_k = t_0 + \delta \cdot k \\
h(t,\delta) = \tau(t+\delta) - \tau(t) \\
Y_{k+1} = Y_k + r \cdot h(t,\delta) \cdot x(\tau(t_k)) \\
\qquad \cdot [1 - x(\tau(t_k))/x_m]
\end{bmatrix}
\end{bmatrix}
$$

Algorithm 4.3 Numerically integrating the logistic population model using an Eulerian approximation.

Table 4.1: Absolute error obtained by solving a population model using the Euler method.

T_0	δ=1	δ=2	δ=4	δ=5	δ=8	δ=10
35	21	44	101	140	297	387
40	34	74	179	239	405	489
45	60	127	258	315	444	502
50	70	133	229	265	342	377
55	42	74	121	138	176	194

Table 4.2: Relative error obtained by solving a population model using the Euler method.

T_0	δ=1	δ=2	δ=4	δ=5	δ=8	δ=10
35	.05	.10	.18	.22	.49	.48
40	.07	.13	.27	.34	.52	.61
45	.11	.21	.39	.47	.63	.70
50	.15	.27	.46	.53	.67	.72
55	.15	.27	.44	.50	.62	.67

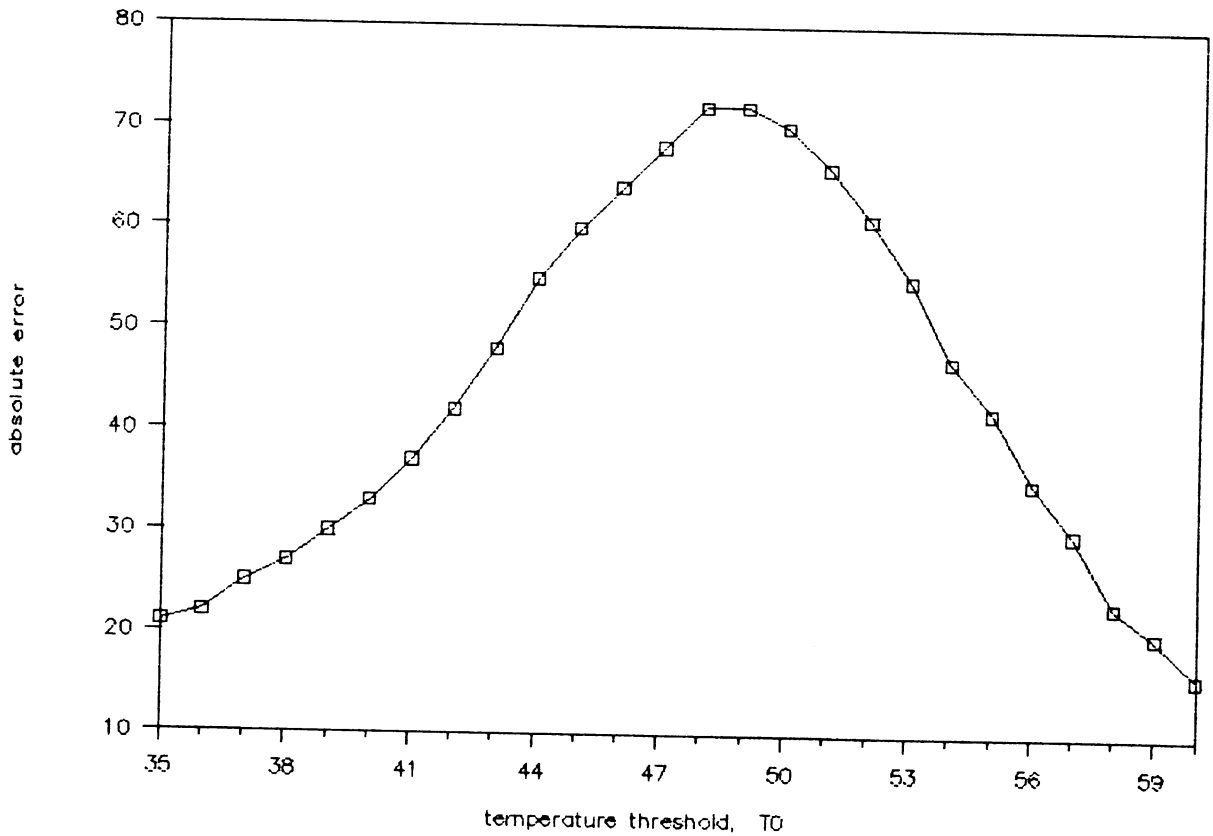


Figure 4.10: The absolute error in a population model using a Eulerian approximation as a function of threshold. δ=1.

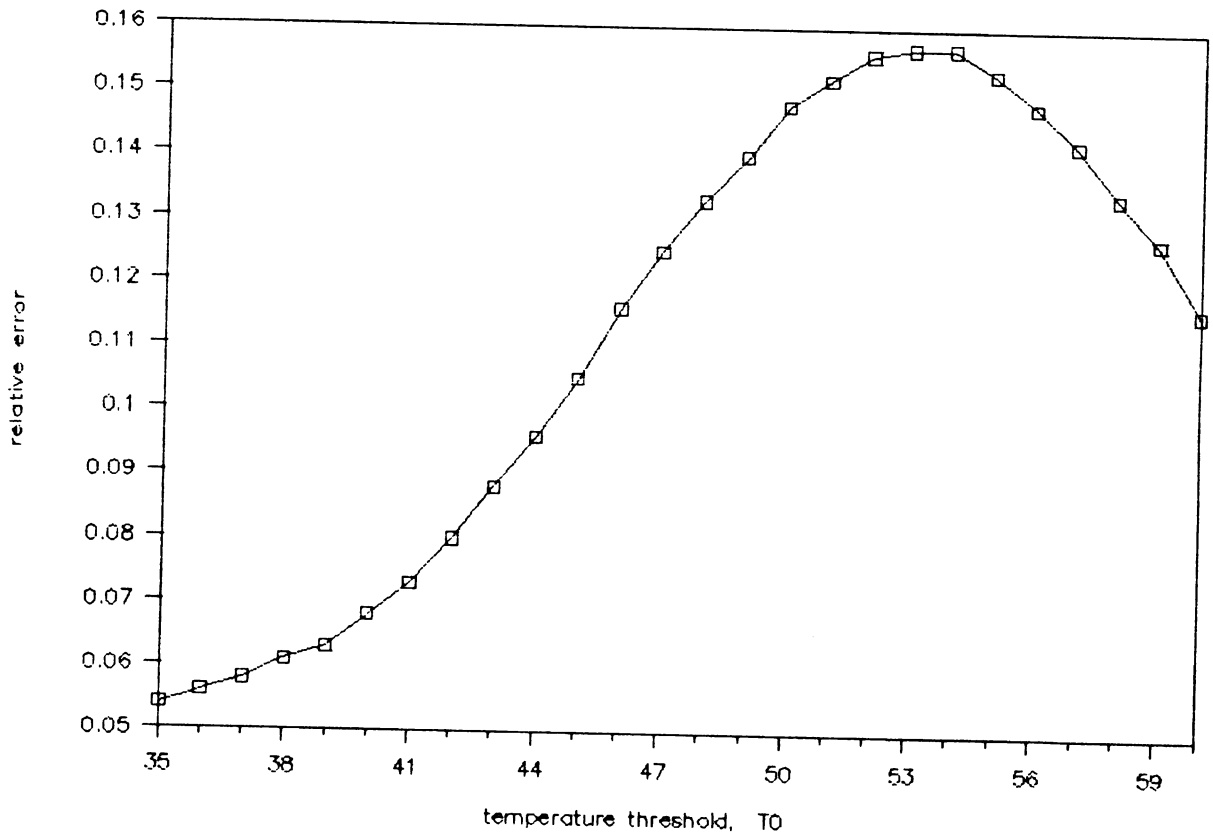


Figure 4.11: The relative error in a population model using a Eulerian approximation as a function of threshold. $\delta=1$.

RUNGE-KUTTA METHODS

A set of popular alternatives to Euler are the Runge-Kutta methods. Typically we use those of fourth order and thus the error should be reduced considerably. This method can be applied to the situation at hand as follows.

Again consider the system [4.29]:

$$\frac{d}{d\tau} x(\tau) = r \cdot g(x(\tau)), \qquad x(0)=x_0$$

$$\tau = \int_{\Gamma(t)} [T(\mu)-T_0]d\mu$$

The Runge-Kutta procedure is given as Algorithm 4.4. This iterative formulation is very similar to the traditional Runge-Kutta but has been applied to the composite function $\tau(t)$. Since $\tau(t)$ is as a rule a most friendly function, this algorithm presents no problem in practice.

Table 4.3 summarizes the results of a series of Runge-Kutta runs applied to the Logistic model defined in Example One. We notice the great improvement in accuracy, even for larger step size δ. Again there is a (so be it less so)

pronounced dependency on the threshold T_0. This is illustrated in Figure 4.12.

In summary, numerical integration techniques can lead to fairly accurate solutions using Runge-Kutta. For an arbitrary temperature profile, however, Euler integration tends to be quite inaccurate, even for the smallest step size δ=1 days. Accuracy depends on the threshold T_0 as well as step size δ. Thus, models using differing thresholds (as in multi-species populations) will have additional relative inaccuracies built in.

While Runge-Kutta techniques require additional computation time (four function calls rather than one), the fourth order error is excellent.

$$
\begin{array}{l}
\text{DO } k=0,n \\
\quad t_k = t_0+\delta\cdot k \\
\quad h_k = \tau(t_k+\delta)-\tau(t_k) \\
\quad\quad K_1 = r\cdot h_k\cdot g(\tau(t_k)) \\
\quad\quad K_2 = r\cdot h_k\cdot g(\tau(t_k)+\tfrac{1}{2}K_1) \\
\quad\quad K_3 = r\cdot h_k\cdot g(\tau(t_k)+\tfrac{1}{2}K_2) \\
\quad\quad K_2 = r\cdot h_k\cdot g(\tau(t_k)+K_3) \\
\quad Y_{k+1} = Y_k + (K_1+2\cdot K_2+2\cdot K_3+K_4)/6
\end{array}
$$

Algorithm 4.4 Numerically integrating a generic population model (expressed as equation [4.16]) using the Runge-Kutta approximation.

Table 4.3: Absolute error obtained by solving a population model using the Runge-Kutta method.

T_0	δ=1	δ=2	δ=4	δ=5	δ=8	δ=10
35	.012	.159	1.724	3.601	18.279	28.745
40	.012	.162	1.747	3.565	14.993	28.275
45	.013	.177	1.954	3.990	15.852	29.606
50	.010	.138	1.587	3.291	13.423	24.472
55	.004	.053	.638	1.354	5.866	10.935

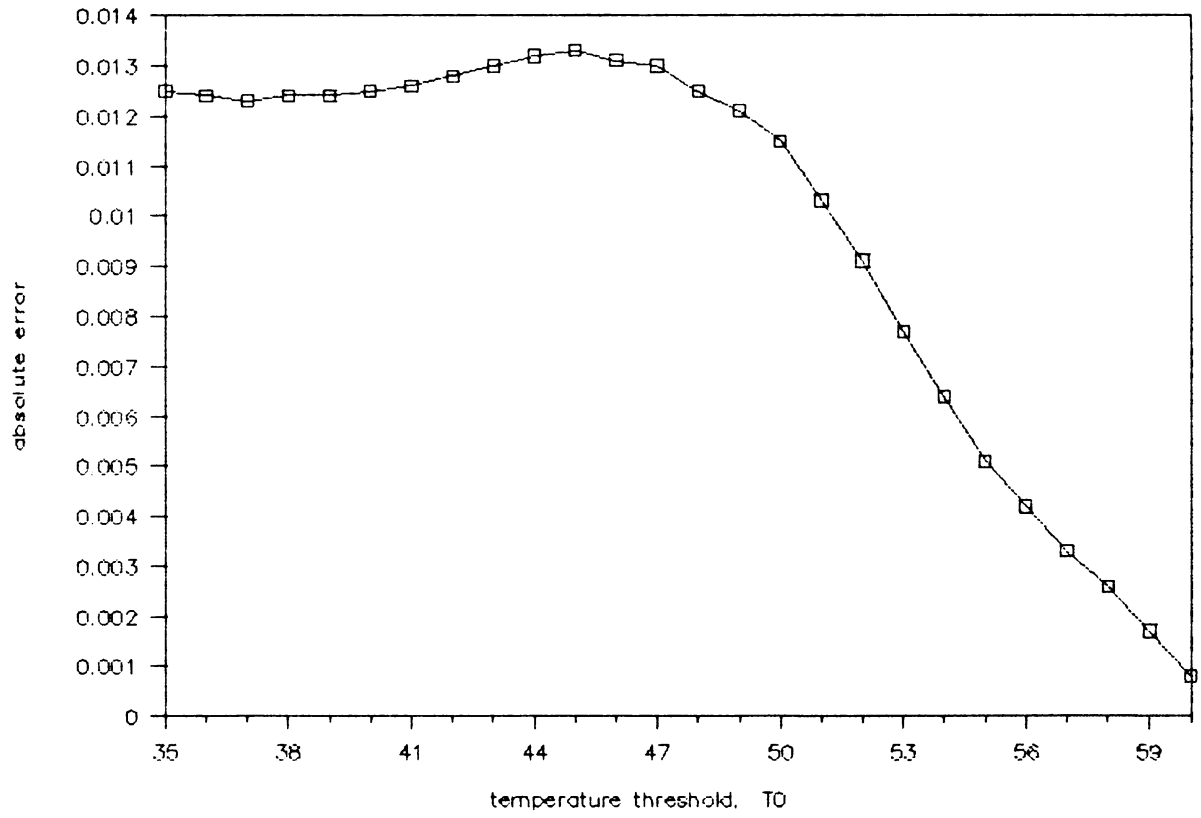


Figure 4.12: **The absolute error of a population model using a Runge-Kutta approximation as a function of threshold. δ=1.**

CURVE FITTING METHODS

Curve fitting is a popular device when the mathematical form of the temperature profile is assumed a'priori. By hypothesizing an aesthetically pleasing form, one can obtain an equation that is often analytically soluble. In addition to the elegance thus achieved, quick and easily comprehendible results are sometimes achieved.

EXAMPLE TWO

Consider the mean temperature data recorded for Lansing, Michigan during 1985 as graphed in Figure 4.13. Ignoring the apparent noise component, intuition implies a sinusoidal carrier with period of p=365 days.

Thus, we assume a temperature profile of

$$[4.46] \qquad T(t) = a + b\cdot\sin(wt) + c\cdot\cos(wt)$$

$$\text{for } w = 2\pi/365 \approx 0.0172142$$

The parameters a,b,c are yet to be determined. To do this, we will use a Least Squares approach. Let (t_k,T_k), $k=1,2,\ldots,p$ be data points. Then we define an objective function $f(a,b,c)$ as follows:

$$[4.47] \qquad f(a,b,c) = \sum_{k=1}^{p} [T_k-a-b\cdot\sin(wt_k)-c\cdot\cos(wt_k)]^2$$

Function $f(a,b,c)$ is to be minimized for a certain choice of parameters a,b,c. This is done using elementary calculus. Differentiation yields:

$$\sum_{k=1}^{p} [T_k - a - b\cdot\cos(wt_k) - c\cdot\sin(wt_i)] = 0$$

$$\sum_{k=1}^{p} [T_k - a - b\cdot\cos(wt_k) - c\cdot\sin(wt_i)]\cdot\sin(wt_k) = 0 \qquad [4.48]$$

$$\sum_{k=1}^{p} [T_k - a - b\cdot\cos(wt_k) - c\cdot\sin(wt_i)]\cdot\cos(wt_k) = 0$$

This can be put in linear form as...

$$a\cdot p\cdot\sin(wt_k) + b\cdot\sum_{k=1}^{p}\sin(wt_k) + c\cdot\sum_{k=1}^{p}\cos(wt_k) = \sum_{k=1}^{p}\tau_k$$

$$a\cdot\sum_{k=1}^{p}\sin(wt_k) + b\cdot\sum_{k=1}^{p}\sin^2(wt_k) + c\cdot\sum_{k=1}^{p}\cos(wt_k)\cdot\sin(wt_k) = \sum_{k=1}^{p}\tau_k\cdot\sin(wt_k) \qquad [4.49]$$

$$a\cdot\sum_{k=1}^{p}\cos(wt_k) + b\cdot\sum_{k=1}^{p}\sin(wt_k)\cdot\cos(wt_k) + c\cdot\sum_{k=1}^{p}\cos^2(wt_k) = \sum_{k=1}^{p}\tau_k\cdot\sin(wt_k)$$

This is best summarized in matrix form, as is shown in equation [4.50]:

$$\begin{bmatrix} p & \sum_{k=1}^{p} \sin(wt_k) & \sum_{k=1}^{p} \cos(wt_k) \\ \sum_{k=1}^{p} \sin(wt_k) & \sum_{k=1}^{p} \sin^2(wt_k) & \sum_{k=1}^{p} \cos(wt_k)\cdot\sin(wt_k) \\ \sum_{k=1}^{p} \cos(wt_k) & \sum_{k=1}^{p} \cos(wt_k)\cdot\sin(wt_k) & \sum_{k=1}^{p} \cos^2(wt_k) \end{bmatrix}$$

$$\cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{p} T_k \\ \sum_{k=1}^{p} T_k\cdot\sin(wt_k) \\ \sum_{k=1}^{p} T_k\cdot\cos(wt_k) \end{bmatrix} \qquad [4.50]$$

We are encouraged upon computing the constants; that the off-diagonal elements are very nearly zero. That is to say, since they nearly sum to zero, the data are very close to sinusoidal in nature.

It follows that a=47.0315, b=-7.54652, c=-24.9468. Figure 4.13 is a sketch of the resulting profile T=T(t).

It will be convenient to express this equation in the mathematically equivalent form:

[4.51] $$T(t) = a + d\cdot\cos(wt+\theta)$$

with $a = 47.0315$

$d = -26.06$

$\theta = -0.293753$

$w = 0.0102142$

Now,

[4.52] $$\tau(t) = \int_{\Gamma(t)} [T(\mu)-T_0]d\mu$$

$$= \int_{t_1}^{t_2} [a + d\cdot\cos(w\mu+\theta) - T_0]d\mu$$

$$= \frac{d}{w}\cdot\left[\sin(wt_2+\theta) - \sin(wt_1+\theta)\right]$$

$$+ (d-T_0)\cdot(t_2-t_1)$$

where $$t_1 = \frac{1}{w}\cdot\left[\cos^{-1}[(t_0-a)/d] - \theta\right]$$

$$t_2 = \frac{1}{w}\cdot\left[2\pi - \cos^{-1}[(t_0-a)/d] - \theta\right]$$

Therefore,

$$\tau(t) = \frac{d}{w}\cdot\left[\sin(wt_2+\theta) - \sin(wt_1+\theta)\right]$$

$$+ \frac{2}{w}\cdot(d-T_0)\cdot\left[\pi - \cos^{-1}[(t_0-a)/d]\right]$$

This function is displayed in Figure 4.14 for threshold $T_0=40^{\circ}$. As a comparison, the actual physiological time is also plotted.

Immediate problems come to light. Figure 4.14 shows clearly that the curve-fitted formulation lags significantly behind the empirical data. Even worse, as soon as it catches up, the curve-fitted result stops and the total degree-days for the two methods finish significantly different!

Figure 4.15 compares the two methods at the chronological time t=180 days. Obviously there is a significant relationship. As the threshold T_0 increases, the two methods are much the same. However, for low and moderate thresholds, the error is significant.

However, the real test of compatibility is how the system solution itself, not the time base, is affected. This is illustrated, using a threshold of $T_0=40^{\circ}$, in Figure 4.16.

Here we see that the early cut-off of the curve-fitting technique has but a relatively small effect. However, there is a highly significant difference at intermediate time values, as shown in Figures 4.22. The curve-fitted solution will result in considerably less growth.

It seems apparent that the curve-fitted (via idealized modeling and parameter estimation) solution produces inferior results. This and the numerical methods are produced in summary for various step sizes in Table 4.4. Here, as in the previous examples, $x_0=10$ and $x_m=1000$.

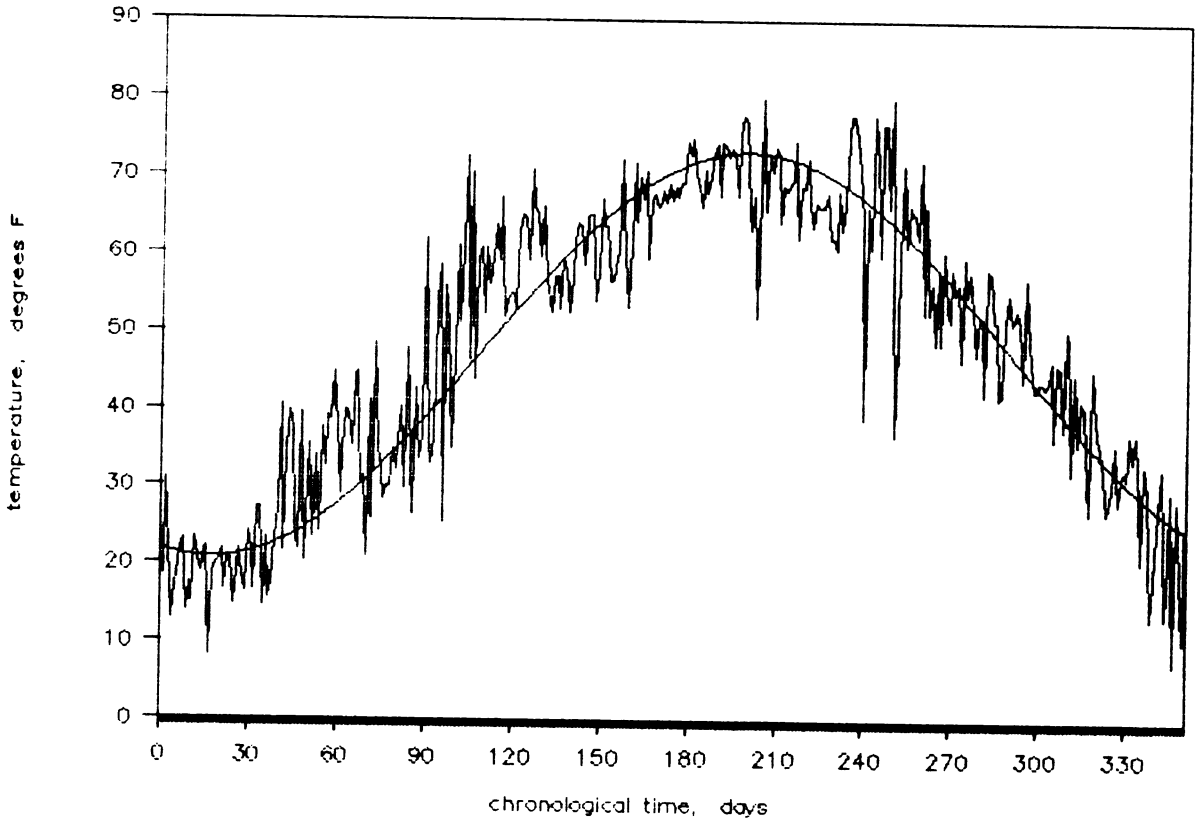


Figure 4.13: The empirical temperature profile for Lansing, Michigan (1985) superimposed on the sinusoidal curve fitting approximation.

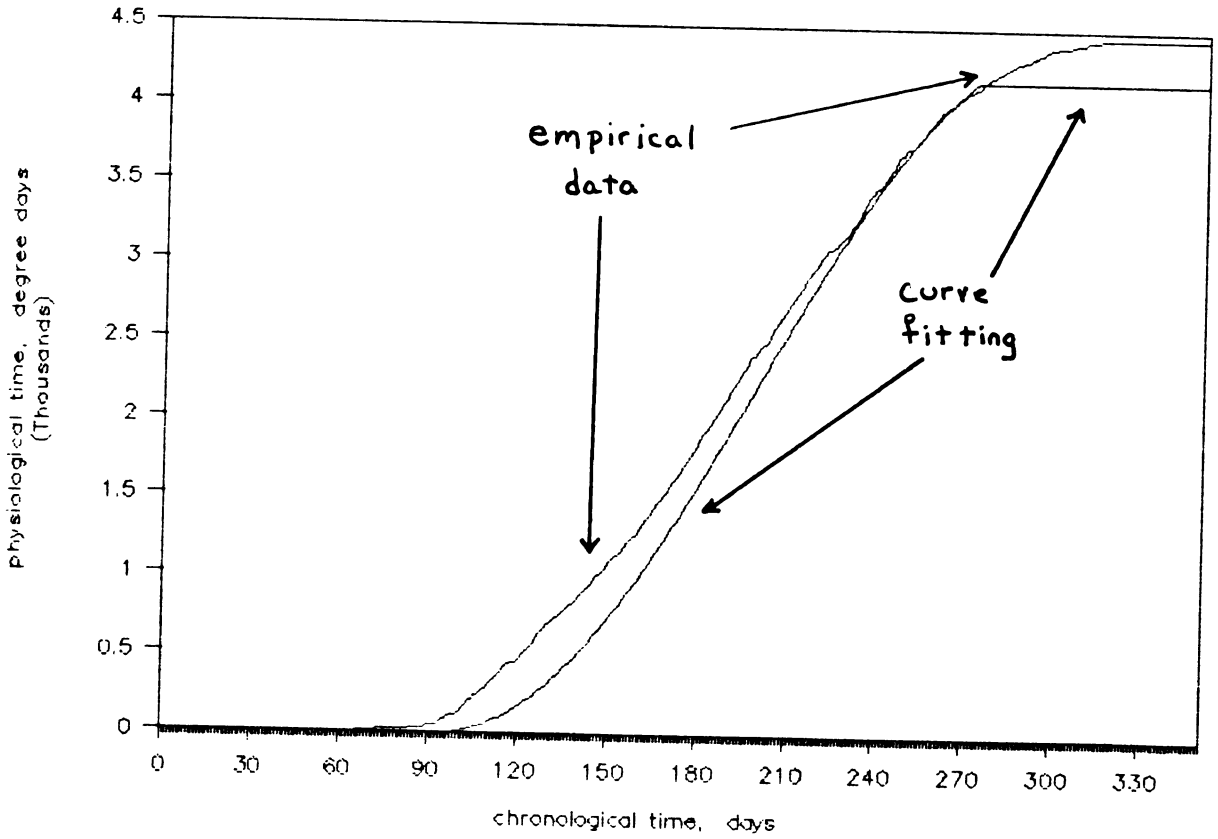


Figure 4.14: Physiological time as computed using empirical data and a sinusoidal curve fitting approximation.

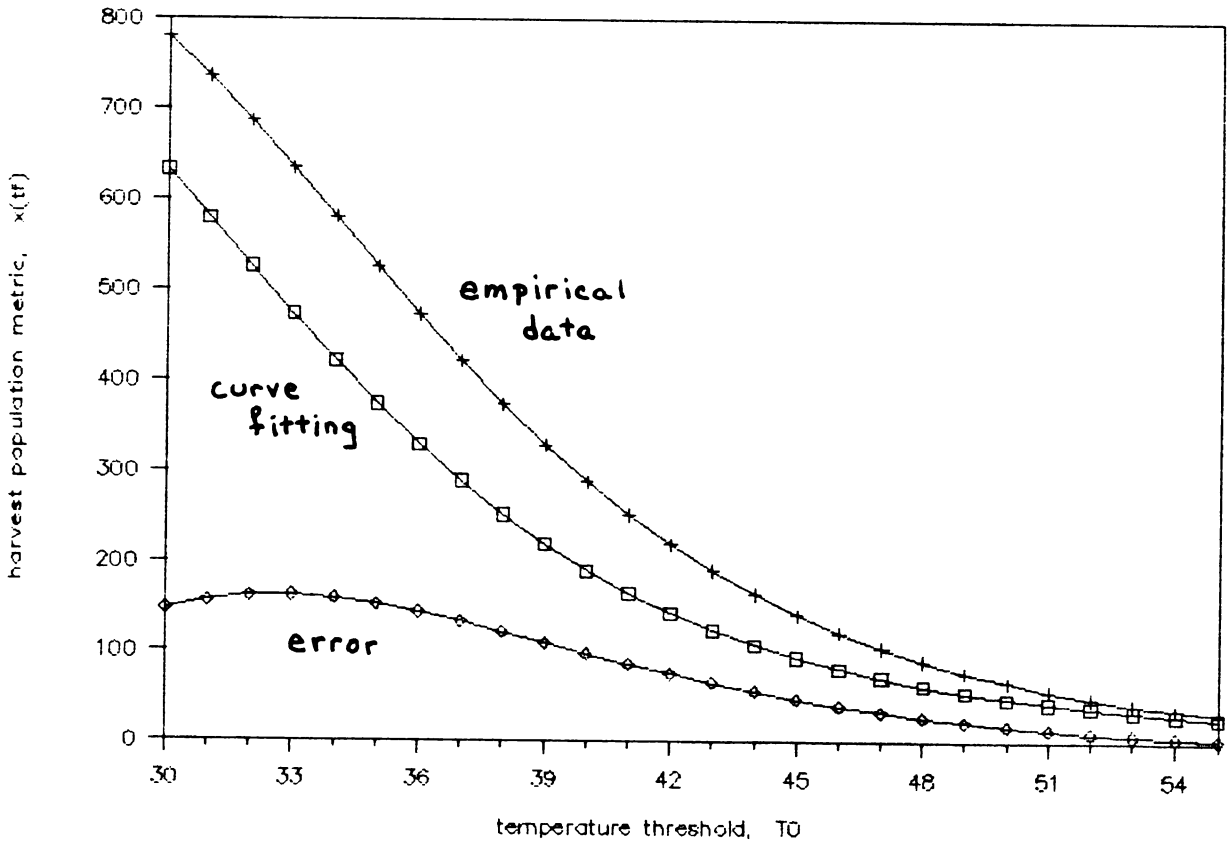


Figure 4.15: Population size at horizon (t_f=300 days) using an empirical temperature profile and a sinusoidal curve fitting profile as a function of threshold.

Table 4.4: Comparison of integration techniques of the logistic model.

		.. Empirical Data ..			... Curve Fitting ...		
delta	day	Form	Euler	R.K.	Form	Euler	R.K.
1	0	10.00	10.00	10.00	10.00	10.00	10.00
	35	10.00	10.00	10.00	10.00	10.00	10.00
	70	10.00	10.00	10.00	10.29	10.29	10.29
	105	10.70	10.70	10.70	15.30	15.20	15.30
	140	26.57	26.24	26.57	52.88	51.32	52.88
	175	146.30	139.73	146.30	229.25	218.43	229.25
	210	624.95	607.02	624.95	702.86	688.14	702.86
	245	930.51	927.98	930.51	934.37	932.28	934.37
	280	976.86	976.49	976.86	979.37	979.23	979.37
	315	976.86	976.49	976.86	986.58	986.55	986.58
	350	976.86	976.49	976.86	986.72	986.69	986.72
2	0	10.00	10.00	10.00	10.00	10.00	10.00
	35	10.00	10.00	10.00	10.00	10.00	10.00
	70	10.00	10.00	10.00	10.29	10.29	10.29
	105	10.83	10.82	10.83	16.25	16.04	16.25
	140	26.57	25.92	26.57	52.88	50.16	52.88
	175	154.18	140.66	154.18	239.06	218.39	239.06
	210	624.95	588.44	624.95	702.86	674.12	702.86
	245	933.72	928.76	933.72	938.73	935.12	938.73
	280	976.86	976.07	976.86	979.37	978.94	979.37
	315	976.86	976.07	976.86	986.58	986.40	986.58
	350	976.86	976.07	976.86	986.72	986.55	986.72
5	0	10.00	10.00	10.00	10.00	10.00	10.00
	35	10.00	10.00	10.00	10.00	10.00	10.00
	70	10.00	10.00	10.00	10.29	10.29	10.29
	105	10.70	10.69	10.70	15.30	15.00	15.30
	140	26.57	25.05	26.57	52.88	47.04	52.88
	175	146.30	118.13	146.30	229.25	185.61	229.24
	210	624.95	530.20	624.93	702.86	627.95	702.84
	245	930.51	914.64	930.51	934.37	922.35	934.37
	280	976.86	974.40	976.86	979.37	977.95	979.37
	315	976.86	974.40	976.86	986.58	985.91	986.58
	350	976.86	974.40	976.86	986.72	986.06	986.72

Table 4.4 (cont'd)

		.. Empirical Data ..			... Curve Fitting ...		
delta	day	Form	Euler	R.K.	Form	Euler	R.K.
7	0	10.00	10.00	10.00	10.00	10.00	10.00
	35	10.00	10.00	10.00	10.00	10.00	10.00
	70	10.00	10.00	10.00	10.29	10.29	10.29
	105	10.70	10.69	10.70	15.30	14.93	15.30
	140	26.57	24.54	26.57	52.88	45.42	52.88
	175	146.30	109.61	146.28	229.25	172.65	229.23
	210	624.95	490.95	624.88	702.86	595.79	702.82
	245	930.51	905.36	930.48	934.37	916.50	934.35
	280	976.86	973.03	976.85	979.37	977.49	979.36
	315	976.86	973.03	976.85	986.58	985.67	986.57
	350	976.86	973.03	976.85	986.72	985.82	986.71
10	0	10.00	10.00	10.00	10.00	10.00	10.00
	35	10.00	10.00	10.00	10.00	10.00	10.00
	70	10.00	10.00	10.00	10.29	10.29	10.29
	105	11.43	11.37	11.43	18.01	17.00	18.01
	140	26.57	23.85	26.57	52.88	42.82	52.87
	175	189.72	123.96	189.61	287.65	190.79	287.53
	210	624.95	434.53	624.69	702.86	542.17	702.70
	245	944.88	914.45	944.76	946.98	925.28	946.90
	280	976.86	969.70	976.80	979.37	975.63	979.34
	315	976.86	969.70	976.80	986.72	984.81	986.70
	350	976.86	969.70	976.80	986.72	984.81	986.70

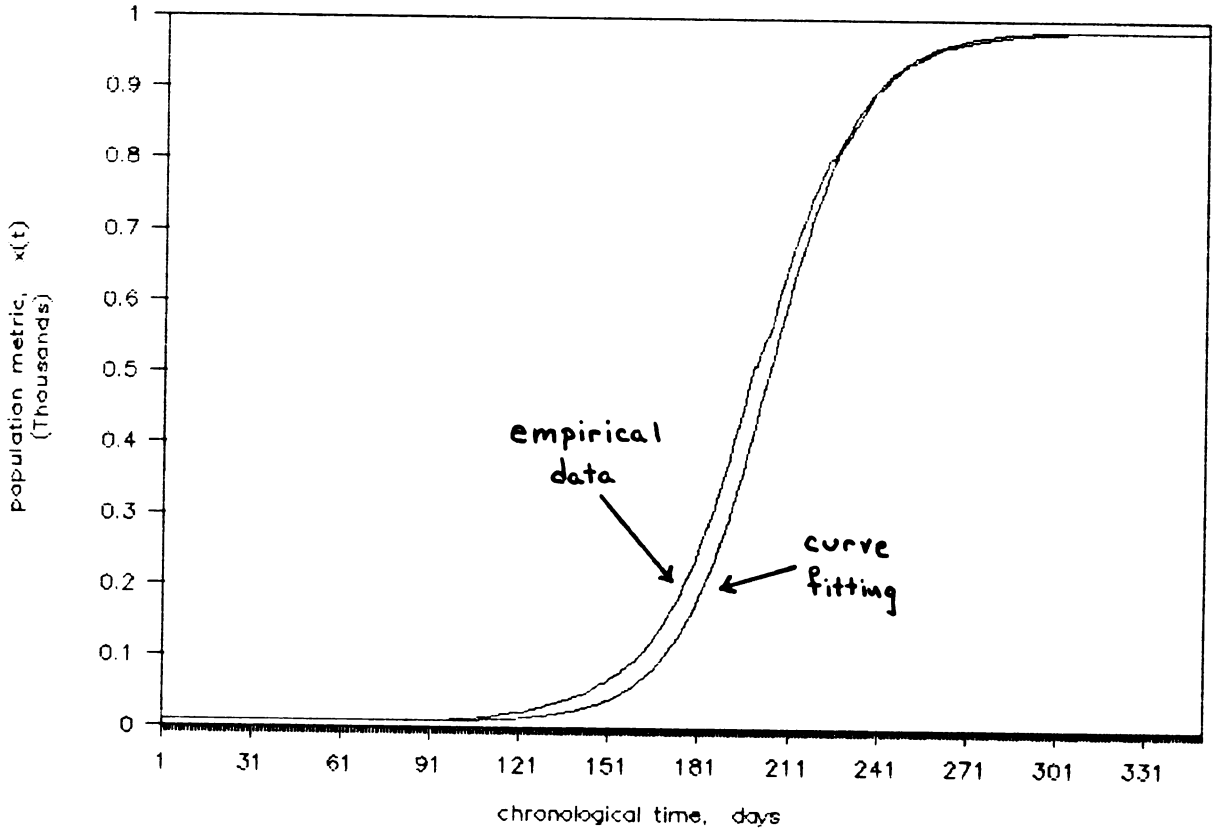


Figure 4.16: Population metric as computed using empirical temperature profile and a sinusoidal curve fitting approximation.

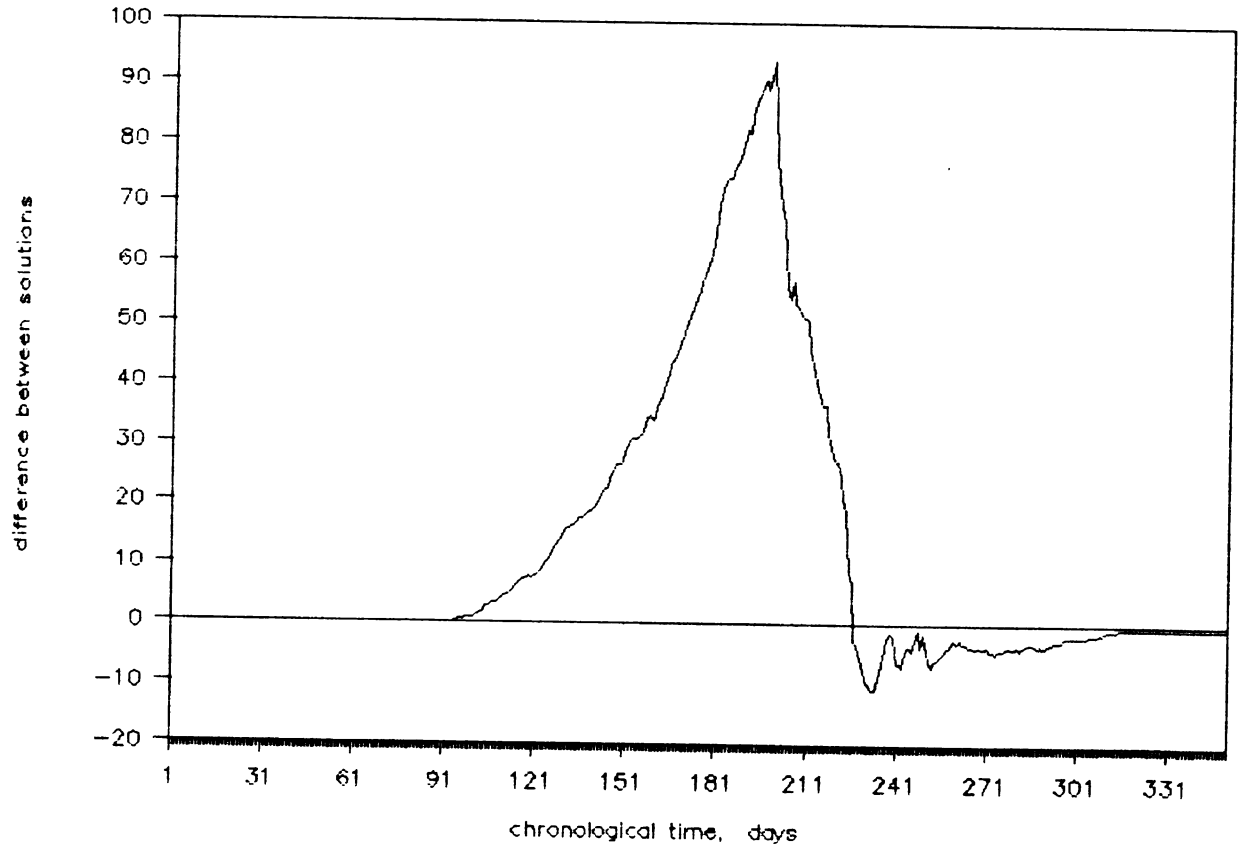


Figure 4.17a: Difference between population metrics as computed by the empirical temperature profile and a curve fitting approximation. $T_0=35^{o}$ F.

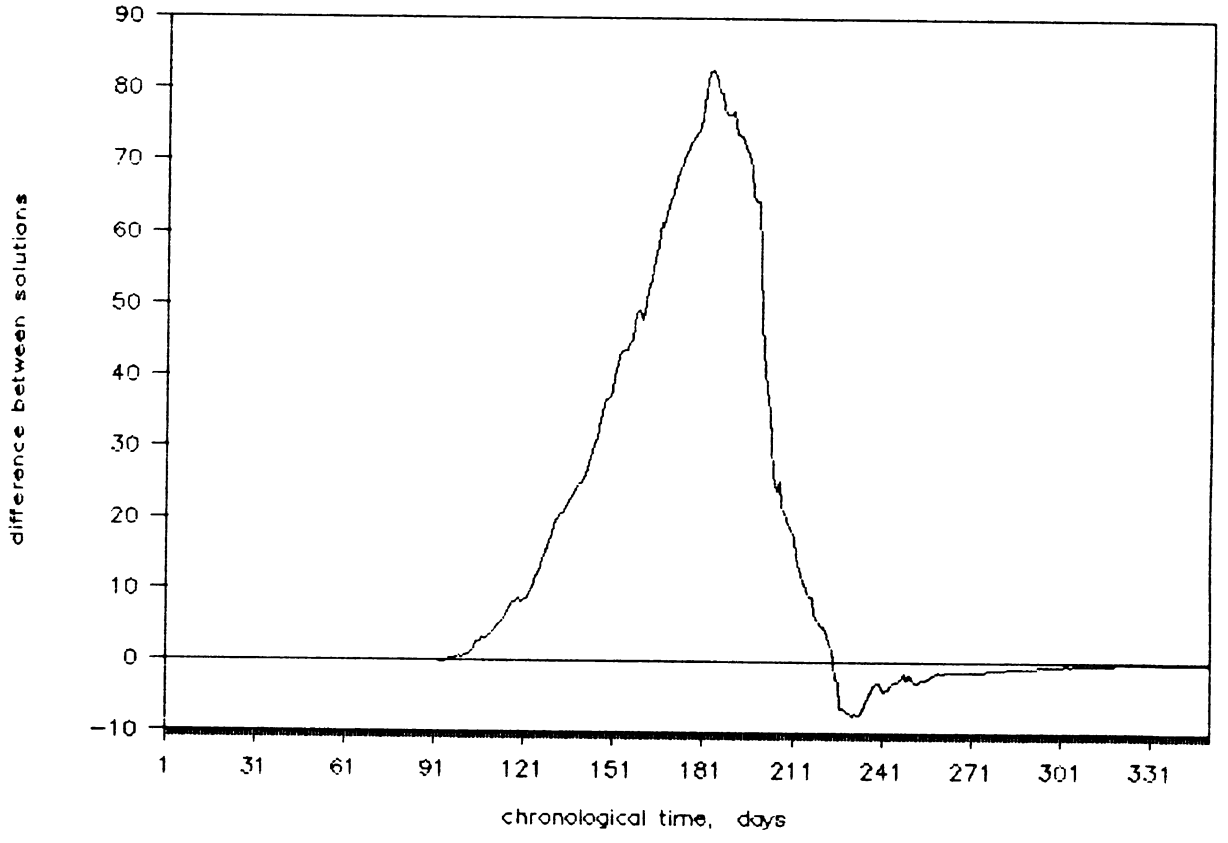


Figure 4.17b: Difference between population metrics as computed by the empirical temperature profile and a curve fitting approximation. $T_0=40^o$ F.

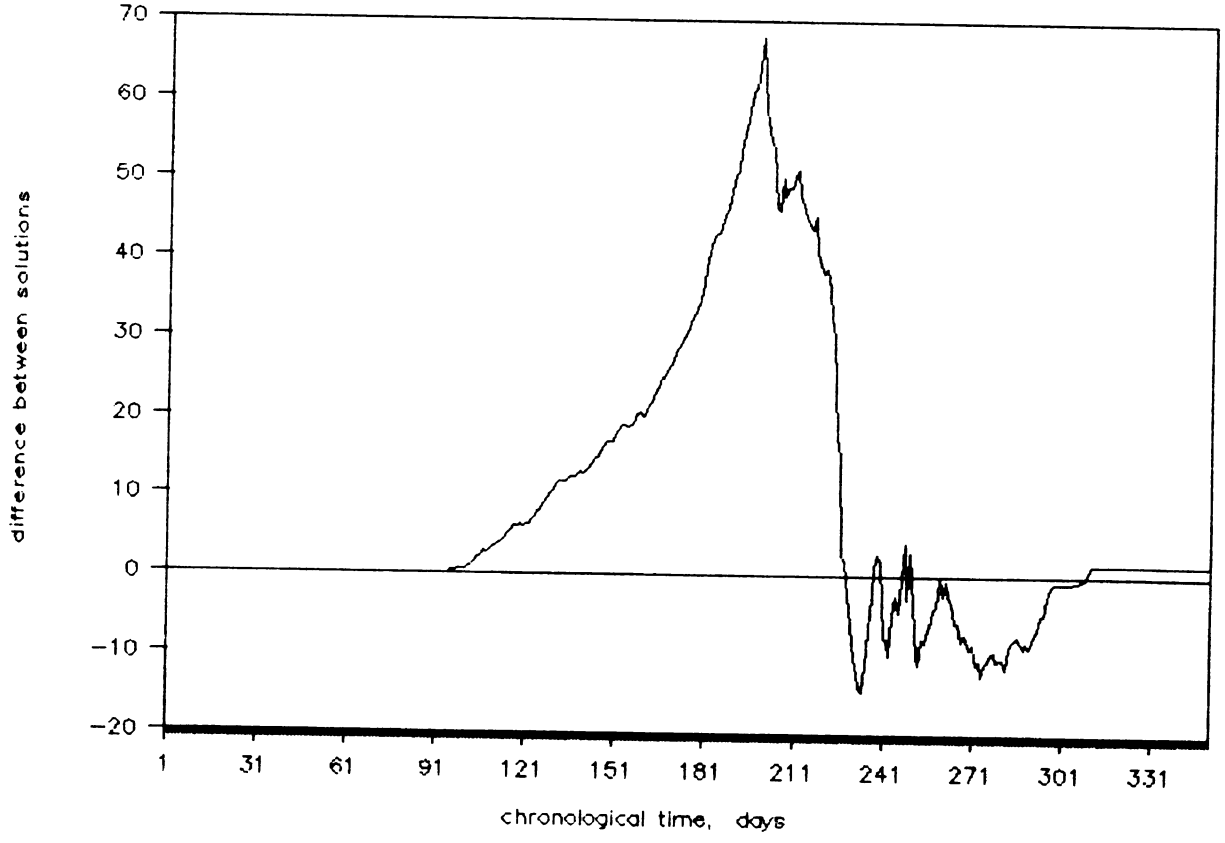


Figure 4.17c: Difference between population metrics as computed by the empirical temperature profile and a curve fitting approximation. $T_0=45^{o}$ F.

CHAPTER FIVE

PARAMETER ESTIMATION

INTRODUCTION

Unlike many physical models, biological models do not lend themselves to intuitively obvious values for their constants. To use a mathematical model it is first necessary to determine values for the various modeling parameters. For models using the heat-unit hypothesis this will necessitate finding thresholds T_k and rates r_k in addition to the explicit species interaction constants.

A traditional means by which to estimate parameters is that of least squares. For classical (non heat-unit) models this is done as follows:

> Suppose the actual population (or appropriate growth metric) y_k is known for chronological times t_k $k=1,\ldots,q$. If $y(t)$ is the population predicted by our model, we wish to minimize the total least-squares difference

$$[5.1] \qquad rms^2 = \frac{1}{q} \cdot \sum_{k=1}^{q} [\, y_k - y(t_k) \,]^2$$

> for the q data points $(t_1,y_1),\ldots,(t_q,y_q)$.

If the populations y_k are known with varying degrees of certainty, an estimate of the standard deviation σ_k can weight the above sum.

$$[5.2] \qquad rms^2 = \frac{1}{q} \cdot \sum_{k=1}^{q} \left[\frac{y_k - y(t_k)}{\sigma_k} \right]^2$$

Again we minimize rms^2 with respect to the parameters incorporated in y(t).

The above process can lead to a system of nonlinear equations. In only the most elementary of models can we determine analytic, closed form solutions. Thus, numerical techniques usually need to be employed.

This process is often static; that is, there are no differential-integral equations with which to contend. This is the case if y(t) is modeled by an algebraic expression with parameters. However, if we introduce the physiological time τ, it will never be the case, save in extremely contrived circumstances.

It is also easy to generalize on the basic heat-unit model. The simple single threshold model with two parameters T_0, r described by equation [5.1] is but one of many possible. Any system identification must also be applicable to this more general class.

The objectives of this chapter are to do the following:

i. Describe a more general class of heat-unit models.

ii. Derive an analytic procedure for use in evaluating certain model parameters.

iii. Develop the nonlinear techniques that will be employed later.

iv. Demonstrate these techniques on classical models.

v. Derive the fundamental equations for parameter identification using the heat-unit hypothesis.

vi. Illustrate the general approach as applied to both single species populations and multi-species populations.

ENDOGENOUS AND EXOGENOUS SYSTEM PARAMETERS

Consider the single species logistic model using the heat-unit hypothesis described earlier:

$$[5.3] \qquad \frac{d}{d\tau} x(\tau) = r \cdot x(\tau) \cdot \left[1 - \frac{x(\tau)}{x_m} \right]$$

$$\text{and} \quad x(0) = x_0,$$

$$\tau(t) = \frac{1}{r} \cdot \int_{\Gamma(t)} \beta(\mu) d\mu$$

Since we view the physiological time τ as this species' time scale, the presence of other populations or environmental considerations should have no effect on parameters defining the biological activity $\beta(T)$. In the single threshold model considered thus far,

$$[5.4] \qquad \beta(t) = \begin{cases} 0, & T(t) \leq T_0 \\ r \cdot [T(t) - T_0], & T(t) > T_0 \end{cases}$$

these would be r and T_0. On the other hand, parameter x_m (the system capacity) is a measure of this species' tolerance to population density. Therefore, we consider these two types of parameters separately.

Parameters which define the physiological time are said to be **endogenous** constants. Again, these are independent of the species' environment. While x_m does not directly depend on the existence of other populations, it certainly is dependent upon the environment. Thus, we consider it to be a **exogenous** constant.

It is the goal of this chapter to consider the system identification problem. For a given temperature profile, we will describe methods for evaluating both the endogenous and exogenous parameters. Once this is done the system will be well defined and a solution is thus attainable.

DEFINING THE ENDOGENOUS SYSTEM PARAMETERS

The biological activity need not be defined only as equation [5.4]. In general it can be any fairly well behaved non-negative function $\beta(T) \geq 0$ for $t \geq 0$.

It is worth noting that this definition of $\beta(T)$ is not differentiable everywhere. We will continue in this vein, requiring the function to be differentiable **almost everywhere** and usually **piecewise linear**. Let us refer to the model in which biological activity is described by equation [5.4] as a **TYPE I** model. We now define several alternative approaches to biological activity.

TYPE II MODELS

Consider a first generalization. Suppose there exist thresholds T_0 and T_2 where there is no biological activity for $T \leq T_0$ or $T \geq T_2$. Between these two thresholds the activity is proportional to temperature with positive rate r_0 between T_0, T_1 and negative rate r_1 between T_1, T_2. Formally this [biological] activity function can be expressed as:

[5.5]
$$\beta(t) = \begin{cases} 0, & T(t) \leq T_0 \\ r_0 \cdot [T(t)-T_0], & T_0 < T(t) \leq T_2 \\ r_1 \cdot [T(t)-T_1], & T_2 < T(t) < T_1 \\ 0, & T(t) \geq 0 \end{cases}$$

[5.6]
$$\text{where} \quad T_2 = (r_0 T_0 - r_1 T_1)/(r_0 - r_1)$$
by continuity.

A graph of this activity function is shown in Figure 5.1. The physiological time generalizes to:

[5.7]
$$\tau(t) = \int_{\Gamma_0(t)} [T(\mu)-T_0]d\mu + \int_{\Gamma_1(t)} [T(\mu)-T_1]d\mu$$

$$= \frac{1}{r} \cdot \int_{\Gamma_0(t)} \beta(\mu)d\mu + \frac{1}{r} \cdot \int_{\Gamma_1(t)} [T(\mu)-T_1]d\mu$$

$$\equiv \tau_0(t) + \tau_1(t)$$

Thus,

$$[5.8] \qquad \int_{\Gamma(t)} \beta(T(\mu))d\mu = r_0 \cdot \tau_0(t) + r_1 \cdot \tau_1(t)$$

where the action sets $\Gamma_0(t)$, $\Gamma_1(t)$ are defined as:

$$[5.9] \qquad \Gamma_0(t) = \{t>t_0: T_0<T(t)<(r_0T_0-r_1T_1)/(r_0-T_1)\}$$

$$\Gamma_1(t) = \{t>t_0: (r_0T_0-r_1T_1)/(r_0-r_1)<T(t)<T_1\}$$

Generalizing this concept further, we define a set of minimal requirements for the biological activity. They are listed in the following definition:

Definition 5.1: A function $\beta(T)$ is said to be a **strict activity function** if the following conditions are met:

i. $\beta(T)\geq 0$ for all T.

ii. There exist thresholds T_0 and T_1 such that:

$\beta(T)=0$ for $T\leq T_0$ and $T\geq T_1$.

iii. $\beta(T)$ is continuous.

iv. $\beta(T)$ is piecewise differentiable.

By dropping condition (iii.) we define a (non strict) **activity function** and by tightening condition (iv.) to be piecewise linear we define a **piecewise linear activity function.**

We can approximate strict activity functions by the use of piecewise linear activity functions.

Unless the specific form of such a function is known a'priori, one possible approach is to approximate $\beta(T)$ with n inscribed line segments. Such an approach constitutes a piecewise continuous activity function and is illustrated in Figure 5.2. It may be mathematically represented as:

$$[5.10] \qquad \beta(T) = \begin{cases} 0, & T(t) \le T_0 \\ r_0 \cdot [T(t)-a_0], & T_0 < T(t) \le T_1 \\ r_1 \cdot [T(t)-a_1], & T_1 < T(t) \le T_2 \\ \quad \vdots & \quad \vdots \\ r_{n-1} \cdot [T(t)-a_{n-1}], & T_{n-1} < T(t) \le T_n \\ 0, & T(t) \ge T_n \end{cases}$$

where the constants $a_0=T_0, a_1, \ldots, a_{n-1}=T_n$ are chosen so that $\beta(T)$ is continuous at $T_1, \ldots, T_{n-1}$. Therefore,

$$r_i[T_{i+1}-a_i] = r_{i+1}[T_{i+1}-a_{i+1}]$$

[5.11]
$$a_{i+1} = r_i a_i / r_{i+1} + T_{i+1}(1 - r_i / r_{i+1}), \qquad i=1,\ldots,n-2$$

where $a_0 = T_0$, $a_{n-1} = T_n$

It is possible to define a physiological time for this more general continuous type model. As in the Type II model we proceed using the biological activity function:

$$\tau(t) = \int_{\Gamma(t)} \beta(T(\mu))\, d\mu$$

$$= \int_{\Gamma_0(t)} [T(\mu) - a_0]d\mu + \cdots + \int_{\Gamma_{n-1}(t)} [T(\mu) - a_{n-1}]d\mu$$

[5.12]
$$= \frac{1}{r} \cdot \int_{\Gamma_0(t)} \beta(T(\mu))d\mu + \cdots + \frac{1}{r_{n-1}} \cdot \int_{\Gamma_1(t)} \beta(T(\mu))d\mu$$

$$\equiv \tau_0(t) + \cdots + \tau_{n-1}(t)$$

In general we write...

$$\int_{\Gamma(t)} \beta(T(\mu))d\mu = r_0 \cdot \tau_0(t) + \cdots + r_{n-1} \cdot \tau_{n-1}(t)$$

[5.13]

$$= \sum_{k=0}^{n-1} r_k \cdot \tau_k(t)$$

where the action sets are defined as...

[5.14] $\Gamma_k(t) = \{t>t_0: T_k<T(t)<T_{k-1}\}, \qquad k=0,\ldots,n-1$

Unfortunately, as an intuitive assist this formalism is of questionable value. While it might lead to more precision, it does not present further insight. Therefore, we will consider two more specific examples.

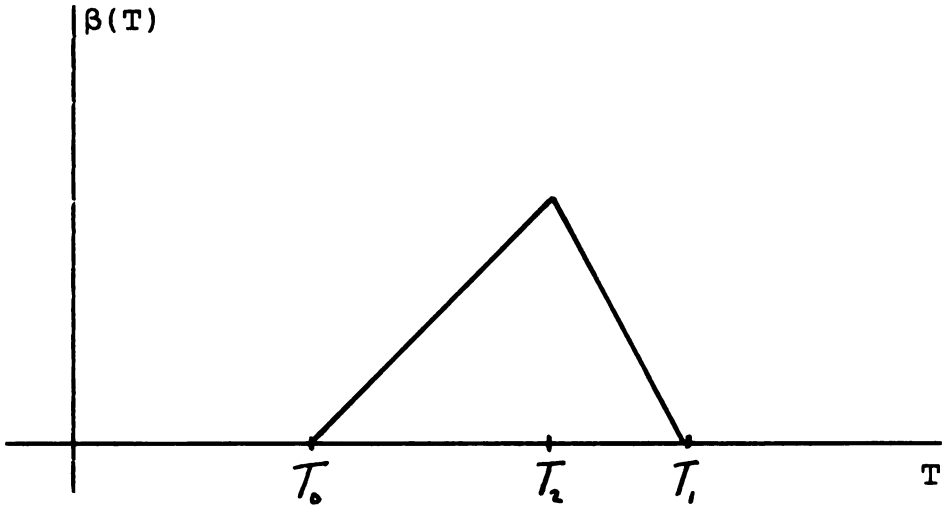


Figure 5.1 The Biological Activity function for a Type II model.

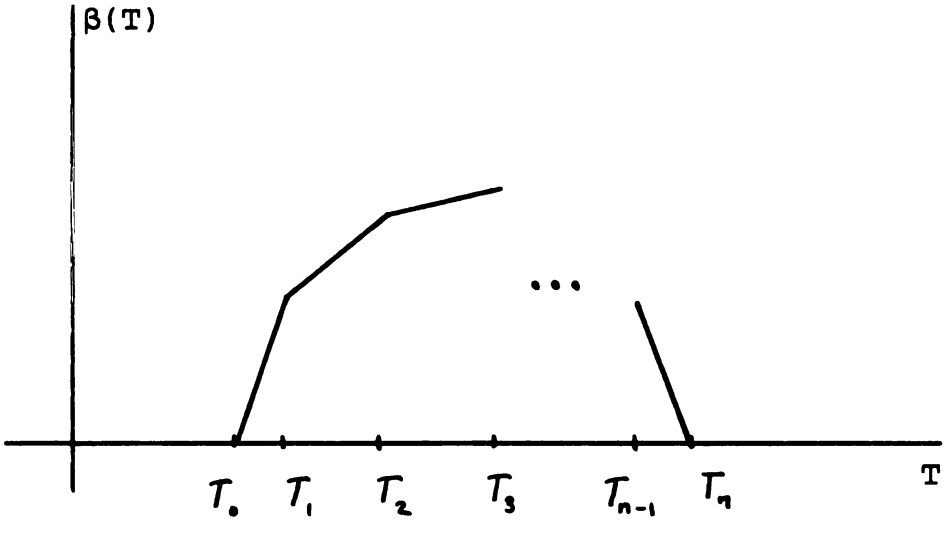


Figure 5.2 The Biological Activity function for a general piecewise continuous model.

TYPE III MODELS

If the biological activity has lower and upper thresholds T_0, T_1 respectively and the activity is proportional to T between these thresholds, $\beta(T)$ takes the following form:

$$[5.15] \qquad \beta(T) = \begin{cases} 0, & T(t) \le T_0 \\ r \cdot [T(t) - T_0], & T_0 < T(t) \le T_1 \\ 0, & T(t) \ge T_1 \end{cases}$$

This Type III [biological] activity function is sketched in Figure 5.3.

Defining the action set as

$$[5.16] \qquad \Gamma(t) = \{t > t_0 : T_0 < T(t) \le T_1\}$$

the physiological time is given by

$$[5.17] \qquad \tau(t) = \frac{1}{r} \cdot \int_{\Gamma(t)} [T(\mu) - T_0] d\mu$$

Notice that there are three parameters to be evaluated here as opposed to two in the Type I formulation and four in Type II models.

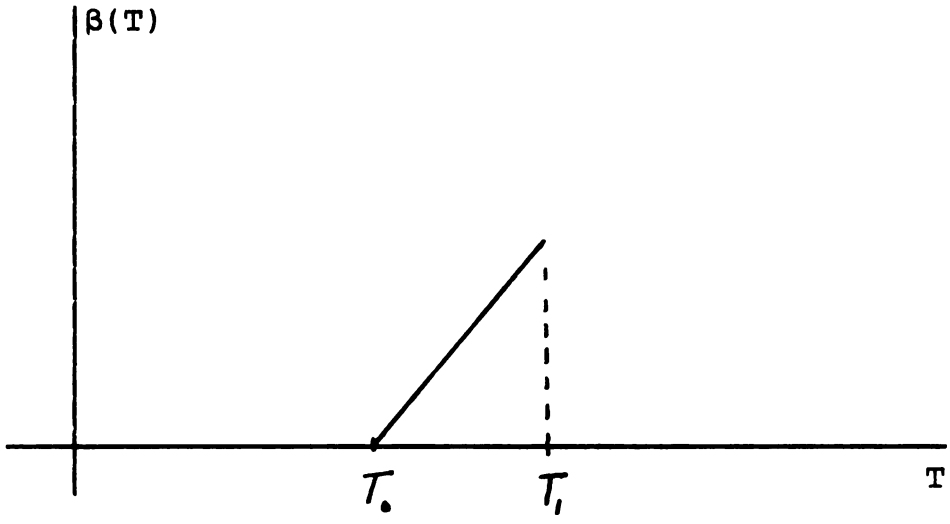


Figure 5.3 The Biological Activity function for a Type III model.

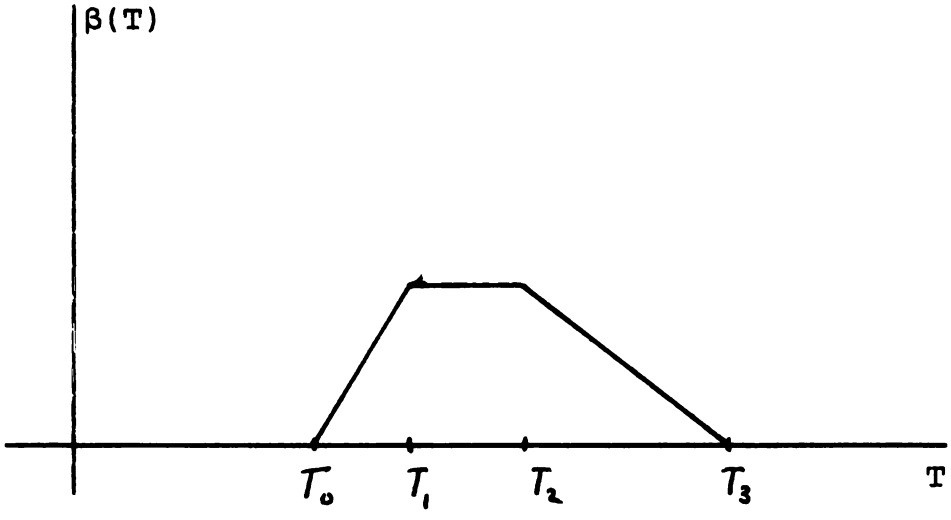


Figure 5.4 The Biological Activity function for a Type IV model.

TYPE IV MODELS

Consider $\beta(T)$ as defined by the graph in Figure 5.4. Mathematically this may be written by:

$$\beta(T) = \begin{cases} 0, & T(t) \le T_0 \\ r_0 \cdot [T(t) - T_0], & T_0 < T(t) \le T_1 \\ r_1, & T_1 < T(t) \le T_2 \\ r_2 \cdot [T(t) - T_3], & T_2 < T(t) \le T_3 \\ 0, & T(t) \ge T_3 \end{cases} \qquad [5.18]$$

Continuity conditions imply

$$\begin{aligned} r_0 \cdot (T_1 - T_0) &= r_1 \\ r_2 \cdot (T_2 - T_3) &= r_1 \end{aligned} \qquad [5.19]$$

This gives

$$\begin{aligned} T_1 &= T_0 + r_1/r_0 \\ T_2 &= T_3 + r_1/r_0 \end{aligned} \qquad [5.20]$$

Therefore, the action sets are:

$$\Gamma_0(t) = \{t>t_0: T_0<T(t)\leq r_1/r_0+T_0\}$$

[5.21] $$\Gamma_1(t) = \{t>t_0: r_1/r_2+T_0<T(t)\leq r_1/r_2+T_3\}$$

$$\Gamma_2(t) = \{t>t_0: r_1/r_2+T_3<T(t)\leq T_3\}$$

and the physiological time is

[5.22] $$\tau(t) = \int_{\Gamma_0(t)}[T(\mu)-T_0]d\mu + \int_{\Gamma_1(t)}d\mu + \int_{\Gamma_2(t)}[T(\mu)-T_2]d\mu$$

We note that here there are five parameters: T_0, T_3, r_0, r_1 and r_2 to be evaluated in this model.

Each of these models and their characteristics are summarized in Tables 5.1, 5.2, 5.3, 5.4 and 5.5.

Table 5.1 Summary of Type I model characteristics.

Biological activity function:

$$\beta(t) = \begin{cases} 0, & T(t) \leq T_0 \\ r \cdot [T(t) - T_0], & T(t) > T_0 \end{cases}$$

Physiological time:

$$\tau(t) = \frac{1}{r} \cdot \int_{\Gamma(t)} \beta(\mu) d\mu$$

Action set:

$$\Gamma(t) = \{t > t_0 : T(t) > T_0\}$$

Activity function/Physiological time relationship:

$$\int_{\Gamma(t)} \beta(T(\mu)) d\mu = r \cdot \tau(t)$$

Table 5.2: Summary of Type II model characteristics.

Biological activity function:

$$\beta(t) = \begin{cases} 0, & T(t) \le T_0 \\ r_0 \cdot [T(t) - T_0], & T_0 < T(t) \le T_2 \\ r_1 \cdot [T(t) - T_1], & T_2 < T(t) < T_1 \\ 0, & T(t) \ge 0 \end{cases}$$

Physiological time:

$$\tau_0(t) = \frac{1}{r} \cdot \int_{\Gamma_0(t)} \beta(\mu) d\mu$$

$$\tau(t) = \frac{1}{r} \cdot \int_{\Gamma_1(t)} \beta(\mu) d\mu$$

$$\tau(t) = \tau_0(t) + \tau_1(t)$$

Action sets:

$$\Gamma_0(t) = \{t > t_0 : T_0 < T(t) < (r_0 T_0 - r_1 T_1)/(r_0 - T_1)\}$$

$$\Gamma_1(t) = \{t > t_0 : (r_0 T_0 - r_1 T_1)/(r_0 - r_1) < T(t) < T_1\}$$

$$\Gamma(t) = \Gamma_0(t) \cup \Gamma_1(t)$$

Activity function/Physiological time relationship:

$$\int_{\Gamma(t)} \beta(T(\mu)) d\mu = r_0 \cdot \tau_0(t) + r_1 \cdot \tau_1(t)$$

Table 5.3: Summary of Type III model characteristics.

Biological activity function:

$$\beta(T) = \begin{cases} 0, & T(t) \leq T_0 \\ r \cdot [T(t) - T_0], & T_0 < T(t) \leq T_1 \\ 0, & T(t) \geq T_1 \end{cases}$$

Physiological time:

$$\tau(t) = \frac{1}{r} \cdot \int_{\Gamma(t)} [T(\mu) - T_0] d\mu$$

Action set:

$$\Gamma(t) = \{t > t_0 : T_0 < T(t) \leq T_1\}$$

Activity function/Physiological time relationship:

$$\int_{\Gamma(t)} \beta(T(\mu)) d\mu = r \cdot \tau(t)$$

Table 5.4 Summary of Type IV model characteristics.

Biological activity function:

$$\beta(T) = \begin{cases} 0, & T(t) \le T_0 \\ r_0 \cdot [T(t) - T_0], & T_0 < T(t) \le T_1 \\ r_1, & T_1 < T(t) \le T_2 \\ r_2 \cdot [T(t) - T_3], & T_2 < T(t) \le T_3 \\ 0, & T(t) \ge T_3 \end{cases}$$

where $T_1 = T_0 + r_1/r_0$,

$T_2 = T_3 + r_1/r_0$

Physiological time:

$$\tau_0(t) = \frac{1}{r} \cdot \int_{\Gamma_0(t)} \beta(\mu) d\mu$$

$$\tau_1(t) = \frac{1}{r} \cdot \int_{\Gamma_1(t)} \beta(\mu) d\mu$$

$$\tau_2(t) = \frac{1}{r} \cdot \int_{\Gamma_2(t)} \beta(\mu) d\mu$$

$$\tau(t) = \tau_0(t) + \|\Gamma_1(t)\| + \tau_2(t)$$

Table 5.4 (cont'd):

Action sets:

$$\Gamma_0(t) = \{t>t_0: T_0<T(t)\leq r_1/r_0+T_0\}$$

$$\Gamma_1(t) = \{t>t_0: r_1/r_2+T_0<T(t)\leq r_1/r_2+T_3\}$$

$$\Gamma_2(t) = \{t>t_0: r_1/r_2+T_3<T(t)\leq T_3\}$$

$$\Gamma(t) = \Gamma_0(t) \cup \Gamma_1(t) \cup \Gamma_2(t)$$

Activity function/Physiological time relationship:

$$\int_{\Gamma(t)} \beta(T(\mu))d\mu = r_0\cdot\tau_0(t) + r_1\cdot\tau_1(t) + r_2\cdot\tau_2(t)$$

Table 5.5 Summary of general continuous model characteristics.

Biological activity function:

$$\beta(T) = \begin{cases} 0, & T(t) \le T_0 \\ r_0 \cdot [T(t)-a_0], & T_0 < T(t) \le T_1 \\ r_1 \cdot [T(t)-a_1], & T_1 < T(t) \le T_2 \\ \quad \vdots & \quad \vdots \\ r_{n-1} \cdot [T(t)-a_{n-1}], & T_{n-1} < T(t) \le T_n \\ 0, & T(t) \ge T_n \end{cases}$$

where $a_{i+1} = r_i a_i / r_{i+1} + T_{i+1}(1 - r_i / r_{i+1})$, $i=1,\ldots,n-2$

$a_0 = T_0$, $a_{n-1} = T_n$

Physiological time:

$$\tau_k(t) = \frac{1}{r} \cdot \int_{\Gamma_k(t)} \beta(\mu) d\mu$$

where $k=0,1,\ldots,n-1$

$$\tau(t) = \sum_{k=0}^{n-1} \tau_k(t)$$

Table 5.5 (cont'd)

Action sets:

$$\Gamma_k(t) = \{t>t_0 : T_k<T(t)<T_{k-1}\}$$

where $k=0,1,\ldots,n-1$

$$\Gamma(t) = \bigcup_{k=0}^{n-1} \Gamma_k(t)$$

Activity function/Physiological time relationship:

$$\int_{\Gamma(t)} \beta(T(\mu))d\mu = \sum_{k=0}^{n-1} r_k \cdot \tau_k(t)$$

DETERMINING THE ENDOGENOUS PARAMETERS

Once a model for the biological activity has been chosen, we must next estimate the endogenous parameters. Traditionally only the simplest models are used and "eyeball" estimation techniques are employed. In this section we will develop an analytic approach to this problem and illustrate such a procedure using model I. A more general numeric procedure for all such models will be described and illustrated in further sections.

Since we are only estimating the endogenous parameters the system to be solved is simply the single species exponential model. In other words, these parameters are determined assuming no species-species or species-environment interactions. Using the given species physiological time $\Gamma(t)$, the description is as follows:

$$[5.23] \qquad \frac{d}{d\tau} x(\tau) = r \cdot x(\tau), \qquad x(0) = x_0$$

$$\text{where} \qquad \beta(T(t)) = r \cdot \frac{d}{dt} \tau(t)$$

This is easily integrable as

$$[5.24] \qquad \ln[x(t)/x_0] = \int_{\Gamma(t)} \beta(T(\mu))d\mu$$

Rather than use equation [5.1] directly, we will minimize on the natural logarithm of the residuals. This is a standard mathematical convenience. Since the logarithm function is one to one, minimizing the difference of logarithms is approximately the same as minimizing the difference of the data directly. We denote this objective function as rms^2 (root mean square), which is the average squared deviation of the model compared to the data.

$$[5.25] \qquad rms^2 = \frac{1}{q} \cdot \sum_{k=1}^{q} \left[\ln(y_k)-\ln(x(t_k))\right]^2$$

Normalizing our data, we let $x_0=1$ and substitute:

$$[5.26] \qquad rms^2 = \frac{1}{q} \cdot \sum_{k=1}^{q} \left[\ln(y_k) - \int_{\Gamma(t_k)} \beta(T))d\mu\right]^2$$

IDENTIFICATION FOR TYPE I MODELS

Consider a Type I endogenous model. This is the standard heat unit measure using only a lower threshold T_0. The two parameters to be estimated are T_0 and r.

$$[5.27] \qquad rms^2 = \frac{1}{q} \cdot \sum_{k=1}^{q} \left[\ln(y_k) - r \cdot \int_{\Gamma(t_k)} (T(\mu) - T_0) d\mu \right]^2$$

Since rms_2 is a differentiable function of two parameters, minimization can be accomplished when the partial derivatives (with respect to r) vanish:

$$[5.28] \qquad \sum_{k=1}^{q} [\ln(y_k) - r \cdot \tau(t_k)] \tau(t(_k)) = 0$$

Solving for r,

$$[5.29] \qquad r = \frac{\sum_{k=1}^{q} \ln(y_k) \cdot \tau(t_k)}{\sum_{k=1}^{n} \tau^2(t_k)}$$

Similarly, differentiating with respect to T_0 yields:

$$[5.30] \qquad \sum_{k=1}^{q} [\ln(y_k) - r\cdot\tau(t_k)]\cdot\int_{\Gamma(t_k)} d\mu = 0$$

Recalling that $\int_{\Gamma(t)} d\mu = \|\Gamma(t)\|$,

$$[5.31] \qquad \sum_{k=1}^{q} \ln(y_k)\cdot\|\Gamma(t_k)\| = r\cdot\sum_{k=1}^{q} \tau(t_k)\cdot\|\Gamma(t_k)\|$$

Again solving for r,

$$[5.32] \qquad r = \frac{\sum_{k=1}^{q} \ln(y_k)\cdot\|\Gamma(t_k)\|}{\sum_{k=1}^{q} \tau(t_k)\cdot\|\Gamma(t_k)\|}$$

In principle we need only combine equations [5.29] and [5.32] and then solve for T_0. Unfortunately this is not trivial in practice! We press ahead with a special class of temperature profiles.

PERIODIC TEMPERATURE PROFILES

Consider a class of periodic temperature profiles with period p and n cycles. Let T_0 be an arbitrary Type I threshold. Figure 5.5 illustrates a typical temperature profile, the corresponding physiological time and action sets.

It follows that the norm of $\Gamma(t)$ is given by:

$$[5.33] \qquad \|\Gamma(t)\| = \begin{cases} 0, & t<t_1 \\ t-t_1, & t_1 \le t<t_2 \\ t_2-t_1, & t_2 \le t<t_3 \\ t-t_3+t_2-t_1, & t_3 \le t<t_4 \\ t_4-t_3+t_2-t_1, & t_4 \le t<t_5 \\ \quad \vdots & \\ \sum_{k=1}^{2n-2} (-1)^k \cdot t_k, & t_{2n-2} \le t<t_{2n-1} \\ t - \sum_{k=1}^{2n-1} (-1)^k \cdot t_k, & t_{2n-1} \le t<t_{2n} \end{cases}$$

Since the profile is periodic with period p,

$$[5.34] \qquad t_k = \begin{cases} t_1 + \tfrac{1}{2}p(k-1), & k \text{ odd} \\ t_2 + \tfrac{1}{2}p(k-2), & k \text{ even} \end{cases}$$

Two specific cases follow.

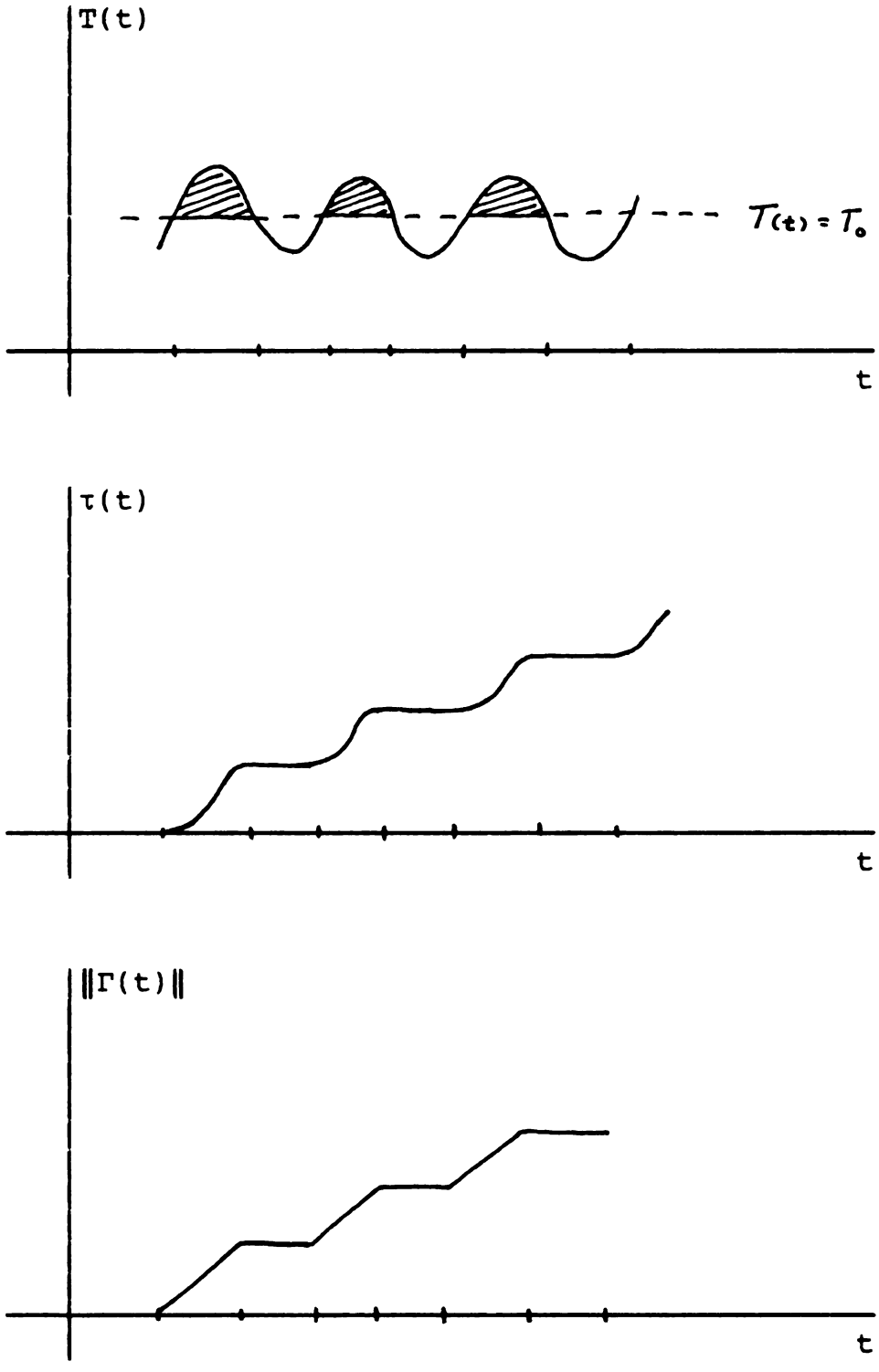


Figure 5.5 A typical periodic temperature profile and the corresponding physiological time and action sets.

CONSTANT MAGNITUDE TEMPERATURE PROFILES

From a practitioner's point of view, there are two choices by which to determine model parameters. The first is to use field data. If this is the case, data determining endogenous parameters will soon become contaminated with exogenous effects (i.e. species-species and species-environment interactions). Certainly this is a viable methodology, but one which requires considerable care. For early times (those close to $t=t_0$), the endogenous effects will dominate those of the environment. On the other hand, as time increases, exogenous effects will become increasingly apparent. From an analyst's perspective, a time interval during which endogenous effects are exemplified must be defined at the outset.

A second approach is to use laboratory data. Presumably in a laboratory environment, all exogenous effects can be screened, leaving only the species' own mechanism to be effected. Obviously this method has advantages: The analyst need make no subjective decision as to the endogenous time interval. Also, temperature profiles can be idealized in order to allow for mathematical simplification of the formulation.

Let us consider an idealized laboratory situation. A species is introduced at time $t=t_0$. Further let us assume

the temperature profile is binary. That is, there are effectively two temperatures working in a periodic fashion: one temperature is above the presumed threshold while the other below.

In this scenario the organism will only grow during the odd intervals using a constant temperature $T(t)=A$.

Defining the **active interval** as $I_k = \{t>t_0: t_k<t\leq t_{k+1}\}$,

$$[5.35] \qquad \|\Gamma(t)\| = \begin{cases} 0, & t \in I_0 \\ t-t_1, & t \in I_1 \\ t_2-t_1, & t \in I_2 \\ t-t_3+t_2-t_1, & t \in I_3 \\ t_4-t_3+t_2-t_1, & t \in I_4 \\ \quad\vdots & \\ \sum_{k=1}^{2n-2} (-1)^k \cdot t_k, & t \in I_{2n-2} \\ t - \sum_{k=1}^{2n-1} (-1)^k \cdot t_k, & t \in I_{2n1-1} \end{cases}$$

Similarly, the physiological time is given by equation

[5.36].

$$\text{[5.36]} \qquad \tau(t) = \begin{cases} 0, & t \in I_0 \\ (A-T_0)\cdot(t-t_1), & t \in I_1 \\ (A-T_0)\cdot(t_2-t_1), & t \in I_2 \\ (A-T_0)\cdot(t-t_3+t_2-t_1), & t \in I_3 \\ (A-T_0)\cdot(t_4-t_3+t_2-t_1), & t \in I_4 \\ \quad\vdots & \\ (A-T_0)\cdot \sum_{k=1}^{2n-2} (-1)^k \cdot t_k, & t \in I_{2n-2} \\ (A-T_0)\cdot[t - \sum_{k=1}^{2n-1} (-1)^k \cdot t_k], & t \in I_{2n-1} \end{cases}$$

It is evident from the above that

$$\text{[5.37]} \qquad \tau(t) = (A-T_0)\cdot\|\Gamma(t)\| \qquad \text{for } t>t_0$$

Substituting [5.37] into equation [5.32] we find that equations [5.32] and [5.37] are equivalent. This implies that all thresholds will be equally valid and that the calculation of T_0 is independent of the data! This is clearly unsatisfactory and unfortunately this particular laboratory procedure is meaningless.

SINUSOIDAL TEMPERATURE PROFILES

We now consider the class of sinusoidal temperature profiles. Let T(t) be of the following form:

$$T(t) = A + B\cdot\cos(wt) \tag{5.38}$$

Also, let T_0 be an arbitrary threshold between the temperature extremes aver n cycles. By defining

$$z = (T_0-A)/B, \tag{5.39}$$

$$c = \cos^{-1}(z), \tag{5.40}$$

the set of end points $\{t_k\}$ defining the action and inaction intervals are given as follows:

$$\begin{aligned} t_1 &= c/w \\ t_2 &= (2\pi-c)/w \\ t_3 &= (2\pi+c)/w \\ t_4 &= (4\pi-c)/w \\ &\vdots \\ t_{2n-1} &= [2(n-1)\pi+c]/w \\ t_{2n} &= (2n\pi-c)/w \end{aligned} \tag{5.41}$$

The norm is explicitly expressed by

$$[5.42] \qquad \|\Gamma(t)\| = \begin{cases} 0, & t<c/w \\ t - c, & c/w \le t<(2\pi-c)/w \\ 2\cdot(\pi-c)/w, & (2\pi-c)/w \le t<(2\pi+c)/w \\ t - 3\cdot c/w, & (2\pi+c)/w \le t<4\pi-c)/w \\ 4\cdot(\pi-c)/w, & (4\pi-c)/w \le t<(4\pi+c)/w \\ \quad\vdots & \end{cases}$$

An explicit formula for the physiological time is also straight forward. It, too, depends on a case argument, which goes as follows:

$$[5.43] \qquad \tau(t) = 0 \qquad \text{for } t < t_1$$

$$[5.44] \qquad \tau(t) = \int_{t_1}^{t} [A+B\cdot\cos(w\mu)-T_0]d\mu$$

$$= A\cdot(t-t_1) + B\cdot w^{-1}\cdot[\sin(w\mu)]\Big|_{t_1}^{t}$$

$$= A\cdot[t-\cos^{-1}(z)] + B\cdot[\sin(wt)-\sqrt{(1-z^2)}]/w$$

$$\text{for } t_1 \le t < t_2$$

$$[5.45] \qquad \tau(t) = \int_{t_1}^{t_2} [A+B\cdot\cos(w\mu)-T_0]d\mu$$

$$= A\cdot(t_2-t_1) + B\cdot[\sin(wt_2) - \sin(wt_1)]/w$$

$$= 2\cdot[\ A\cdot\pi - c - B\cdot\sqrt{(1-z^2)}\ ]$$

$$\text{for} \quad t_2 \le t < t_3$$

Similarly,

$$[5.46] \qquad \tau(t) = A\cdot[t-3\cdot c/w] + B\cdot[\sin(wt) - 3\cdot\sqrt{(1-z^2)}]$$

$$\text{for } t_3 \le t < t_4$$

$$[5.47] \qquad \tau(t) = 4\cdot[A\pi - A\cdot c - B\cdot\sqrt{(1-z^2)}]$$

$$\text{for } t_4 \le t < t_5$$

Most generally,

$$[5.48] \qquad \|\Gamma(t)\| = \begin{cases} k\cdot(\pi-c)/w, & k \text{ even} \\ t - k\cdot c/w, & k \text{ odd} \end{cases}$$

$$[5.49] \qquad \tau(t) = \begin{cases} k\cdot[A - A\cdot c - B\cdot\sqrt{(1-z^2)}], & k \text{ even} \\ A\cdot(t-k\cdot c/w) + B\cdot[\sin(wt) - k\cdot\sqrt{(1-z^2)}], & k \text{ odd} \end{cases}$$

$$[5.50] \qquad \sum_{i=1}^{k} (-1)^i \cdot t_i = \begin{cases} k \cdot (\pi - c)/w, & k \text{ even} \\ -k \cdot c/w, & k \text{ odd} \end{cases}$$

The most general case for periodic models is shown in Tables 5.6 and 5.7.

For the best estimate (in the least squares sense) of T_0 and r, equations [5.32], [5.37] must be simultaneously the same:

$$[5.51] \qquad \frac{\sum_{k=1}^{q} \ln(y_k) \cdot \tau(t_k)}{\sum_{k=1}^{q} \tau^2(t_k)} = \frac{\sum_{k=1}^{q} \ln(y_k) \cdot \|\Gamma(t_k)\|}{\sum_{k=1}^{q} \tau(t_k) \cdot \|\Gamma(t_k)\|}$$

No explicit solution to this equation has been found. However, a rough approximation can be derived in the next section.

Table 5.6: Summary Physiological time results for the sinusoidal temperature profile model.

$$\tau(t) = \begin{cases} 0 & \text{for } t < t_1 \\ A\cdot[t-c/w] + B\cdot[\sin(wt) - \sqrt{(1-z^2)}]/w & \text{for } t_1 \le t < t_2 \\ 2\cdot[\, A\cdot\pi - c - B\cdot\sqrt{(1-z^2)}\,] & \text{for } t_2 \le t < t_3 \\ A\cdot[t-3\cdot c/w] + B\cdot[\sin(wt) - 3\cdot\sqrt{(1-z^2)}]/w & \text{for } t_3 \le t < t_4 \\ 4\cdot[A\pi - A\cdot c - B\cdot\sqrt{(1-z^2)}] & \text{for } t_4 \le t < t_5 \\ \vdots & \\ A\cdot[t-(2k-1)\cdot c/w] + B\cdot[\sin(wt) - (2k-1)\sqrt{(1-z2)}]/w & \text{for } t_{2k-1} \le t < t_{2k} \\ 2k\cdot[A\pi - A\cdot c - B\cdot\sqrt{(1-z_2)}] & \text{for } t_{2k} \le t < t_{2k+1} \end{cases}$$

$$\tau(t) = \begin{cases} k\cdot[A - A\cdot c - B\cdot\sqrt{(1-z^2)}], & k \text{ even} \\ A\cdot(t-k\cdot c/w) + B\cdot[\sin(wt) - k\cdot\sqrt{(1-z^2)}], & k \text{ odd} \end{cases}$$

Table 5.7 Summary Action set Norm results for the sinusoidal temperature model.

$$\|\Gamma(t)\| = \begin{cases} 0, & t < t_1 \\ t - c, & t_1 \le t < t_2 \\ 2\cdot(\pi-c)/w, & t_2 \le t < t_3 \\ t - 3\cdot c/w, & t_3 \le t < t_4 \\ 4\cdot(\pi-c)/w, & t_4 \le t < t_5 \\ \quad\vdots & \\ t - (2k-1)\cdot c/w, & t_{2k-1} \le t < t_{2k} \\ 2k\cdot(\pi-c)/w, & t_{2k} \le t < t_{2k+1} \end{cases}$$

$$\|\Gamma(t)\| = \begin{cases} k\cdot(\pi-c)/w & \text{for } k \text{ even} \\ t - k\cdot c/w, & \text{for } k \text{ odd} \end{cases}$$

THE DATA-INDEPENDENT APPROXIMATION

Although not necessary, certainly a sufficient condition for equation [5.51] is that

$$[5.52] \qquad \tau(t_k) = A \cdot \frac{\|\Gamma(t)\|}{1+a}$$

where a is a constant parameter for data points (t_k, y_k), $k=1,\ldots,n$. This is certainly not generally the case since the norm $\|\Gamma(t)\|$ is piecewise linear while the physiological time $\tau(t)$ is usually not.

However, approximating a solution of this form has one major advantage: it is independent of the data! Thus, determining the threshold T_0 will depend only on the temperature profile. This will hardly give consistently accurate results, but will still provide order of magnitude approximations while parameter a will measure the goodness of fit of $\tau(t)$ compared $A \cdot \|\Gamma(t)\|/(1+a)$.

In order for equation [5.51] to hold (see Figure 5.37), we first require equality to hold on each plateau:

[5.53] $$2\cdot k\cdot[A\pi - A\cdot c - B\cdot\sqrt{(1-z^2)}] = A/(1+a)\cdot 2\cdot k/w\cdot(\pi - c)$$

for interval I_{2k}: $[t_{2k}, t_{2k+1})$

Elementary algebra produces:

$$a\cdot A\cdot(\pi - c) = B\cdot(1+a)\cdot\sqrt{(1-z^2)}$$

[5.54]

$$T_0 = A - B\cdot\cos[(1+a^{-1})\cdot\sqrt{(B^2-(T_0-A)^2)}]$$

Equation [5.54] can be solved numerically by simple iterative stationary point methods. Given values for A, B and a, Algorithm 5.1 produces such a solution for T_0.

Since A and B are given temperature profile parameters, it remains but to determine a. A most likely (in the statistical sense) value for a, assuming random data can be found by minimizing the difference between $\tau(t)$ and $A\cdot\|\Gamma(t)\|/(1+a)$. We choose to force closeness by minimizing the square of the area enclosed by the two functions between successive plateaus as shown in Figure 5.6.

Consider the odd interval:

$$I_{2k-1} : [t_{2k-1}, t_{2k})$$

$$[2\pi\cdot(k-1)+c]/w \leq t < (2\pi\cdot k - c)/w$$

The crucial difference is

$$\text{diff} = \tau(t) - A\cdot\|\Gamma(t)\|/(1+a)$$

[5.55]

$$= A\cdot[t-(2k-1)\cdot c/w]$$

$$+ B\cdot[\sin(wt)-(2k-1)\cdot\sqrt{(1-z^2)}]/w$$

$$- A\cdot[t-(2k-1)\cdot c/w]$$

The area (to be minimized) is:

$$\text{area} = \sum_{k=1}^{q} \text{diff}^2$$

$$= \int_{[2(k-1)\pi+c]/w}^{(2k\pi-c)/w} \Big[A\cdot[t-(2k-1)\cdot c/w] + B\cdot[\sin(wt) - (2k-1)$$

$$\cdot\sqrt{(1-z^2)}]/w - A\cdot[t-(2k-1)\cdot c/w]/(1+a)\Big]^2 dt$$

[5.56]

$$= \frac{1}{w^2}\cdot\int_{2(k-1)\pi+c}^{2k\pi-c} \Big[A\cdot[t-(2k-1)\cdot c] + B\cdot[\sin(t) -$$

$$(2k-1)\cdot\sqrt{(1-z^2)}] - A\cdot[t-(2k-1)\cdot c]/(1+a)\Big]^2 dt$$

$$= \frac{1}{w^2\cdot(1+a)^2}\cdot\int_{2(k-1)\pi+c}^{2k\pi-c}\Big[\,a\cdot A\cdot[t-(2k-1)\cdot c] + (1+a)\cdot B\cdot[\sin(t) - (2k-1)\cdot\sqrt{(1-z^2)}]\Big]^2 dt$$

The derivative of the [area] with respect to parameter a must vanish at the optimal point. Thus,

$$[5.57]\qquad D_a[\text{area}] = \frac{2}{w^2\cdot(1+a)^2}\cdot\int_{2(k-1)\pi+c}^{2k\pi-c}\Big[\,a\cdot A\cdot[t-(2k-1)\cdot c + (1+a)\cdot B\cdot[\sin(t)-(2k-1)\cdot\sqrt{(1-z^2)}]\Big]$$

$$\cdot\Big[\,A\cdot[t-(2k-1)\cdot c] + B\cdot[\sin(t) - (2k-1)\cdot\sqrt{(1-z^2)}]\Big]\,dt = 0$$

Solving for a results in equation [5.58].

$$[5.58] \quad a = \frac{-B\cdot\int_{2(k-1)\pi+c}^{2k\pi-c}\left[\sin(t)-(2k-1)\cdot\sqrt{(1-z^2)}\right]\cdot\left[A\cdot[t-(2k-1)\cdot c] + B\cdot[\sin(t)-(2k-1)\cdot\sqrt{(1-z^2)}]\right]\,dt}{\int_{2(k-1)\pi+c}^{2k\pi-c}\left[A\cdot[t-(2k-1)\cdot c] + B\cdot[\sin(t) - (2k-1)\cdot\sqrt{(1-z^2)}]\right]\,dt}$$

These integrals, while complicated, are elementary and an explicit formula can be found for a. However, since this is not a particularly general result and this derivation was for illustrative purposes only, this section concludes with a numerical example.

$T_0 = A$

DO UNTIL $|T_1 - T_0| < \varepsilon$

$T_2 = T_0$

$T_0 = A - B \cdot \cos\{(1+1/a) \cdot [B^2 - (T_0 - A)^2]/a\}$

PRINT T_0

$\varepsilon \approx 0.01$

Algorithm 5.1 Computing the threshold T_0 for a Type I model using the data independent solution approximation.

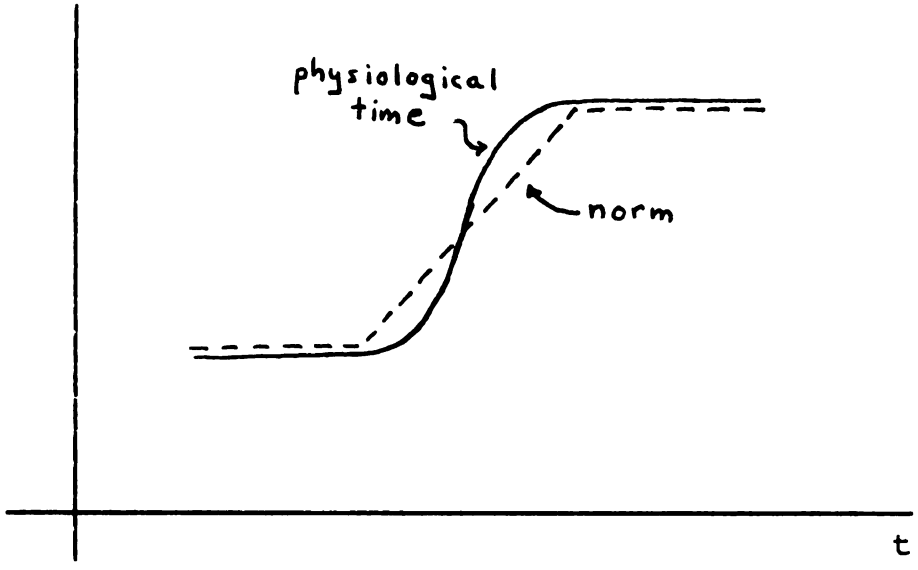


Figure 5.6 Relationship between the Physiological time $\tau(t)$ and the Action set Norm $\|\Gamma(t)\|$.

EXAMPLE

Consider an annual model (n=1) with a sinusoidal temperature profile of extremes 75^{o} and 45^{o} (A=60, B=15). Equation [5.54] can be solved for various values of a. Algorithm 5.1 produces the results shown in Table 5.8.

A numerical solution of equation [5.58] (off-the-shelf software using Rhomberg integration) produces the table of data in Table 5.9.

Graphs summarizing Tables 5.8 and 5.9 are shown in Figure 5.9. The approximate solution is easily read as the point of intersection of the two curves. Equation [5.29], in conjunction with the data:

$$(t_k, y_k), \qquad k=1,\ldots,q$$

if these points were given, calculates the rate r.

The value of a has a convenient interpretation. Recalling that A is the mean of the temperature profile, the relative difference between the proportionality constant, A/(1+a), and profile mean A is:

$$[5.59] \qquad \text{Relative Difference} = \frac{A - A/(1+a)}{A/(1+a)}$$

$$= a$$

Thus, $a \approx 0.25$ implies a fairly modest difference between $\tau(t)$ and $A \cdot \|\Gamma(t)\|/(1+a)$. The threshold of $T_0 \approx 53.7^{\circ}$ is rather stable and the data will not change this approximation to T_0 appreciably.

Table 5.8 Threshold T_0 as a function of parameter a determined by Algorithm 5.5.

a	T_0
0.33	45.46
0.30	48.01
0.27	50.85
0.25	53.62
0.22	56.30
0.20	58.80
0.18	61.34
0.15	63.67
0.13	65.86

Table 5.9 Parameter a as a function of threshold T_0 as determined by equation 5.47.

T_0	a
50	0.12779
51	0.13574
52	0.14203
53	0.14683
54	0.15029
55	0.15251
56	0.15362
57	0.15370
58	0.15287
59	0.15123

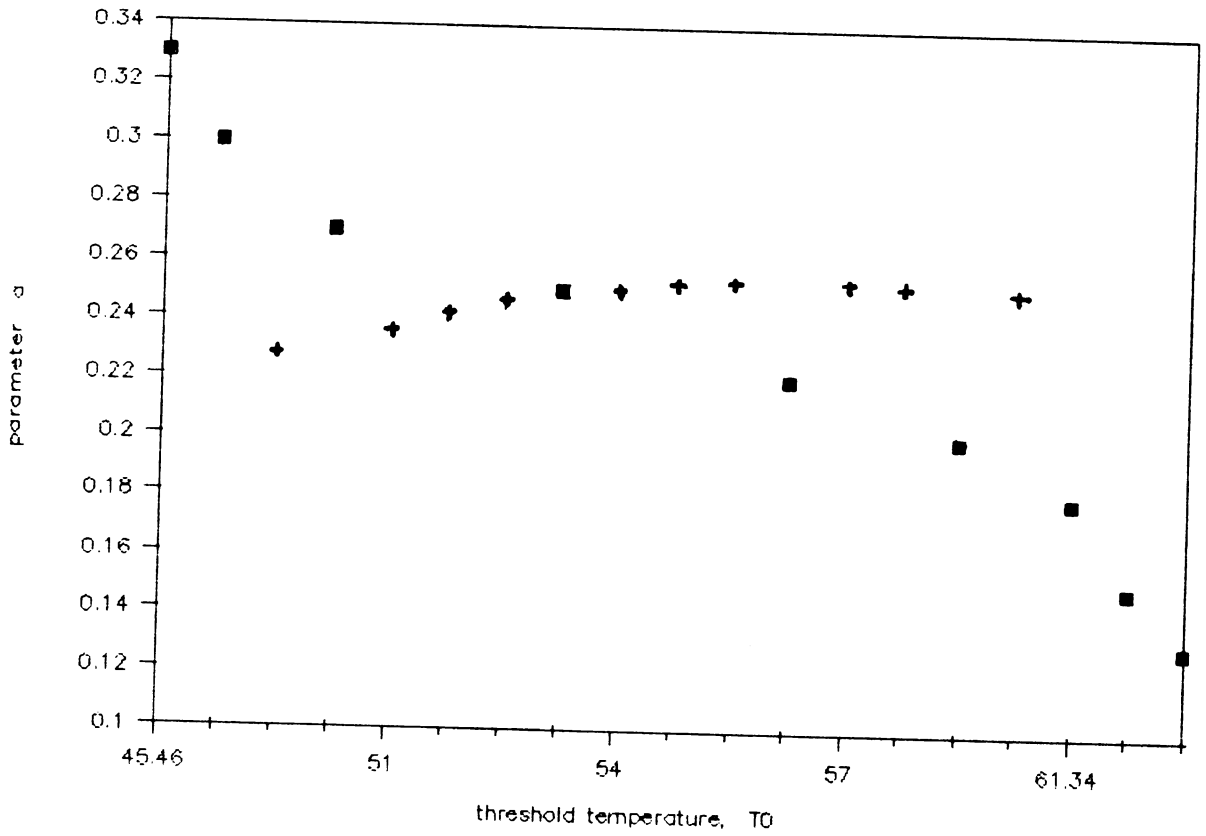


Figure 5.7: Plots of Tables 5.6 and 5.7 showing point of intersection for data independent solution:
$a \approx 0.25$, $T_0 \approx 53.7^{o}$.

• algorithm 5.5

\+ equation [5.58]

SUMMARY OF THE DATA INDEPENDENT METHOD

The above method for calculation is theoretically tempting but as a practical matter there are several inherent limitations.

i. A hypothetical temperature profile is assumed. Clearly the sinusoid is an oversimplification of reality. The only reflection of the true temperature profile is in the values of constants A, and B.

ii. The value of threshold T_0 is found independently of the data (t_k,y_k), $k=1,\ldots,q$. This is clearly unsatisfactory unless the assumption

$$\tau(t = A\cdot\|\Gamma(t)\|/(1+a)$$

is a particularly good approximation.

iii. The degree to which the above assumption is valid is reflected in the value of a. If $a\approx 0$, this can produce reasonable threshold values. Otherwise, if a is not close to zero, specific data can lead to significantly contradictory results.

iv. The integral for evaluating a is generally non-trivial.

A more straight forward method for evaluating the endogenous parameters is needed. Such a procedure is developed in the next chapter.

MICHIGAN STATE UNIV. LIBRARIES
31293005712207

2241153

THESIS
V. 2

MICHIGAN STATE UNIVERSITY LIBRARIES
3 1293 00571 2199

LIBRARY
Michigan State
University

MSU
LIBRARIES

RETURNING MATERIALS:
Place in book drop to remove this checkout from your record. FINES will be charged if book is returned after the date stamped below.

POPULATION DYNAMICS OF BIOLOGICAL SYSTEMS USING THE HEAT UNIT HYPOTHESIS

Volume II

By

Francis Lynn Severance

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Electrical Engineering

1988

CHAPTER SIX

NONANALYTIC OPTIMIZATION METHODS

INTRODUCTION

Because the chronological time $\tau(t)$ is empirical in nature, any practical method for determining the modeling parameters will need to be numerical. Fortunately there exist several excellent numerical optimization techniques that can be made applicable to this problem. This section demonstrates feasibility by implementing the necessary algorithms and applies these programs to sample data sets.

In practice this process is not automatic. It will be seen that both biological and numerical intuition are still necessary requisites and there is still room for creativity.

Also, unlike the analytic approach we will now be able to apply the actual temperature profile and data points.

THE ENDOGENOUS TYPE I PROBLEM

Recall that for the endogenous parameter problem it is convenient to minimize the objective function:

$$[6.1] \qquad rms^2 = \frac{1}{q} \cdot \sum_{k=1}^{q} [\ln(y_k)-\ln(x(t_k))]^2$$

$$= \frac{1}{q} \cdot \sum_{k=1}^{q} [\ln(y_k/x(t_k))]^2$$

Further, we assume an exponential model to determine the endogenous parameters. This is summarized by equation [6.2]:

$$[6.2] \qquad \ln(x(t)/x_0) = \int_{\Gamma(t)} \beta(T(\mu))d\mu$$

Combining equations [6.1] and [6.2],

$$[6.3] \qquad rms^2 = \frac{1}{q} \cdot \sum_{k=1}^{q} \left[\ln(y_k) - \int_{\Gamma(t)} \beta(T(\mu))d\mu \right]^2$$

For the Type I problem,

$$[6.4] \qquad \int_{\Gamma(t)} \beta(T(\mu))d\mu = r\cdot\tau(t)$$

Thus,

$$[6.5] \qquad rms^2 = \frac{1}{q} \cdot \sum_{k=1}^{q} [\ln(y_k - r\cdot\tau(t_k)]^2$$

The function rms^2 is a function of two independent variables: T_0 (implicitly defined within $\tau(t)$) and r. By differentiating, optimization characteristics are easily determined.

$$[6.6] \qquad D_r(rms^2) = \frac{-2}{q} \cdot \sum_{k=1}^{q} [\ln(y_k)\cdot\tau(t_k) - r\cdot\tau^2(t_k)]$$

$$[6.7] \qquad D^2{}_r(rms^2) = \frac{2}{q} \cdot \sum_{k=1}^{q} \tau^2(t_k)$$

Equation [6.6] vanishes when

$$[6.8] \qquad r = \frac{\sum_{k=1}^{q} \ln(y_k)\cdot\tau(t_k)}{\sum_{k=1}^{q} \tau^2(tk)}$$

Since equation [6.7] is positive definite for a given T_0, rms^2 has a single absolute minimum. Therefore, we simply consider the character of rms^2 as a function of T_0 to complete this analysis.

While the above formalism might look forbidding, Eulerian integration leads to rather simple algorithms. These are discussed below and summarized as Algorithm 6.1.

i. The Physiological Time...

Suppose we are given the temperature profile $T(t)$, $t=1,2,\ldots,p$, initiation time t_0 and threshold T_0. Then

[6.9] $$\tau(t) = \int_{\Gamma(t)} [T(\mu)-T_0]d\mu$$

where $\Gamma(t) = \{t>t_0: T(t)>T_0\}$

is computed by Algorithm 6.1a.

ii. The Objective Function...

Suppose the data points

$$t_k, y_k \ , \quad k=1,2,\ldots,q.$$

are given. Then the Root Mean Square (rms) objective function is calculated using equation [6.5]. The rms is a measure of the residual per data point and can be thought of as the average modeling error. It is this quantity that must be minimized . Algorithm 6.1b computes the rms.

iii. The Population Model...

Assuming normalized data ($y_1=1$) we let $x_0=1$ too. Assuming an exponential model, equation [6.2] is solved by Algorithm 6.1c using Eulerian integration.

iv. Outputing the model results...

The population model may be compared against the empirical data by simply outputing the

model results and the source data. This is illustrated as Algorithm 6.1d.

As an illustration of applying these algorithms, consider the phenological data for a crop shown in Table 6.1. Seventeen data points that encompass the growth segment of a single season are shown. It is desired to model this data using the heat unit hypothesis. To begin, let us apply the algorithms outlined in Figure 6.1. This will, of course, require a simple Type I (single temperature threshold) model.

Algorithms 6.1 are applied on the first eleven (q=11) data points using an initiation time t_0=80 days and the temperature profile of 1985 for Lansing, Michigan (recall Figure 1.5). The results in Table 6.2 were achieved for various thresholds. This error is also plotted in Figure 6.1.

Notice the characterization of rms^2 as a function of T_0. Obviously there is one relative (and absolute) minimum of rms as a function of T_0. In this case the minimum occurs at threshold $T_0 \approx 40^o$ and $r \approx 0.001733$. In this way it is clearly a simple matter to graphically determine those parametric values that best fit the data.

Even so, while this graphical method may be satisfactory, we strive to automate the minimization process by using numerical optimization techniques.

(a.)
DO t=1,t_0 { $\tau(t) = 0$
DO t=t_0+1,p {
  $T(t)>T_0$ { $\tau(t) = \tau(t-1)+T(t)-T_0$
  $T(t)\leq T_0$ { $\tau(t) = \tau(t-1)$

(b.)
$s_1=0$
$s_2=0$
DO k=1,q {
  $s_1 = s_1 + \ln(y_k)\cdot\tau(t_k)$
  $s_2 = s_2 + \tau^2(t_k)$
$r=s_1/s_2$
$s_3=0$
DO k=1,q { $s_3 = s_3 + [\ln(y_k) - r\cdot\tau(t_k)]^2$
rms = $\sqrt{s_3/q}$

(c.)
DO t=1,t_0 { $x(t)=x_0$
DO t=t_0+1,p { $x(t)=x(t-1)\cdot(1+r\cdot[\tau(t)-\tau(t-1)])$

(d.)
DO k=1,q { PRINT k,t_k,y_k,$x(t_k)$

Algorithm 6.1: Computing critical elements for the Type I model.

a. deterministic physiological time,
b. Root Mean Square (rms) objective,
c. the [exponential] population model,
d. outputing the model results.

Table 6.1: Sample population data used in Type I model example.

k	t_k (days)	y_k (population)
1	100	1.000
2	110	1.030
3	117	1.060
4	122	1.210
5	129	1.780
6	141	2.630
7	150	4.319
8	156	4.784
9	161	6.919
10	172	9.860
11	183	12.927
12	192	15.167
13	201	20.095
14	211	27.925
15	229	29.649
16	246	31.046
17	253	31.299

Table 6.2: Results of simulations using Algorithms 6.1 on the data of Table 6.2 for 40 different temperature thresholds T_0.

T_0	r	rms
30.5	.001010289	.006603120
31.0	.001034551	.006440828
31.5	.001059961	.006272517
32.0	.001086597	.006098066
32.5	.001114546	.005917409
33.0	.001143902	.005730570
33.5	.001174769	.005537701
34.0	.001207260	.005339134
34.5	.001241498	.005135466
35.0	.001277618	.004927659
35.5	.001315483	.004721992
36.0	.001354861	.004526066
36.5	.001395760	.004343272
37.0	.001438296	.004175656
37.5	.001482592	.004025705
38.0	.001528745	.003896698
38.5	.001576857	.003792093
39.0	.001626966	.003715895
39.5	.001679221	.003670722
40.0	.001733785	.003658849
40.5	.001790793	.003682092
41.0	.001850394	.003741307
41.5	.001912745	.003836353
42.0	.001977991	.003966104
42.5	.002046243	.004128374
43.0	.002117790	.004321090
43.5	.002192848	.004541848
44.0	.002271650	.004788256
44.5	.002354449	.005058020
45.0	.002441521	.005349026
45.5	.002533166	.005659345
46.0	.002629694	.005987076
46.5	.002731329	.006329442
47.0	.002838601	.006686389
47.5	.002951940	.007056843
48.0	.003071820	.007439910
48.5	.003198756	.007835124
49.0	.003333320	.008242164
49.5	.003476144	.008660492
50.0	.003627934	.009089644

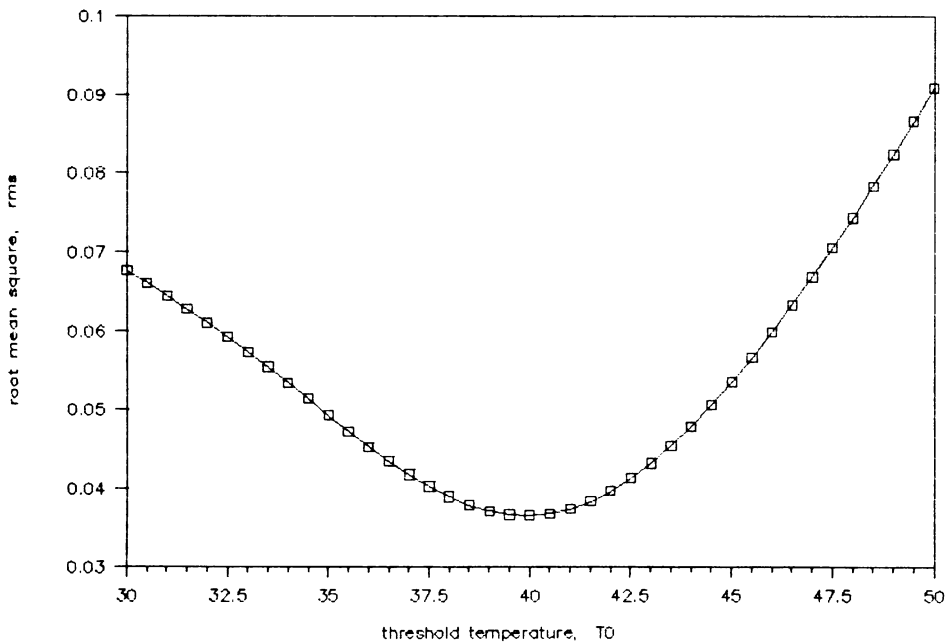


Figure 6.1: **Graphs of objective rms as a function of T_0, reproducing Table 7.5.**

LINE SEARCH METHODS

The basic Type I problem requires the identification of two parameters: T_0 and r. By equation [6.8] this can be reduced to only a single parameter T_0. The resulting objective (equation [6.5]) is clearly nonlinear since T_0 doesn't even occur explicitly in the formula! Also, the expression for the (total) derivative $d[rms^2]/dT_0$ is unmanageable.

Thus, we choose to implement an optimization algorithm from a class of nonlinear, unconstrained, single variable optimization methods called line searching without derivatives. In particular, we describe and implement the so called **golden section** method (Bazaraa, Shetty 1979).

This golden section method can be shown to be quite efficient. It, like all line searches, requires the analyst to provide an initial search **interval of uncertainty** $[a_1,b_1]$ within which the minimum of on objective function F is known to reside. The algorithm proceeds by establishing a sequence of refined internals $[a_k,b_k]$ which are nested within the old intervals:

$$[a_k,b_k] \quad [a_{k+1},b_{k+1}]$$

and continues until the interval length is smaller than a pre-specified **length of uncertainty** len.

[6.10] $$b_k - a_k < \text{len}$$

Formally this is done as shown in Algorithm 6.2.

Each of these algorithms have been implemented. Using the data in Table 6.1 and q=11 data points, the following parameter values were obtained:

$q = 11$

$T_0 = 38.75$

$r = 0.001598$

rms = 0.3198

They are consistent with those found by graphical approximation from Figure 6.1. The model results are tabulated and sketched in Table 6.5 and Figure 6.4 respectively.

Certainly this example could have been done using a different number of data points. The choice of q=11 was made simply because the exponential assumption breaks down at extreme times due to environmental effects. This presents the analyst with a fundamental methodological question: How much (chronological) time elapses until the

endogenous phase is complete? A suitable answer to this general question will require expert biological judgement and further research. However, further examples are yet illustrative.

The golden section algorithm was implemented (in the form a computer program) and executed for various numbers of data points: q=5,6,...,17. Tables 6.3, 6.4, 6.5, 6.6 and 6.7 tabulate results for q=5,8,11,14 and 17 respectively while Figures 6.2, 6.3, 6.4, 6.5 and 6.6 sketch the empirical population data against that predicted by the model.

In order to observe the dynamics at work, a summary of results are tabulated for each of the runs at q=5,6,...,17. These are listed in Table 6.8. Graphs for the optimal objectives (rms), growth rates (r) and temperature thresholds (T_0) as functions of data horizon (q) are plotted in Figures 6.7, 6.8 and 6.9.

It is easy to see that the very initial data achieves an excellent fit (q=5, rms=0.009216) while the model totally breaks down at the q=17 (rms=1.7605). "Eyeballing" the graphs confirm this fact. Intermediate q-values show reasonable fits with moderate rms-values.

Figure 6.7 plots the optimal rms as a function of q.

Ignoring q=5, it will be noticed that there are three rather distinct stages. Explanation of these stages follow.

Stage 1: Endogenous: $6 \leq q \leq 9$ ...

Choosing q over this interval results in significantly low rms-values. Figure 6.7 also shows an excellent model fit. Clearly in this region we see purely endogenous effects. The corresponding thresholds (Figure 6.9) vary between 43^{o} and 45^{o}. The exponential model is excellent.

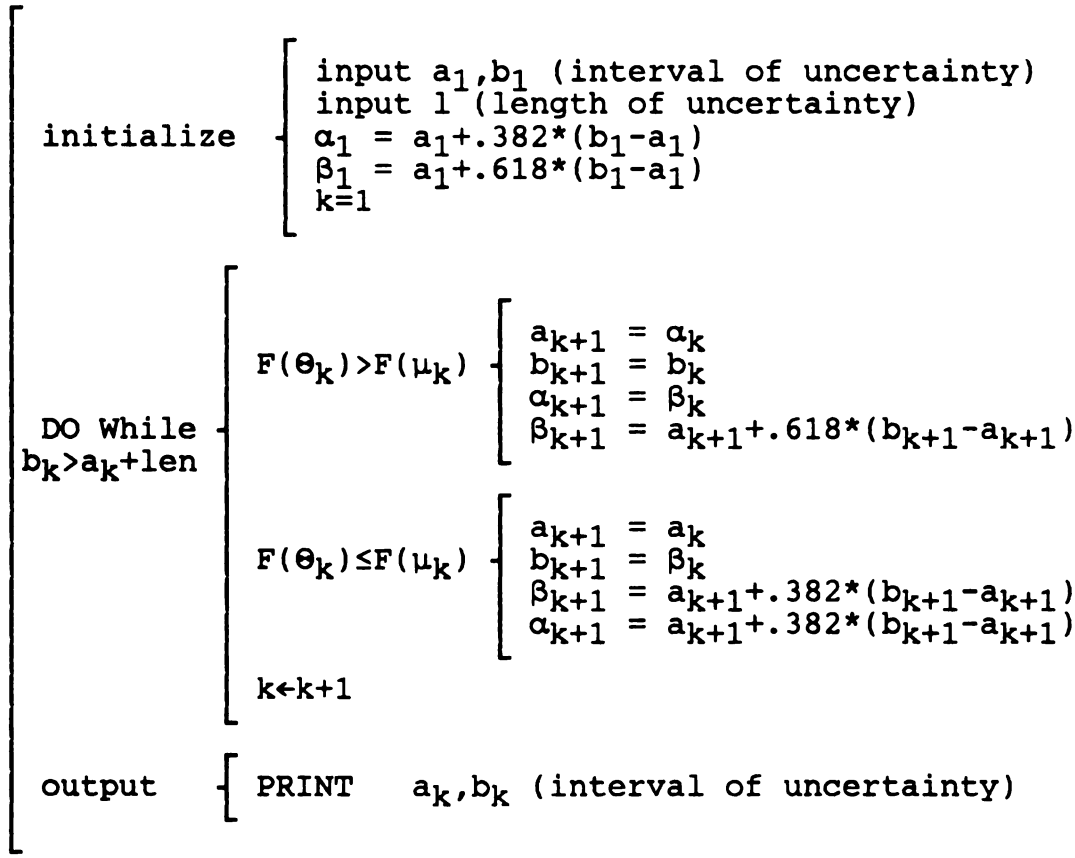
Stage 2: Transition: $10 \leq q \leq 14$...

Choosing q on this interval shows increasing rms-values. While there is not a total break-down in the exponential model, parameters evaluated under this assumption will be somewhat suspect. Probably we see the increasing results of environmental density effects, but no inter-species interaction.

Stage 3: Endogenous: $q \geq 15$...

Here we observe the effects of model incompatibility. The large rms-values and corresponding absurd threshold-values (T_0=3.70° F!) indicate that the exponential model is no longer viable.

Again, it is the responsibility of the analyst to determine the data interval appropriate for the endogenous stage.



Algorithm 6.2: The Golden Section Algorithm

Table 6.3 Optimal results using the first five data points of Table 6.1.

t_k	empirical data	model prediction
100	1.000	1.000000
110	1.030	1.000000
117	1.060	1.033089
122	1.210	1.194482
129	1.780	1.763921

Table 6.4 Optimal results using the first eight data points of Table 6.1.

t_k	empirical data	model prediction
100	1.000	1.000000
110	1.030	1.046386
117	1.060	1.157074
122	1.210	1.286690
129	1.780	1.564282
141	2.630	2.464890
150	4.319	3.795257
156	4.784	5.257814

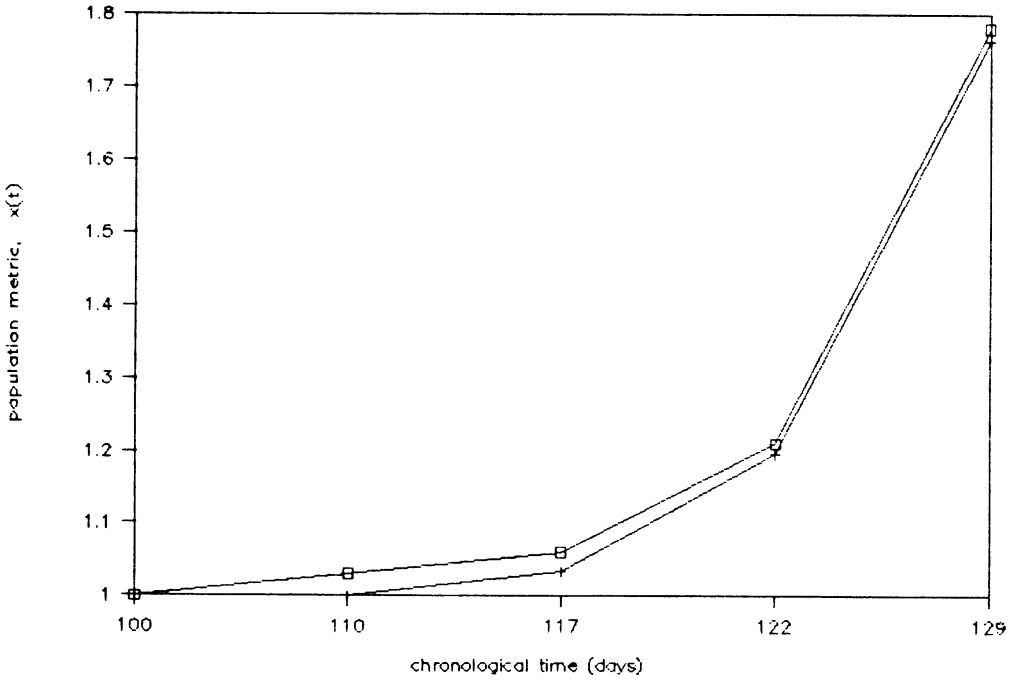


Figure 6.2 Graph comparing empirical data against modelled population metric. q=5.

- • empirical data
- + model prediction

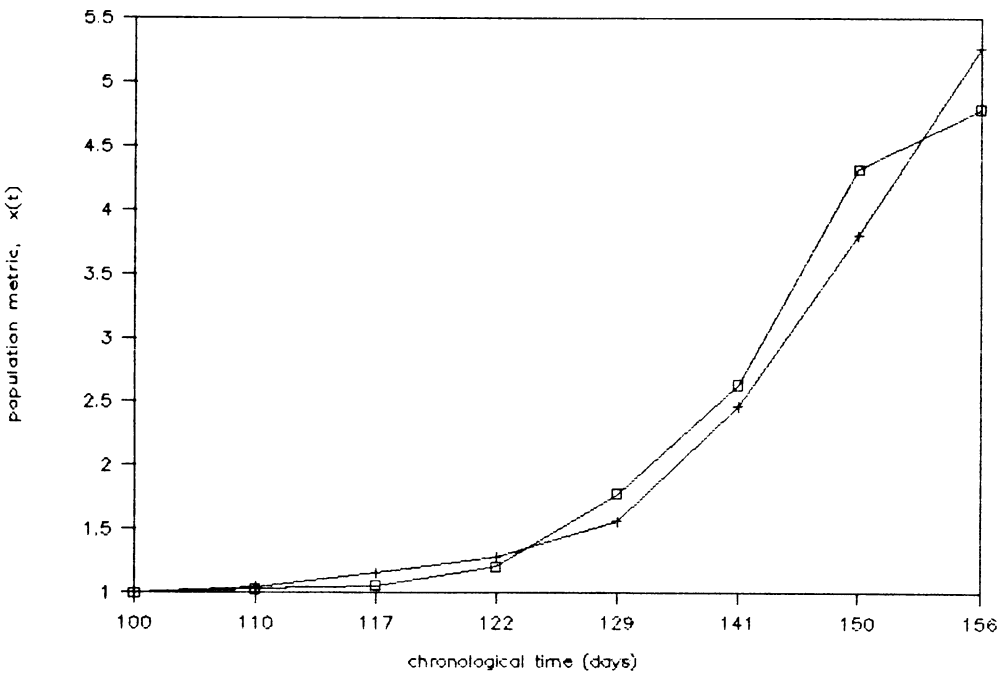


Figure 6.3 Graph comparing empirical data against modelled population metric. q=8.

▪ empirical data

\+ model prediction

Table 6.5 Optimal results using the first eleven data points of Table 6.1.

t_k	empirical data	model prediction
100	1.000	1.035493
110	1.030	1.150449
117	1.060	1.291578
122	1.210	1.432458
129	1.780	1.703675
141	2.630	2.467717
150	4.319	3.444365
156	4.784	4.404234
161	6.919	5.475366
172	9.860	9.154198
183	12.927	15.882200

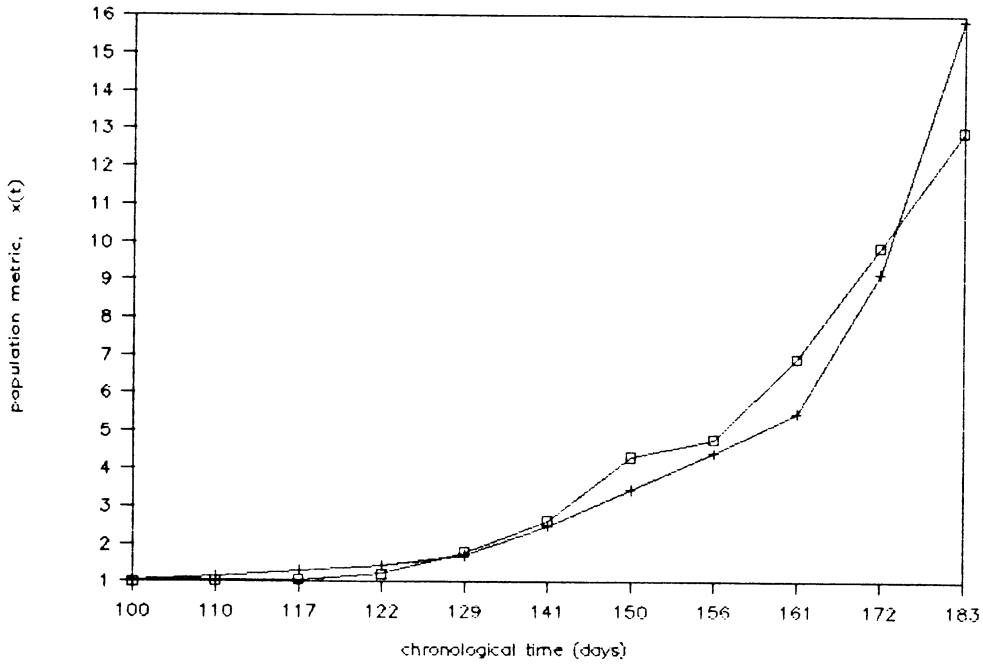


Figure 6.4 Graph comparing empirical data against modelled population metric. q=11.

▪ empirical data

\+ model prediction

Table 6.6 Optimal results using the first fourteen data points of Table 6.1.

t_k	empirical data	model prediction
100	1.000	1.144691
110	1.030	1.306769
117	1.060	1.471703
122	1.210	1.623112
129	1.780	1.895012
141	2.630	2.587265
150	4.319	3.383804
156	4.784	4.107133
161	6.919	4.865857
172	9.860	7.222898
183	12.927	10.975120
192	15.167	15.625620
201	20.095	22.340710
211	27.925	33.183070

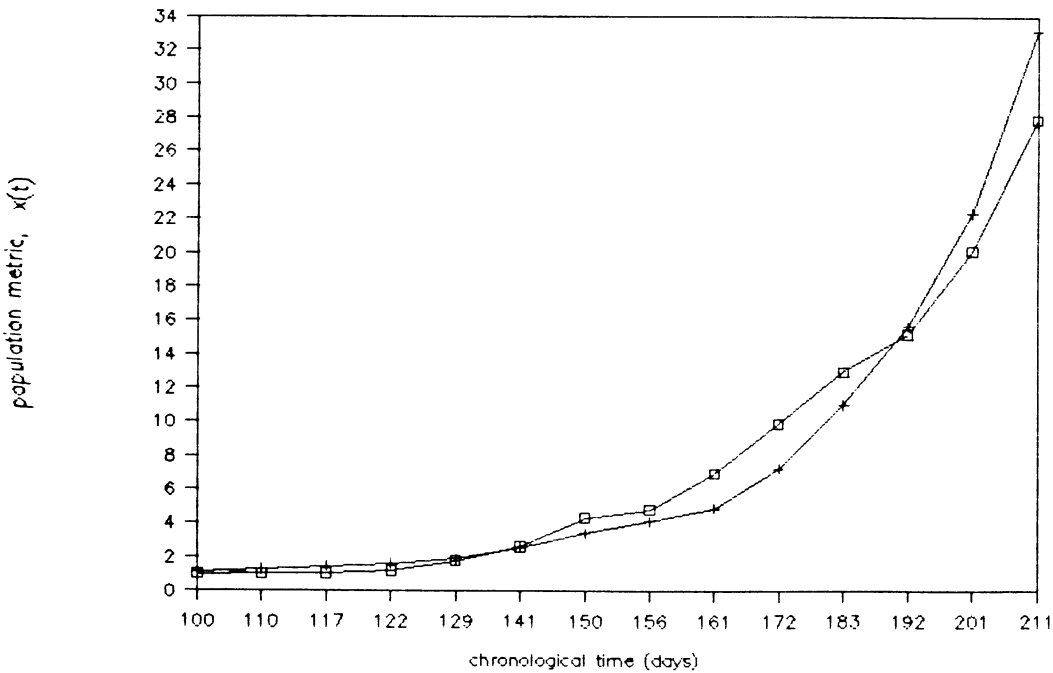


Figure 6.5 Graph comparing empirical data against modelled population metric. q=14.

• empirical data

\+ model prediction

Table 6.7 Optimal results using the seventeen data points of Table 6.1.

t_k	empirical data	model prediction
100	1.000	1.326888
110	1.030	1.568573
117	1.060	1.782085
122	1.210	1.962403
129	1.780	2.262030
141	2.630	2.939783
150	4.319	3.629052
156	4.784	4.201246
161	6.919	4.761976
172	9.860	6.329595
183	12.927	8.493454
192	15.167	10.852020
201	20.095	13.889350
211	27.925	18.259130
229	29.649	29.551470
246	31.046	45.405370
253	31.299	53.649530

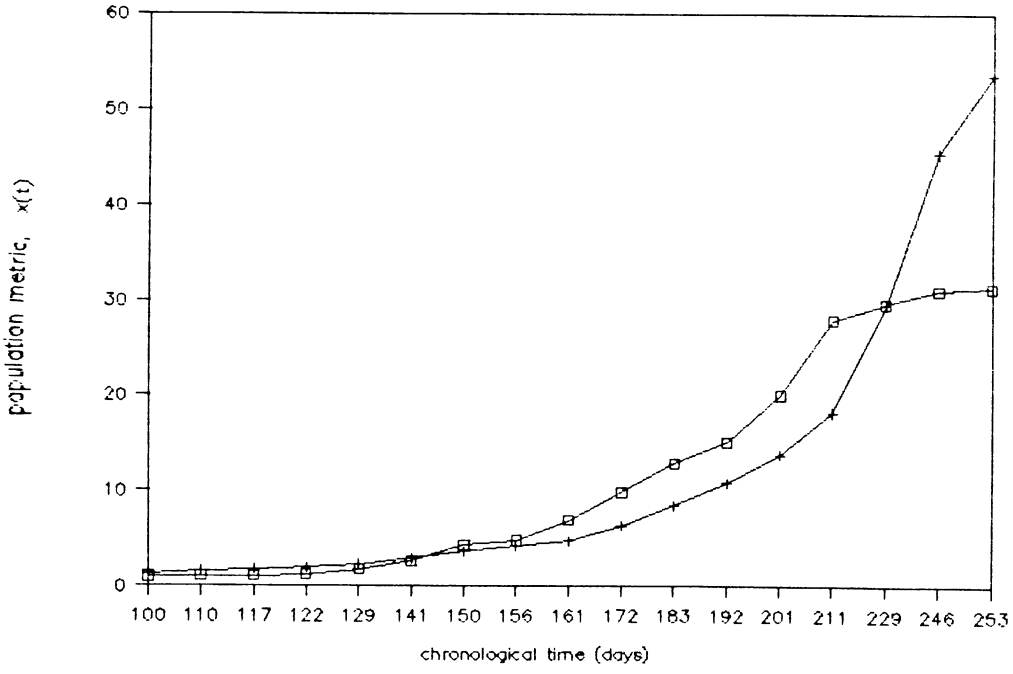


Figure 6.6 Graph comparing empirical data against modelled population metric. q=17.

- ▪ empirical data
- \+ model prediction

Table 6.8 Optimal rms values along with the corresponding T_0 and r values as a function of number of data points q.

q	T_0	r	rms
5	49.09	.010695050	.009216376
6	45.44	.003338706	.037631340
7	44.95	.003254177	.034300950
8	43.68	.002628307	.095865640
9	43.37	.002421788	.081618940
10	41.61	.002073451	.170504000
11	38.75	.001598437	.319811500
12	36.69	.001301794	.449267300
13	34.62	.001126551	.492408300
14	32.98	.001018055	.525416600
15	26.66	.000774287	1.271429000
16	16.32	.000540940	1.759235000
17	3.74	.000400562	1.760493000

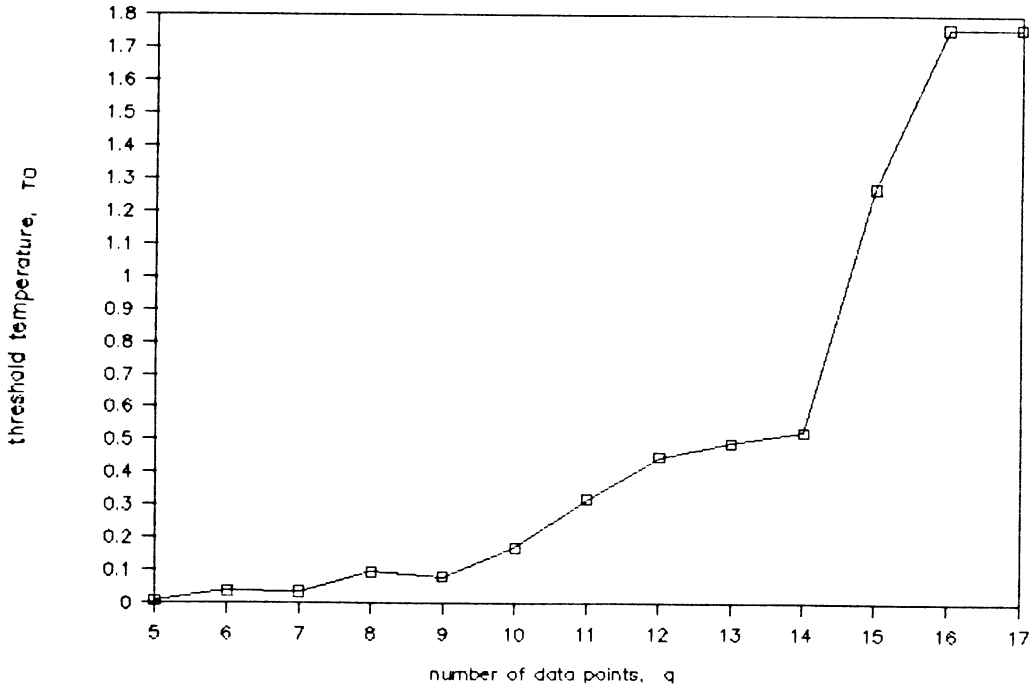


Figure 6.7: Plot of objective rms against horizon time q.

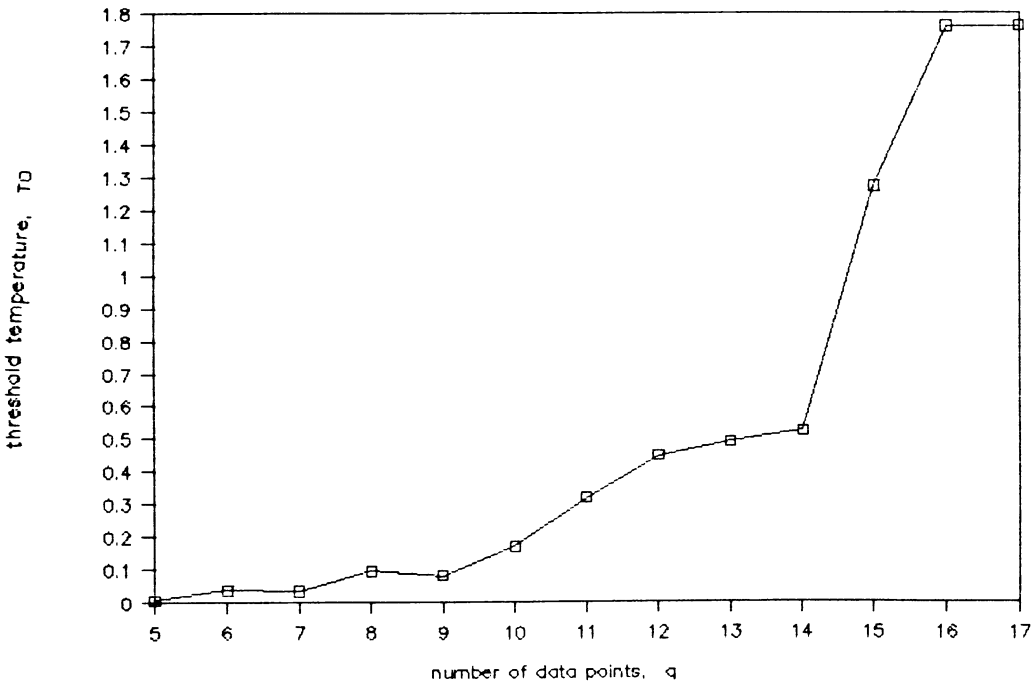


Figure 6.7: Plot of objective rms against horizon time q.

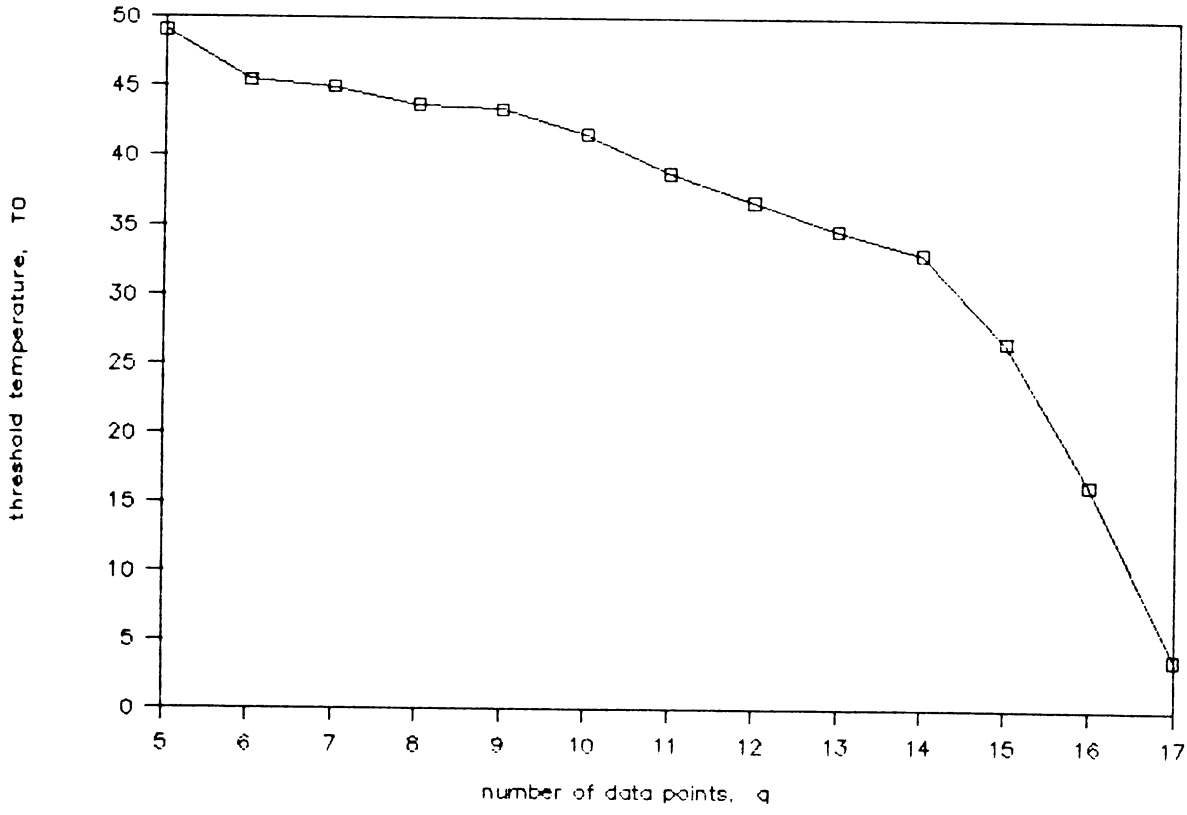


Figure 6.8: Plot of optimal threshold T_0 against horizon time q.

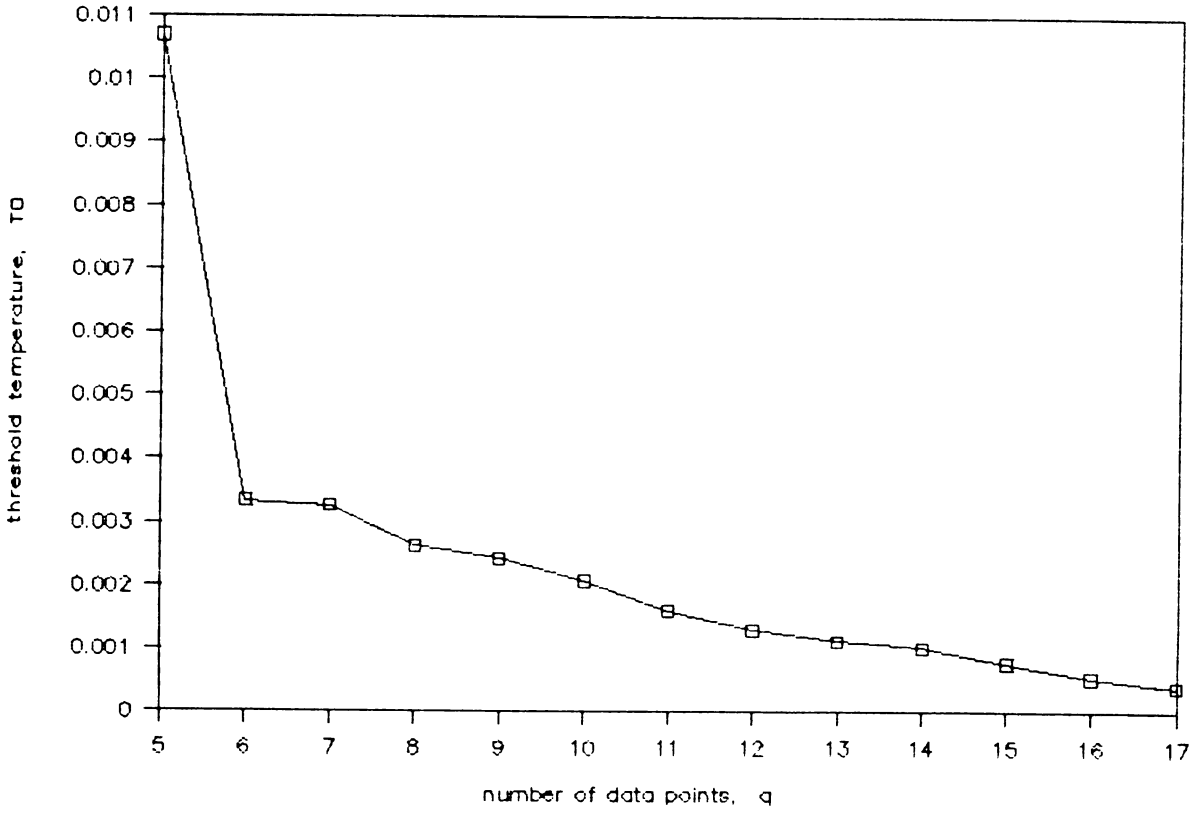


Figure 6.9 Plot of optimal rate r against horizon time q.

THE ENDOGENOUS TYPE III MODEL

A most useful generalization to the fundamental type I heat unit model is that of type III. Here we assume the existence of an upper threshold T_1 as well as the usual lower threshold T_0. Solving this model is similar to that of the type I.

Mathematically, the only difference between these two models is the action set:

[6.11] $$\Gamma(t) = \{t>t_0: T_0< T(t)\leq T_1\}$$

The expression for r is still unchanged. Therefore, by eliminating r from the list of parameters, <u>two</u> thresholds, T_0 and T_1, remain to be determined. For this reason, the routine to evaluate the objective function is very similar. However, rather than a simple line search, we must use a two dimensional (n=2) approach to this problem..

The physiological time evaluation, done by a δ=1 Euler approximation, proceeds by Algorithm 6.3. The objective rms function evaluation, the population model, etc. are still handled as in Algorithm 6.1.

$$
\left[\begin{array}{l}
\text{DO } t=1,t_0 \;\{\; \tau(t)=0 \\
\\
\text{DO } t=t_0+1,p \left\{\begin{array}{ll}
T(t)<T_0 & \{\; \tau(t)=\tau(t-1) \\
T_0 \le T(t)<T_1 & \{\; \tau(t)=\tau(t-1)+T(t)-T_0 \\
T(t)\ge T_1 & \{\; \tau(t)=\tau(t-1)
\end{array}\right.
\end{array}\right.
$$

Algorithm 6.3 Computing the physiological time for a Type III model.

THE POWELL OPTIMIZATION PROCEDURE

For multidimensional optimization (n>1) problems, algorithms become slightly more complex. In addition to determining a step size, it is necessary to determine a new direction at each iteration. The ability to do this efficiently is greatly enhanced if one has knowledge of the objective function's derivative. Without the knowledge, methods for convergence do exist; however, their performance is considerably less. Unfortunately, our problem does not possess convenient derivatives.

One of the solution classes of the multi-variable problem is the so-called direction set methods as put forth by M. J. D. Powell (Press, et al. 1986). From a given initial point, each stage determines a direction and a distance to travel (using line searches) along that direction. No derivatives are required and there are no constraints on the solution space.

As in the golden ratio line search, the analyst must specify certain search parameters. These include the following:

i. dimension of the search space n.
ii. the initial search point $x_1(0),\ldots,x_n(0)$.
iii. the Cauchy convergence limit $e_1,\ldots,e_n$.

The algorithm stops when

[6.12] $$|x_k(i)-x_k(i-1)|<e_k \qquad \text{for } k=1,\ldots,n$$

The logical details are rather complicated, but the algorithm is outlined as a high level diagram as Algorithm 6.4. This has been implemented using a procedure drawn from Kuester and Mize (1973). In using this program the model parameters, search parameters and operation parameters are defined below.

Model Parameters:

number of parameters to be determined: n=2
initialization time: t_0=80
number of time steps: p=365
number of data points: q=17

Operation Parameters:

print option: print search point: x(1),x(2),
objective: rms,
number of function evaluations.

IPRINT=1: after each variable search.
IPRINT=2: after each iteration.
IPRINT=3: after optimum has been found.

maximum iterations before termination: MAXIT=999.

maximum step multiplier each search: ESCALE=2.

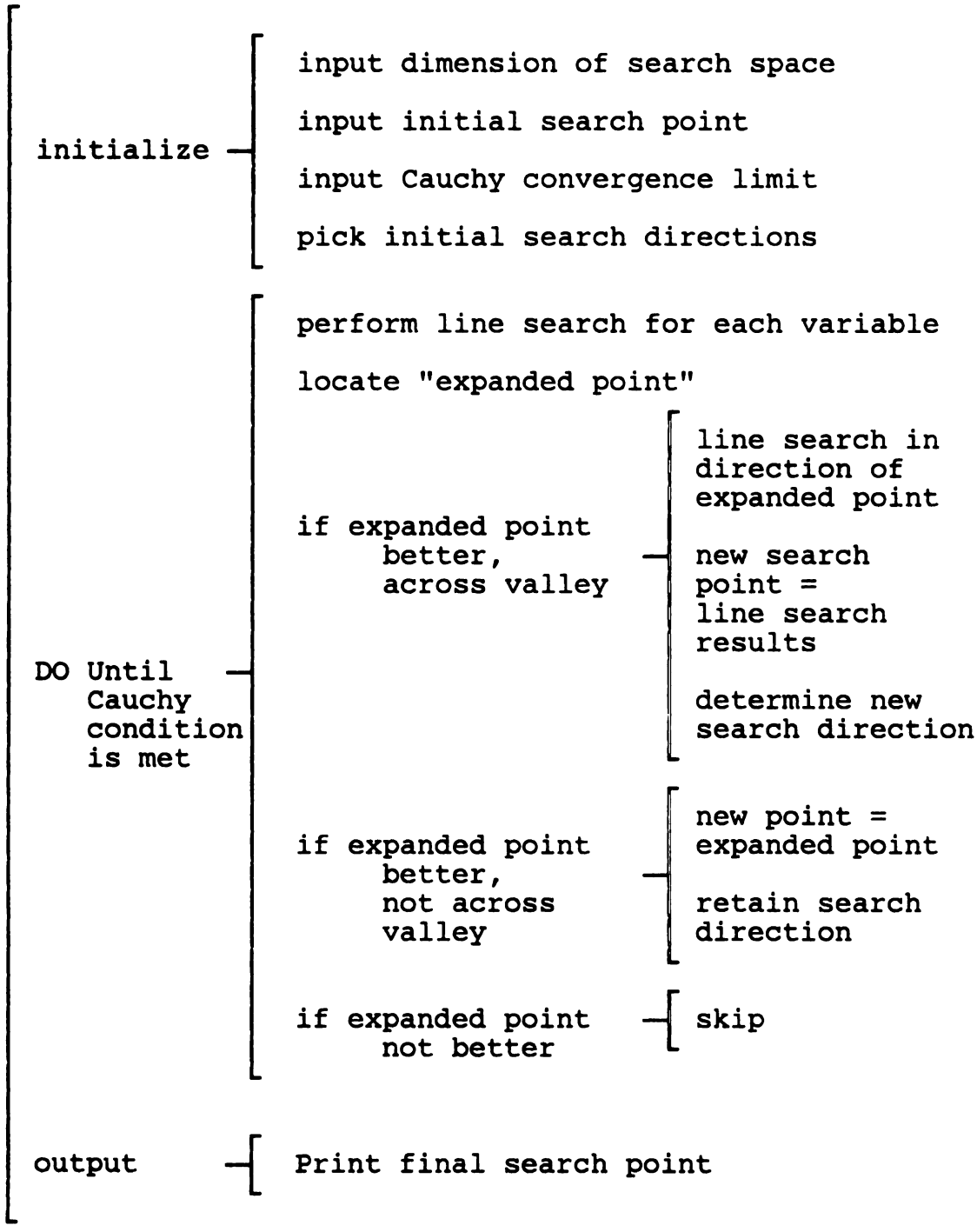
Search Parameters:

initial search point: x(1)=50, x(2)=50
convergence limit for point: e(1)=.5, e(2)=.5

Again using the data in Table 6.1, the Powell algorithm was applied for q=11. The results produce...

$q = 11$
$T_0 = 36.07$
$T_1 = 84.93$
$r = 0.002376$
rms = 0.2334

It should be noticed that the rms-value is indeed superior to that of the Type I model. Not surprisingly, the Type III model leads to a better fit, since there are now two parameters rather than one. This will provide an extra degree of freedom, and in general provide superior results.



Algorithm 6.4 Computing physiological times for multi-species population Type I models.

ENDOGENOUS TYPE II AND IV MODELS

It is possible to apply a similar procedure to other endogenous models. Here we demonstrate an approach to the Type II problem.

We begin by decomposing the physiological time as equation [5.7].

$$[6.13] \qquad \tau(t) = \int_{\Gamma_0(t)} [T(\mu)-T_0]d\mu + \int_{\Gamma_1(t)} [T(\mu)-T_0]d\mu$$

$$= \frac{1}{r_0} \cdot \int_{\Gamma_0(t)} \beta(T(\mu))d\mu + \frac{1}{r_1} \cdot \int_{\Gamma_1(t)} \beta(T(\mu))d\mu$$

$$= \tau_0(t) + \tau_1(t)$$

with action sets:

$$[6.14] \qquad \Gamma_0(t) = \{t>t_0: T_0<T((t)<(r_0T_0-r_1T_1)/(r_0-r_1)$$

$$\Gamma_1(t) = \{t>t: (r_0T_0-r_1T_1)/(r_0-r_1)<T(t)<T_1\}$$

and chronological times:

$$\tau_0(t) = \int_{\Gamma_0(t)} [T(\mu)-T_1]d\mu$$

[6.15]

$$\tau_1(t) = \int_{\Gamma_1(t)} [T(\mu)-T_1]d\mu$$

and biological activity

[6.16] $$\int_{\Gamma_0(t)\cup\Gamma_1(t)} \beta(T(t))d\mu = r_0\tau_0(t) + r_1\tau_1(t)$$

By equation [6.3],

[6.17] $$rms^2 = \frac{1}{q} \cdot \sum_{k=1}^{q} \left[\ln(y_k) - \int_{\Gamma_0(t)\cup\Gamma_1(t)} \beta(T(\mu))d\mu \right]^2$$

$$= \frac{1}{q} \cdot \sum_{k=1}^{q} \left[\ln(y_k) - r_0\tau_0(t_k) - r_1\tau_(t_k)\right]^2$$

Equating the partial derivatives to zero yields...

$$\sum_{k=1}^{q} [\ln(y_k) - r_0\tau_0(t_k) - r_1\tau_1(t_k)]\cdot\tau_0(t_k) = 0$$

[6.18]

$$\sum_{k=1}^{q} [\ln(y_k) - r_0\tau_0(t_k) - r_1\tau_1(t_k)]\cdot\tau_1(t_k) = 0$$

Using matrix notation,

$$\begin{bmatrix} \sum_{k=1}^{q} \tau_0^2(t_k) & \sum_{k=1}^{q} \tau_0(t_k)\cdot\tau_1(t_k) \\ \sum_{k=1}^{q} \tau_0(t_k)\cdot\tau_1(t_k) & \sum_{k=1}^{q} \tau_1^2(t_k) \end{bmatrix} \cdot \begin{bmatrix} r_0 \\ r_1 \end{bmatrix}$$

[6.19]

$$= \begin{bmatrix} \sum_{k=1}^{q} \ln(y_k)\cdot\tau_1(t_k) \\ \sum_{k=1}^{q} \ln(y_k)\cdot\tau_1(t_k) \end{bmatrix}$$

Explicit expressions for r_0 and r_1 are given by

$$r_0 = \frac{[\sum_{k=1}^{q} \tau_1^2(t_k)]\cdot[\sum_{k=1}^{q} \ln(y_k)\cdot\tau_0(t_k)] - [\sum_{k=1}^{q} \tau_0(t_k)\cdot\tau_1(t_k)]\cdot[\sum_{k=1}^{q} \ln(y_k)\cdot\tau_1(t_k)]}{[\sum_{k=1}^{q} \tau_1^2(t_k)]\cdot[\sum_{k=1}^{q} \tau_1^2(t_k)] - [\sum_{k=1}^{q} \tau_0(t_k)\cdot\tau_1(t_k)]^2}$$

[6.20]

$$r_1 = \frac{[\sum_{k=1}^{q} \tau_0^2(t_k)]\cdot[\sum_{k=1}^{q} \ln(y_k)\cdot\tau 1(t_k)] - [\sum_{k=1}^{q} \tau_0(t_k)\cdot\tau_1(t_k)]\cdot[\sum_{k=1}^{q} \ln(y_k)\cdot\tau_0(t_k)]}{[\sum_{k=1}^{q} \tau_1^2(t_k)]\cdot[\sum_{k=1}^{q} \tau_1^2(t_k)] - [\sum_{k=1}^{q} \tau_0(t_k)\cdot\tau_1(t_k)]^2}$$

While four parameters need be deduced, equations [6.20] reduces this to only the two thresholds T_0 and T_1. Therefore, the solution is very similar to the type III model: Apply the Powell procedure with n=2.

Similarly, a type IV model with 5 parameters can be reduced to finding only the high and low thresholds T_0 and T_1. This is true since there are now three relationships for rates r_0, r_1 and r_2.

Again we consider the expression for physiological time:

[6.21] $$\tau(t) = \int_{\Gamma_0(t)} [T(\mu)-T_0]d\mu + \int_{\Gamma_1(t)} d\mu + \int_{\Gamma_2(t)} [T(\mu)-T_0]d\mu$$

with action sets...

$$\Gamma_0(t) = \{t>t_0: T_0<T(t)\leq r_1/r_0+T_0\}$$

[6.22] $$\Gamma_1(t) = \{t>t_0: r_1/r_0+T_0<T(t)\leq r_1/r_2+T_3\}$$

$$\Gamma_2(t) = \{t>t_0: r_1/r_2+T_3<T(t)\leq_{T3}\}$$

and biological activity

[6.23] $$\int_{\Gamma_0 U \Gamma_1 U \Gamma_2} \beta(T(\mu))d\mu = r_0 \cdot \tau_0(t) + r_1 \cdot \|\Gamma_1(t)\| + r_2 \cdot \tau_2(t)$$

By equation [6.3],

[6.24]
$$rms^2 = \frac{1}{q} \cdot \sum_{k=1}^{q} \left[\ln(y_k) - r_0 \cdot \tau_0(t_k) - r_1 \cdot \|\Gamma_1(t_k)\| - r_2 \cdot \Gamma_2(t_k) \right]^2$$

Equating partial derivatives to zero results in...

$$r_0 \cdot \sum_{k=1}^{q} \tau_0^2 + r_1 \cdot \sum_{k=1}^{q} \tau_0 \cdot \|\Gamma_1\| + r_2 \cdot \sum_{k=1}^{q} \tau_0 \cdot \tau_1 = \sum_{k=1}^{q} LN(y_k) \cdot \tau_0$$

[6.26]
$$r_0 \cdot \sum_{k=1}^{q} \tau_0 \|\Gamma_1\| + r_1 \cdot \sum_{k=1}^{q} \|\Gamma_1\|^2 + r_2 \cdot \sum_{k=1}^{q} \tau_2 \|\Gamma_1\| = \sum_{k=1}^{q} LN(y_k) \cdot \|\Gamma_1\|$$

$$r_0 \cdot \sum_{k=1}^{q} \tau_0 \cdot \tau_2 + r_1 \cdot \sum_{k=1}^{q} \tau_2 \cdot \|\Gamma_1\| + r_2 \cdot \sum_{k=1}^{q} \tau_2^2 = \sum_{k=1}^{q} LN(y_k) \cdot \tau_2$$

or by expressing in matrix notation,

$$\begin{bmatrix} \sum_{k=1}^{q} \tau_0^2 & \sum_{k=1}^{q} \tau_0 \cdot \|\Gamma_1\| & \sum_{k=1}^{q} \tau_0 \cdot \tau_1 \\ \sum_{k=1}^{q} \tau_0 \|\Gamma_1\| & \sum_{k=1}^{q} \|\Gamma_1\|^2 & \sum_{k=1}^{q} \tau_2 \|\Gamma_1\| \\ \sum_{k=1}^{q} \tau_0 \cdot \tau_2 & \sum_{k=1}^{q} \tau_2 \cdot \|\Gamma_1\| & \sum_{k=1}^{q} \tau_2^2 \end{bmatrix} \cdot \begin{bmatrix} r_0 \\ r_1 \\ r_2 \end{bmatrix}$$

[6.27]

$$= \begin{bmatrix} \sum_{k=1}^{q} LN(y_k) \cdot \tau_0 \\ \sum_{k=1}^{q} LN(y_k) \cdot \|\Gamma_1\| \\ \sum_{k=1}^{q} LN(y_k) \cdot \tau_2 \end{bmatrix}$$

The above system of equations can be solved explicitly for r_0, r_1 and r_2. As with models types II and III, this reduces to finding T_0 and T_3 (two parameters) and can be accomplished via the Powell algorithm.

This same procedure can even be applied to the more general piecewise linear model described by equation [5.10]. While this would be an interesting mathematical diversion, the

usefulness of this approach is of questionable value. Indeed, current practice seems to use model types I and III exclusively. Until the requisite precision is achieved, there is little need to apply the higher order endogenous types. More importantly, there is little intuitive justification for these multi-threshold models. Therefore, we have held back to only explicit solutions for types I and III.

As was illustrated in the data of 6.1, the use of Types I and III seem accurate and quite sufficient. The far more important question is the methodological one: How much time constitutes the endogenous phase? The answer is not precise, although a promising approach was indicated in the prior analysis.

In any case, once the model type has been defined and the endogenous time horizon understood, the endogenous parameters can be determined and tabulated. If the analysis has been well performed, these parametric values should be fixed. Repeated experiments should produce the similar results. Although this has not been done, it is the hope of the author that continuing research in this area will result in standardized tables of endogenous parameters for biological entities.

EXOGENOUS MODELS

Exogenous parameters can also be determined by using these optimization methods. Even in the most general case where the form of the model is not specified before hand, both the line search and Powell algorithms can be used to advantage.

Since there is no mathematical advantage to minimizing rms^2 on the natural logarithm of the residuals as in equation [6.5], we consider rms^2 as defined by equation [7.1]. Which ever means is used for computing rms^2, the differences between these two approaches will yield very small differences.

THE LOGISTIC MODEL

The logistic model was described earlier. Using the Type I endogenous model we may assume that parameters r and T_0 have already been determined by the endogenous methods outlined above. It remains but to determine the carrying capacity x_m. Since this is the only unknown parameter, we use the line search technique. The program procedure follows.

i. Input

modeling parameters:

threshold temperature: T_0
growth rate: r
initial time: t_0
initial population: x_0

search parameters:
left search point: a
right search point: b
length of uncertainty: len

external source files:
temperature profile: temp.dat
phenological history: phenol.dat

operation parameters: none

ii. Compute physiological time $\tau(t)$, t=1,...,p.

For Type I models this may be accomplished using Algorithm 6.1. As before, Type III models require the adjustment given in Algorithm 6.3. Physiological times for other model types are handled as in the development shown above.

iii. Line search routine (determines x_m using an rms objective function subroutine)

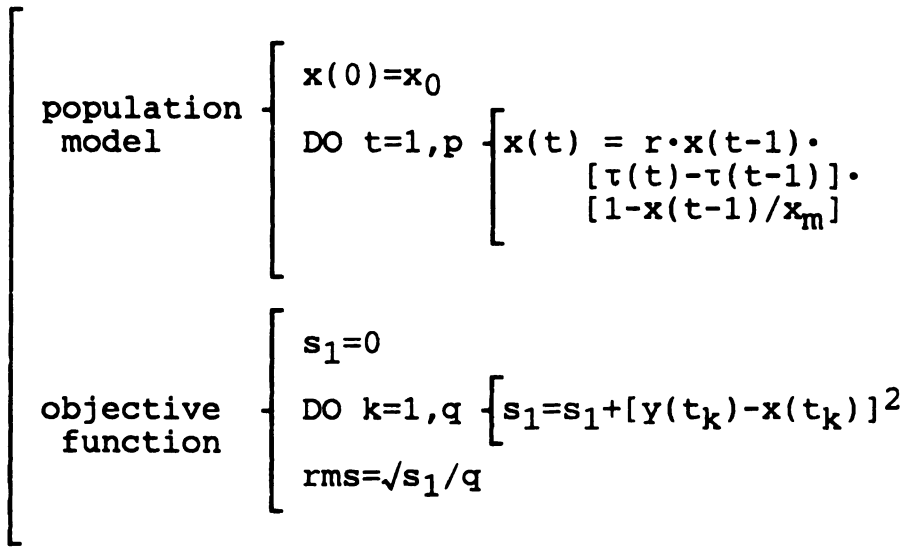
The Golden Section algorithm can still be used.

iv. Objective function subroutine
(input: independent variable: x
output: root mean square: rms)

Algorithm 6.5 is applicable for the logistic model. Obvious extensions exist for other basic models.

Execution of this procedure produces the best value of x_m=187. Using this value along with threshold T_0=38.75° and rate r=0.001598 (determined earlier through endogenous methods) produces rather good results.

The population x(t) is plotted for the annual time horizon and compared to the experimental values t_k, y_k in Figure 6.8. As can be observed, the fit is quite reasonable. Since the rms-value of rms=.3080 shows this to be the best fitting model of the logistic form, we are satisfied that the logistic heat-unit model is most reasonable for the given data.



Algorithm 6.5 Calculating the Logistic population model and objective function for any exogenous model.

THE LOTKA-VOLTERRA MODEL

We now consider the case of two interacting species. Classically, assuming a predator/prey relationship, this leads to the Lotka-Volterra equations:

$$\frac{dx_1(t)}{dt} = r_1 \cdot x_1(t) \cdot \left[1 - \frac{x_2(t)}{x_{1m}}\right]$$

[6.28]

$$\frac{dx_2(t)}{dt} = r_2 \cdot x_2(t) \cdot \left[1 - \frac{x_1(t)}{x_{2m}}\right]$$

where $x_1(t_{10}) = x_{10}$,

$x_2(t_{20}) = x_{20}$

and r_1, r_2 are endogenous growth rate parameters and are exogenous inter-species action parameters. Assuming a type-1 heat-unit model, each species will also have lower temperature thresholds T_1 and T_2.

The solution technique for this system proceeds very much as before. Parameters r_1, r_2, T_1 and T_2 are computed assuming exponential models for each population. Once these are determined, x_{1m} and x_{2m} can be found using Powell. A very workable methodology follows.

i. Physiological times $\tau_1(t)$, $\tau_2(t)$.

Given the temperature profile T(t), a Type I model and initial times t_{10}, t_{20}, Algorithm 6.6 computes the physiological time for t=1,2,...,p.

ii. Objective functions for each species' endogenous phase.

Given data points and time horizons for each species' endogenous phase, r_1, r_2, T_1, T_2 are found assuming the Type I exponential model.

species 1: (t_{1k},y_{1k}), $k=1,\ldots,q_{11}$
species 2: (t_{2k},y_{2k}), $k=1,\ldots,q_{21}$

This is shown as Algorithm 6.7.

Using the line search algorithm to minimize the objective function rms for each species, T_1 and T_2 are determined. Equation [6.8] (or its equivalent) calculates the rates r_0 and r_1.

iii. Objective function for the entire system (both endogenous and exogenous phases):

Given data points for each species:

species 1: (t_{1k},y_{1k}), $k=1,\ldots,q_1$
species 2: (t_{2k},y_{2k}), $k=1,\ldots,q_2$

This objective function is based on equation [7.1] rather than its logarithmic counterpart. There is no longer any mathematical justification to use equation [7.2] and the results are very nearly identical. Since there is no analytical assist here, the price to be paid is longer execution time.

Algorithm 6.5 evaluates [any] rms objective function.

iv. Using the Powell optimization procedure to minimize the rms objective, parameters x_{1m}, x_{2m} can also be determined. This is done using the following input parameters and files.

Input parameters, files

Modeling parameters...

growth rates: r_1, r_2
temperature thresholds: T_1, T_2
initial times: t_{10}, t_{20}
initial populations: x_{10}, x_{20}
number of exogenous parameters: n=2
number of time steps: p=365
number of data points: q_1, q_2

Operation parameters...

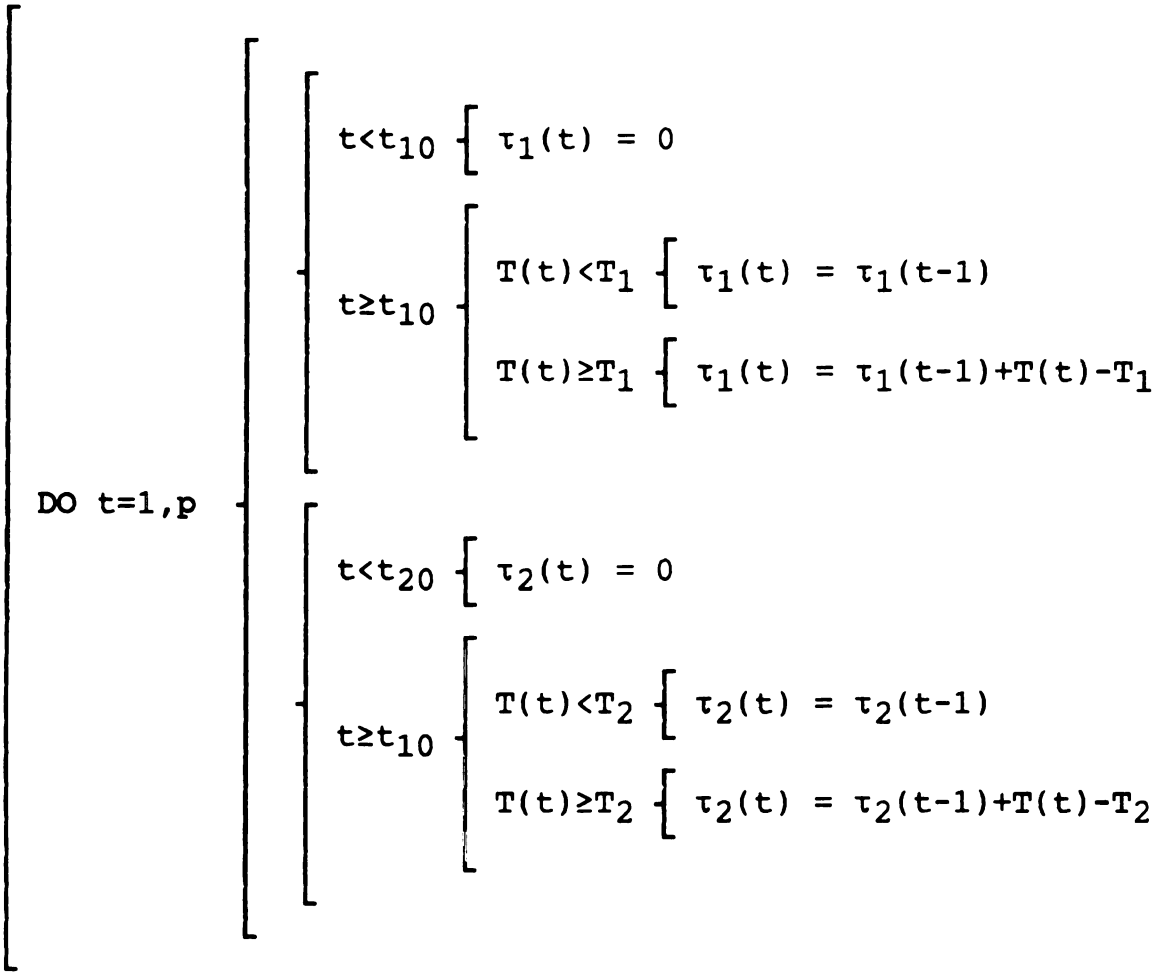
IPRINT=1
MAXIT=999
ESCALE=5

Search parameters...

initial search point: (x(1)=10, x(2)=10
convergence limit point: e(1)=1, e(2)=1

External source files...

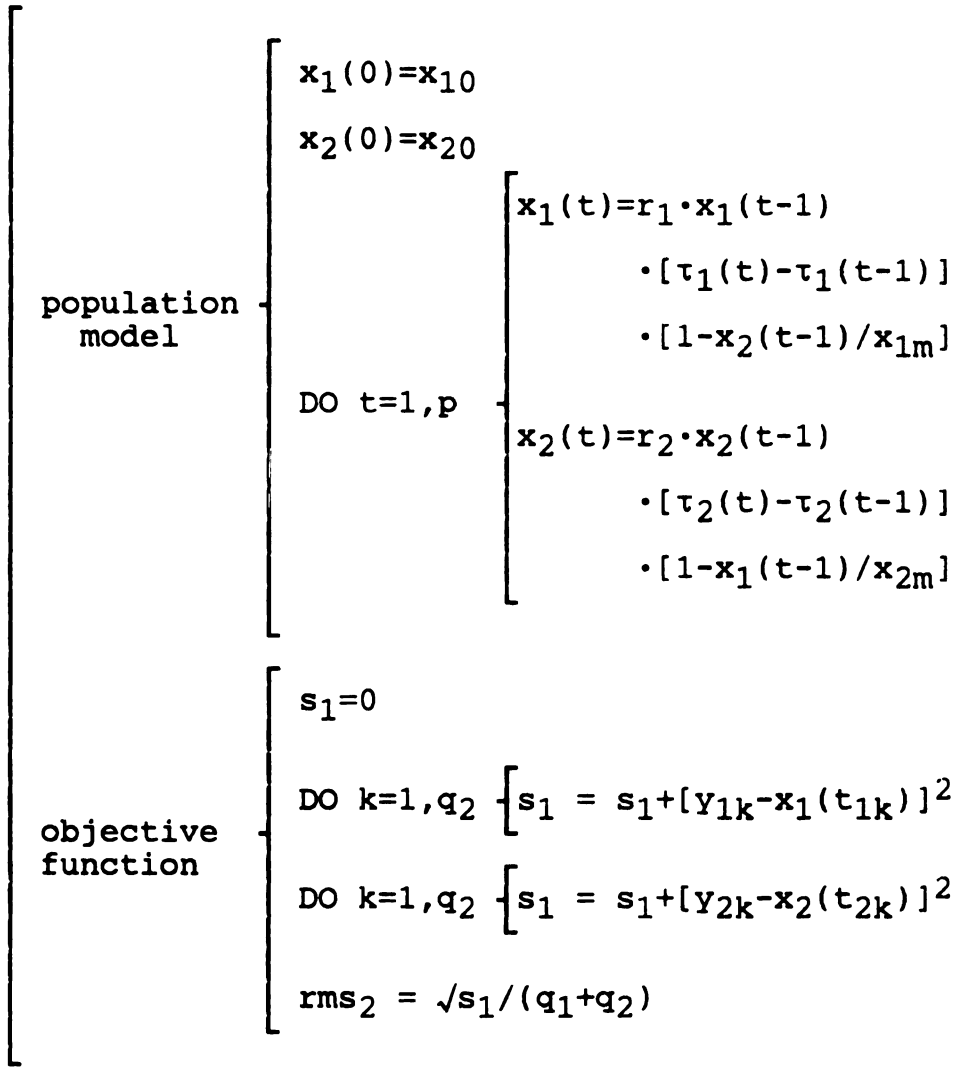
temperature profile: temp.dat
phenological history: phenol.dat



Algorithm 6.6 Computing physiological times for multi-species population Type I models.

$$
\begin{bmatrix}
\begin{bmatrix}
s_1 = 0 \\
s_2 = 0 \\
\text{DO } k=1,q_{11} \begin{bmatrix} s_1 = s_1 + \ln(y_k)\cdot\tau_1(t_{1k}) \\ s_2 = s_2 + \tau^2(t_{1k}) \end{bmatrix} \\
r_1 = s_1/s_2 \\
s_3 = 0 \\
\text{DO } k=1,q_{11} \begin{bmatrix} s_3 = s_3+[LN(y_{1k})-r_1\cdot\tau_1(t_{1k})]^2 \end{bmatrix} \\
\text{rms1} = \sqrt{s_3/q_{11}}
\end{bmatrix} \\
\begin{bmatrix}
s_1 = 0 \\
s_2 = 0 \\
\text{DO } k=1,q_{11} \begin{bmatrix} s_1 = s_1 + LN(y_k)\cdot\tau_1(t_{1k}) \\ s_2 = s_2 + \tau^2(t_{1k}) \end{bmatrix} \\
r_1 = s_1/s_2 \\
s_3 = 0 \\
\text{DO } k=1,q_{11} \begin{bmatrix} s_3 = s_3+[LN(y_{1k})-r_1\cdot\tau_1(t_{1k})]^2 \end{bmatrix} \\
\text{rms1} = \sqrt{s_3/q_{11}}
\end{bmatrix}
\end{bmatrix}
$$

Algorithm 6.7 Calculating the objective rms function for the multi-species population (exponential) model.



Algorithm 6.8 Calculating the objective rms function for any population (both endogenous and exogenous) model.

SUMMARY OF NUMERICAL TECHNIQUES

In any mathematical model it is necessary to determine values for the various constants. If it is possible to express this model in an explicit algebraic form, traditional methods such as least squares may be employed. If the form is linear, it is especially convenient, but even in nonlinear cases, the problem is tractable.

In the case where the model is not algebraically expressible, it is also possible to evaluate parameters using experimental data. This is done by minimizing an objective function comparing the model against the empirical data. For $n=1$ parameter, a line searching technique can be employed while if $n>1$, a multidimensional approach is necessary. These approaches are currently used in practice for a variety of situations.

Applying the heat-unit hypothesis to biological models, however, highlights two very unique circumstances. Applying these numerical methods, therefore, requires an approach specifically adapted to heat-unit models.

First, one of the modeling parameters - the temperature thresholds(s), implicitly defines the independent variable, physiological time. The optimization procedure outlined above was applied specifically to this problem and quite

satisfactory results were achieved. This approach assumed the existence of an initial "endogenous" growth phase characterized by negligible species-environmental interaction. This permitted the use of an initial exponential model.

Second, by assuming this endogenous period, it is necessary to establish the exogenous phase where both species-species and species-environmental interactions are significant. This presents a separate methodological problem to be considered. However, since the time base was previously determined, this mathematically assumes the form of the classical parameter evaluation methods.

Unfortunately, this has introduced a subjective element into the plan. It is now necessary to decree an (arbitrary) time horizon for the endogenous phase. Fortunately, analysis of differing horizons, while not entirely mechanical, does seem to confirm that a fairly distinct horizon does exist.

This chapter has developed the mathematics and algorithms necessary for any parameter evaluation using the heat-unit hypothesis. This has been done for the generalized threshold model types. Also, several examples were given in which parameters were determined and population results shown over time. These results seem to be quite reasonable.

CHAPTER SEVEN

STOCHASTIC WEATHER GENERATION

INTRODUCTION

Stochastic input is characteristic of a large class of system simulators. For biological systems this often takes the form of weather. In this research we have confined our attention to only temperature, but other extensions are certainly possible. Indeed, the motivation for such a generator might consider temperature as an archetype. Therefore, let us look at those properties of temperature that are natural for all stochastic generators.

i. The variations in the annual temperature distribution is time dependent in a regular sort of way. Certainly the seasonal changes are not only part of our experience, but are well documented. If we ignore climatic changes in the large, the local (say fifty years) experience is quite consistent.

ii. For a given time slice, that is we consider any specific day, the distribution is statistically regular. For example, we might consider the mean temperature to be normally distributed and the precipitation to be characterized by a Poisson distribution. Even if this distribution is not easily modeled by one of the classical ones, its behavior is sufficiently pronounced to be well understood.

iii. For any given time slice, the actual temperature is random. We might consider this to be simply statistically independent of other weather or of the previous day's weather. However, in a more precise view we should consider the temperature to be quite dependent on these other entities as well.

With these properties in mind, we create what will be called a **stochastic driver**. The particular stochastic generator we use in this research is called a **stochastic weather generator**, the only difference between them being that the weather generator will generate several random sequences and the stochastic driver, strictly speaking, generates but one.

THE STOCHASTIC DRIVER

The function of a stochastic driver is to generate a sequence of real numbers (temperatures) that will serve as the exogenous input to the system model. This sequence, called the [**generated**] **profile**, needs to be random. That is to say, the sequence generated will be different each time the generator is invoked. In order to be truly useful a stochastic generator must also retain certain statistical properties of an idealized profile distribution called the **profile base.**

We write this using the following notation:

Profile base: X(t)

[7.1]

Profile: x(t)

for t=1,2,...,p

It is important to understand that the profile base is, strictly speaking, a random process. That is, for a given t, X(t) represents a random variable with the usual descriptive statistics. This is to be contrasted with the profile x(t) which is not a random variable, but simply a specific sequence.

On the other hand, we can speak of a set of q historical profiles defined for p time slices:

[7.2] $x_k(t)$ for $t=1,2,\ldots,p$ and $k=1,2,\ldots,q$

For a large number of instances ($q\to\infty$), the historical database should closely resemble the profile base. Thus, in order for a generator to be useful, we would likely require that the mean and variance of the profile base are closely approximated by the generated profile sequences...

$$E[X(t)] \approx \frac{1}{q} \cdot \sum_{k=1}^{q} x_k(t)$$

[7.3]

$$Var[X(t)] \approx \frac{1}{q} \cdot \sum_{k=1}^{q} (x(t)-E[X(t)])^2$$

where $E[\cdot]$ is the expected value,
$Var[\cdot]$ is the variance.

In physical systems it is often possible to hypothesize an aesthetically reasonable distribution for the profile base. In biological systems, however, a theoretical base is usually unrealistic. Even so, for temperature, the history profile is quite consistent and excellent results can be had

by simply using this empirical history distribution as the actual base!

In summary, a stochastic generator simply produces a sequence of numbers whose statistical properties mimics to some predetermined degree the statistics of the profile base. In this chapter we will do the following:

i. Define those stochastic drivers that will be useful in generating a temperature sequence.

ii. Define a hierarchy of generators whose statistical properties will prove to be increasingly precise in a given statistical sense.

iii. Demonstrate these statistical properties by means of examples.

iv. Justify the use of stochastic processes in the computation of physiological time and dynamic systems.

v. Illustrate the applicability of this procedure to population dynamical systems.

We consider an empirical history as a table of **mean temperatures** $M_k(t)$ and **temperature ranges** $R_k(t)$ as in Tables 7.1 and 7.2. It is necessary for us to use both mean temperature and temperature range because of the sampling mechanism outlined in chapter five. This will allow the daily temperature fluxuation to be ignored, but still be accounted for it in computing the number of heat units (degree-days) per day.

We define sample means and variances for each of the these entities as follows:

$$[7.4a] \qquad \mu_M(t) \sim \frac{1}{q} \cdot \sum_{k=1}^{q} M_k(t)$$

$$[7.4b] \qquad \sigma_M(t) \sim \frac{1}{q-1} \cdot \sum_{k=1}^{q} [M_k(t)-\mu_M]^2$$

$$[7.5a] \qquad \mu_R(t) \sim \frac{1}{q} \cdot \sum_{k=1}^{q} R_k(t)$$

$$[7.5b] \qquad \sigma_R(t) \sim \frac{1}{q-1} \cdot \sum_{k=1}^{q} [R_k(t)-\mu_R]^2$$

Certainly for large q, each of these expressions should asymptotically approach their theoretical counterparts on the left. Therefore, we assume ergodic behavior.

Table 7.1: Daily Mean Temperature Empirical History

	day 1	day 2	...	day p
year 1	$M_1(1)$	$M_1(2)$	...	$M_1(p)$
year 2	$M_2(1)$	$M_2(2)$	...	$M_2(p)$
	.	.		.
	.	.		.
	.	.		.
year q	$M_q(1)$	$M_q(2)$	...	$M_q(p)$
mean	$\mu_M(1)$	$\mu_M(2)$	...	$\mu_M(p)$
standard deviate	$\sigma_M(1)$	$\sigma_M(2)$	...	$\sigma_M(p)$

Table 7.2: Daily Temperature Range Empirical History

	day 1	day 2	...	day p
year 1	$R_1(1)$	$R_1(2)$	...	$R_1(p)$
year 2	$R_2(1)$	$R_2(2)$	...	$R_2(p)$
	.	.		.
	.	.		.
	.	.		.
year q	$R_q(1)$	$R_q(2)$	...	$R_q(p)$
mean	$\mu_R(1)$	$\mu_R(1)$	...	$\mu_R(p)$
standard deviate	$\sigma_R(1)$	$\sigma_R(1)$	...	$\sigma_R(p)$

THE DETERMINISTIC TEMPERATURE GENERATOR

Although not useful in this specific project, a most popular temperature generator is the very simple one of a deterministic model. We mention it here because a great majority of models use this! It simply computes the mean for each time slice, then considers this as the "typical" mean temperature and temperature range for the day. Diagrammatically we might consider this as in Figure 7.3.

Such models for producing somewhat characteristic temperatures are indeed used. However, this simple procedure fails to account for the obvious temperature fluxuation and since it is deterministic, no stochastic results are attainable. We, therefore, consider this generator an anachronism and move on to other models.

THEORETICAL DISTRIBUTION TEMPERATURE GENERATORS

One of the properties of temperature distributions generally is that over a local time frame we can assume their distributions to be well defined. That is for a number of repeated years, the descriptive statistics and distributions themselves are effectively constant. This behavior is technically termed **stationarity** and is assumed throughout the rest of this work.

Not surprisingly, the distribution of the mean temperature $M_k(t)$ for fixed t (and varying k) is roughly **Gaussian** distributed. Also, the temperature range $R_k(t)$ can be approximated by a **Gamma** distribution. That this is so is illustrated in Figure 7.2 where the distributions for the given time slice June 1 are shown for q=20 years.

Because of this stationarity property, we can generate the required random sequences by introducing a random variate U[0,1] which is **uniformly** distributed on the interval [0,1]. This particular random number generator is available for almost all programming languages.

For most theoretical distributions there exist algorithms by which we can generate sequences which mimic these distributions by using U[0,1] as a source. For instance, we can generate normally distributed random variables z :

N[0,1] which have mean zero and standard deviation 1 using Algorithm 7.1.

This can be extended to produce a variate x which is normally distributed with mean μ and standard deviation σ:

$$x = \sigma \cdot z + \mu$$

The Gamma distribution is also characterized by two parameters. α is continuous and k is discrete: k=1,2,.... If k=1, the distribution is called exponential and if $k \to \infty$, it reduces to Gaussian (with the mean proportional to the variance: $\mu = \alpha \cdot \sigma^2$)

It is easily shown that the mean and variance are given by...

[7.6]

$$\text{Mean:} \quad \mu = k/\alpha$$

$$\text{Variance:} \quad \sigma^2 = k/\alpha^2$$

Therefore:

[7.7]

$$k = (\mu/\sigma)^2$$

$$\alpha = \mu/\sigma^2$$

It is entirely likely that k will not be integral by equation [7.7]. However, this is not a large problem in

practice since k is fairly large (k ~ 25+) and the mean generated will, therefore, be close to that of the true mean. The algorithm for generating a gamma distribution is shown as Algorithm 7.2 where x is the random variate.

These two procedures, generating a normal distribution $N[\mu,\sigma]$ and generating a gamma distribution $\Gamma[\mu,\sigma]$ are sufficient to create a stochastic weather generator in which very agreeable statistical properties are obtained. First, the mean and standard deviations of the basis profile (as defined by historical data) are retained. Secondly, the (theoretical) distributions desired are also mimicked. Figure 7.4 is a diagram of this procedure.

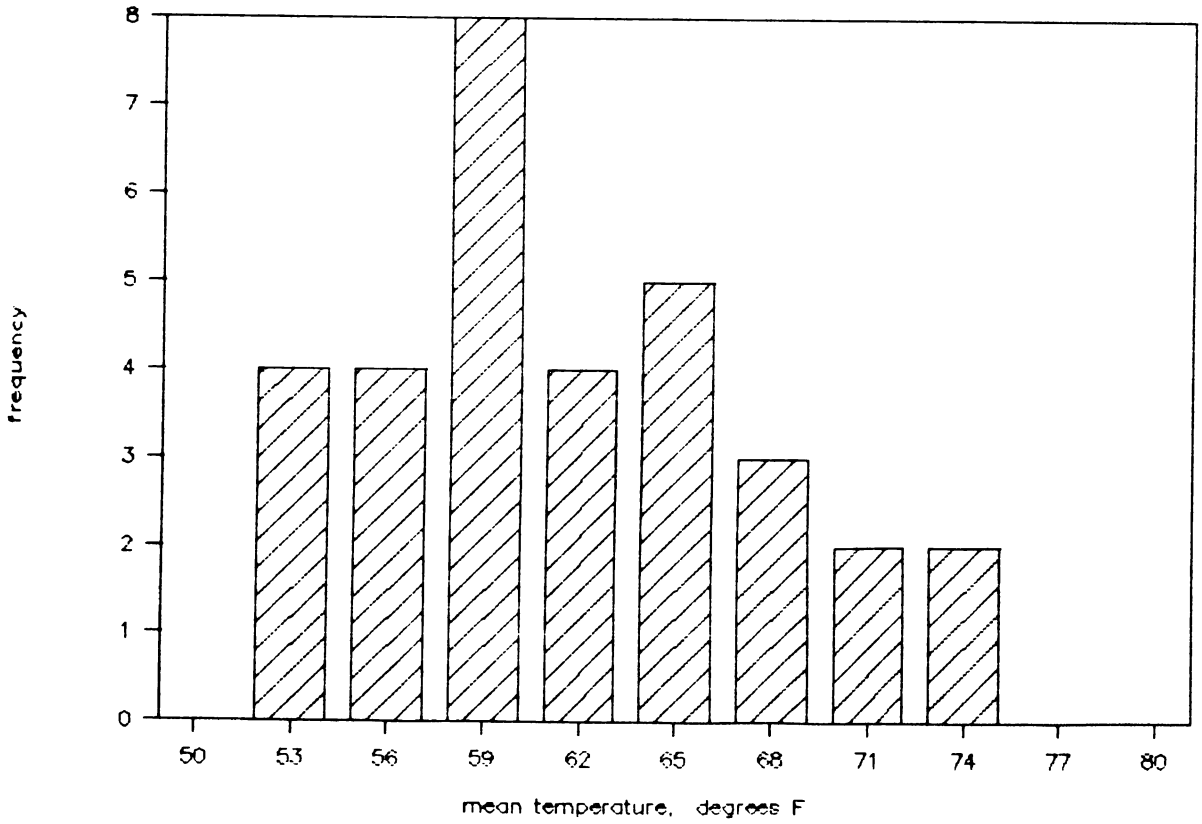


Figure 7.1: Mean Temperature Distribution for the June 1 Time Slice.

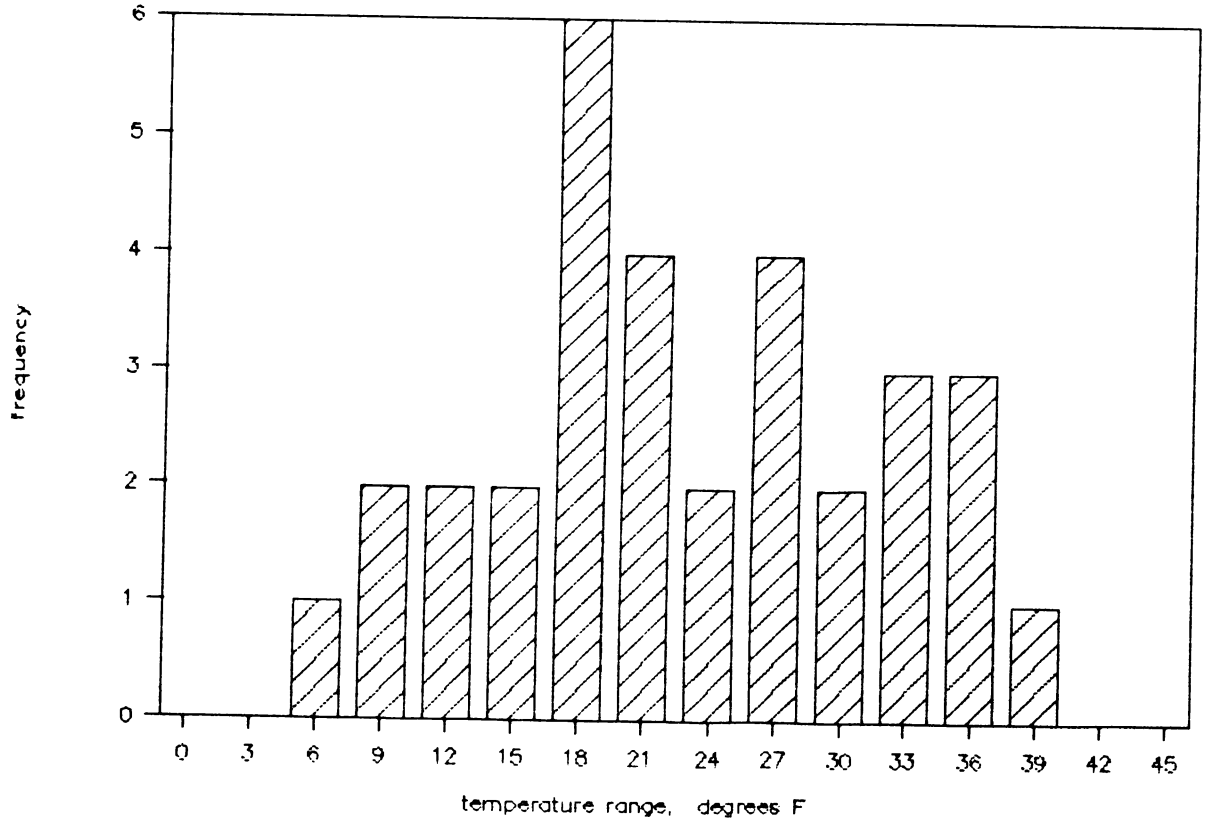


Figure 7.2: Temperature Range Distribution for the June 1 Time Slice.

$$
\begin{cases}
r_1 = \text{RND} \\
r_2 = \text{RND} \\
z = \sqrt{(-2 \cdot \ln(r_1))} \cdot \cos(2\pi \cdot r_2) \\
x = \sigma \cdot z + \mu
\end{cases}
$$

Algorithm 7.1: Generating Gaussian normalized variates z and rectified variates x.

$$z : N[0,1]$$
$$x : N[\mu,\sigma]$$

$$
\begin{cases}
k = \text{int}((\mu/\sigma)^2 + 0.5) \\
s = 1 \\
\text{DO } i=1,k \quad \begin{cases} r = \text{RND} \\ s = s \cdot r \end{cases} \\
x = -\sigma^2 \cdot \ln(s)/\mu
\end{cases}
$$

Algorithm 7.2: Generating a gamma variate x.

$$x : \Gamma[\mu,\sigma]$$

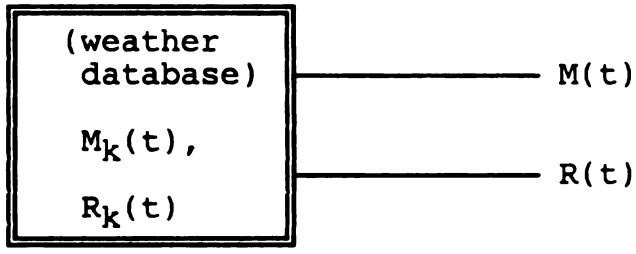


Figure 7.3: The deterministic generator.

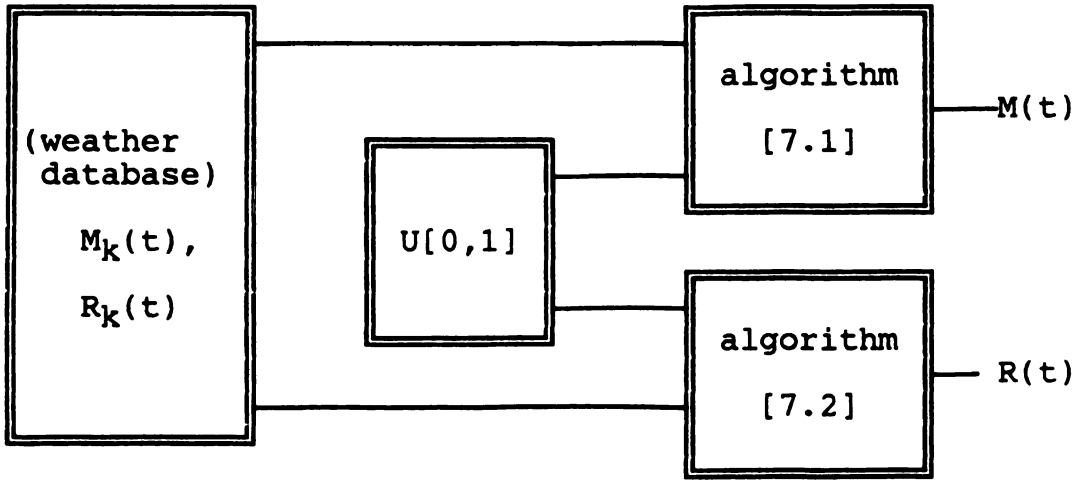


Figure 7.4: The theoretic distribution fitting generator.

EMPIRICAL DISTRIBUTION TEMPERATURE GENERATORS

The assumption that $T_k(t)$ is normally distributed is mathematically attractive, but for a variety of reasons not entirely realistic. Although based on only 22 years historical data, Figures 7.1 and 7.2 show that it may be preferable to model the stochastic process on empirical results rather than the ideal Gaussian and Gamma random variates.

Fortunately, it is possible to reproduce not only the mean and standard deviation from a theoretical distribution, but also the empirical distribution itself! This is accomplished as follows:

> Suppose the density and distribution functions of a given distribution are represented by f(x) and F(x) respectively. Since F(x) is defined over the range [0,1], we can set F(x) = u for a uniformly distributed random variable u : U[0,1]. It is clear that x is uniquely determined by this equation since F is a non-decreasing function. u determines the random variable.

$$x = F^{-1}(u)$$

Then $\text{Prob}\{ X < x \} = F(x)$

$$= \text{Prob}\{ u < F(x) \}$$

$$= \text{Prob}\{ F^{-1}(u) < x \}$$

> Thus, $F^{-1}(u)$ is a variate that has f(x) as its density function.

This can be done for both the mean temperature and temperature range and can be implemented using Algorithm 7.3. This algorithm will produce a temperature distribution

$$S'' = \{x_1'(t),\ldots,x_m'(t)\}$$

identical with S over the long run (i.e. as $m\to\infty$). Also,

$$E[S] = E[S'']$$

[7.8]

$$Var[S] = Var[S'']$$

Thus, not only are the mean and standard deviation of the distribution preserved, but so is the distribution itself! This is an important property since while the mean temperature might be normally distributed, other important quantities most certainly are not!

The essence of this distribution-preserving procedure is summarized by the block diagram in Figure 7.5.

Consider the set of data points

$$S = \{x_1(t),\ldots,x_n(t)\}$$

Sort S into ascending order. Denote this ordered set by S'.

$$S' = \{x_1'(t),\ldots,x_n'(t)\}$$

Let U[0,1] be a random variate uniformly distributed over [0,1].

Then choose a sequence of m trials by...

if $0 < n \cdot u < 1$, choose $x_1'(t)$

if $1 < n \cdot u < 2$, choose $x_2'(t)$

.
.
.

if $n-1 < n \cdot u < n$, choose $x_n'(t)$

Algorithm 7.3: The Empirical Distribution Fitting Algorithm.

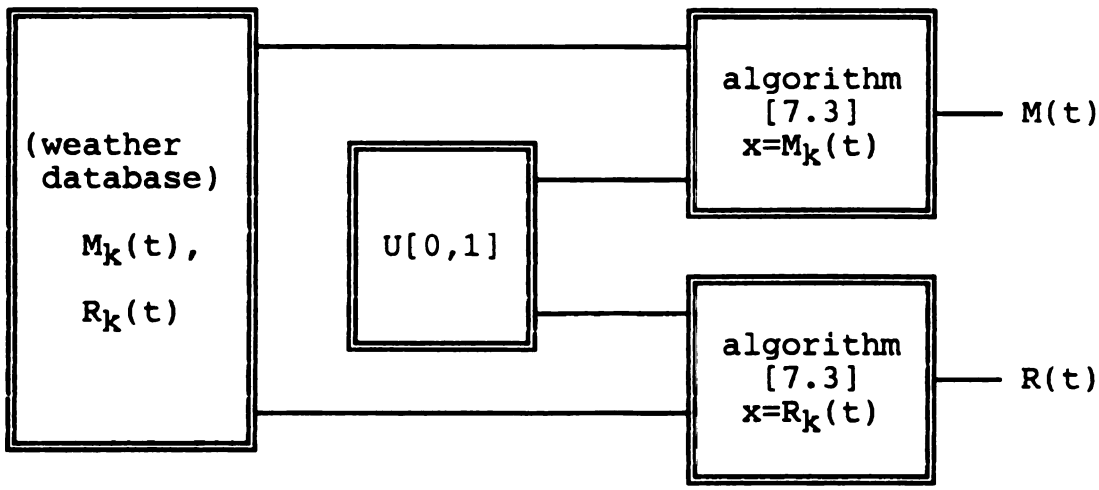


Figure 7.5: The empirical distribution preserving generator.

EXAMPLE ONE

We continue by illustrating with an example. Choosing several typical dates as time slices, each of the above algorithms were implemented to show the comparative statistics for the deterministic, theoretic distribution fitting and empirical distribution fitting results. Table 7.3 summarizes the means and standard deviations for 100 trials using each technique.

It is easily verified that in each case shown there is no statistically significant difference between any of these measures. Therefore, in generating stochastic data, both the theoretical and empirical distribution fitting techniques are quite satisfactory. On the other hand, if it is necessary that the distribution itself also be maintained there is yet a considerable difference.

Consider the mean temperature data as shown in Figure 7.6. While Figure 7.6a might give sufficient intuition to justify the use of a Gaussian distribution (as shown in Figure 7.6b), the empirical distribution fitting as shown in Figure 7.6c is apparently a realistic representation of the original distribution. On the other hand, the 10 years of data used as the basis of the database is hardly sufficient to guarantee that nature is actually similar to Figure 7.6a!

A similar argument holds for temperature ranges. Figure 7.7b shows the theoretic Gamma distribution results and 7.7c gives the empirical distribution fit. Again it is seen that 7.7c faithfully reproduces the 10 years history used as a basis for the example in Figure 7.7a.

Table 7.3: Mean/Standard deviation comparisons for each temperature measure using various generators. (Using 100 trials.)

(April 1)	Mean Temperature	Temperature Range
Deterministic	41.81/12.01	11.40/7.76
Theoretic Distribution	41.65/12.13	11.21/7.81
Empirical Distribution	41.69/11.94	11.34/7.75
Feedback	41.83/11.99	11.40/7.71

(May 1)	Mean Temperature	Temperature Range
Deterministic	54.87/6.57	22.41/7.10
Theoretic Distribution	54.93/6.60	22.62/7.13
Empirical Distribution	54.97/6.48	22.28/7.08
Feedback	54.88/6.54	22.43/7.09

(June 1)	Mean Temperature	Temperature Range
Deterministic	60.81/6.88	26.24/6.65
Theoretic Distribution	60.63/6.92	26.49/6.48
Empirical Distribution	61.59/6.69	26.69/6.51
Feedback	60.86/6.85	26.23/6.54

FEEDBACK GENERATORS

There is one further useful refinement to the temperature generator yet to be considered. In addition to being distribution sensitive, the temperature on a given day is often highly correlated with the temperature of the previous day. This **autocorrelation** effect can be accounted for by using a feedback mechanism. Figure 7.8, where the generated temperature T(t) is produced by feedback on the basis predictor P(t), illustrates this diagrammatically.

In words, this feedback generator can be interpreted as follows:

> The future temperature is directly related to the present temperature (proportionality constant a) and the database predictor (proportionality constant b).

Thus,

[7.9] $$T(t+1) = a \cdot T(t) + b \cdot P(t)$$

where P(t) is the predictor of the basis.

Useful predictors, such as the ones discussed above have statistical properties identical to the temperature. In particular, their expected values should be the same:

[7.10] $E[T(t)] = E[P(t)] = \mu$

Therefore,

$$E[T(t+1)] = E[a \cdot T(t) + b \cdot P(t)]$$

$$= a \cdot E[T(t)] + b \cdot E[P(t)]$$

[7.11] $\mu = a \cdot \mu + b \cdot \mu$

implies that $a + b = 1$.

Thus:

[7.12] $T(t+1) = a\ T(t) + (1-a)\ P(t)$

where P(t) is calculated as in previous examples. For instance, assuming a normal temperature distribution, this reduces to

[7.13] $T(t+1) = a \cdot T(t) + (1-a) \cdot [n \cdot R(t)/\sqrt{8} + M(t)]$

where R(t) is the temperature range,
M(t) is the mean temperature and
z is a normal random variate:

$$z : N[0,1].$$

Consider the extremes for a. If a=0, the future temperature is solely determined by the predictor. This is the situation considered in the deterministic and distribution

preserving generators above. On the other hand, if a=1, the future temperature is simply the same as the present temperature. This is too autocorrelated! It follows that there exists a "best" value for a somewhere between 0 and 1.

In order to find this optimal a we consider a sample year at the location in question. By comparing the calculated temperature profile (using equation [7.13]) against the actual, it is possible to choose a so that the following Root Mean Square (rms) is minimized.

$$[7.14] \qquad rms^2 = \frac{1}{p} \cdot \sum_{t=1}^{p} [T(t)-M(t)]^2$$

$$[7.15] \qquad \approx \frac{1}{p} \cdot \sum_{t=1}^{p} [a\cdot(T(t)-M(t)) + (1-a)\cdot z\cdot R(t)/\sqrt{8}]^2$$

assuming a standardized normal distribution fitting,

$$z : N[0,1]$$

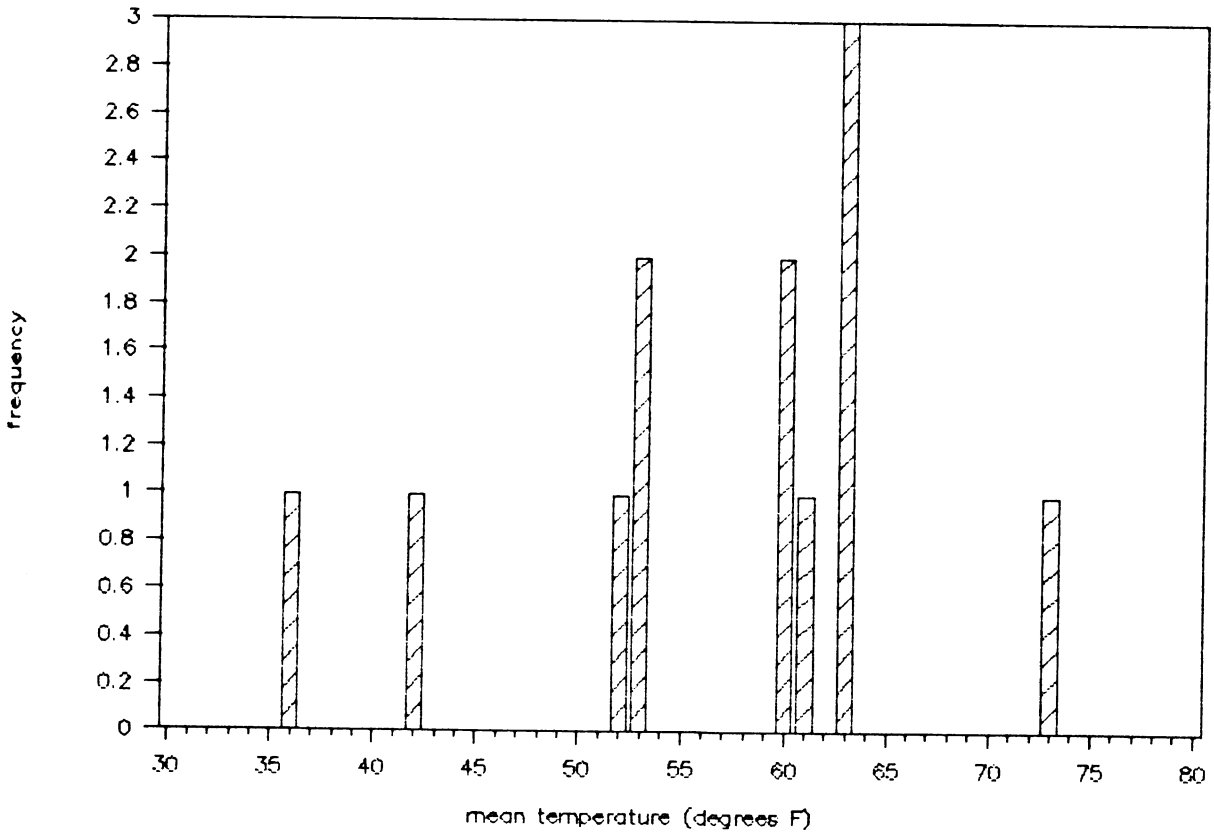


Figure 7.6a: **Mean temperature distribution for the June 1 time slice using the deterministic generator.**

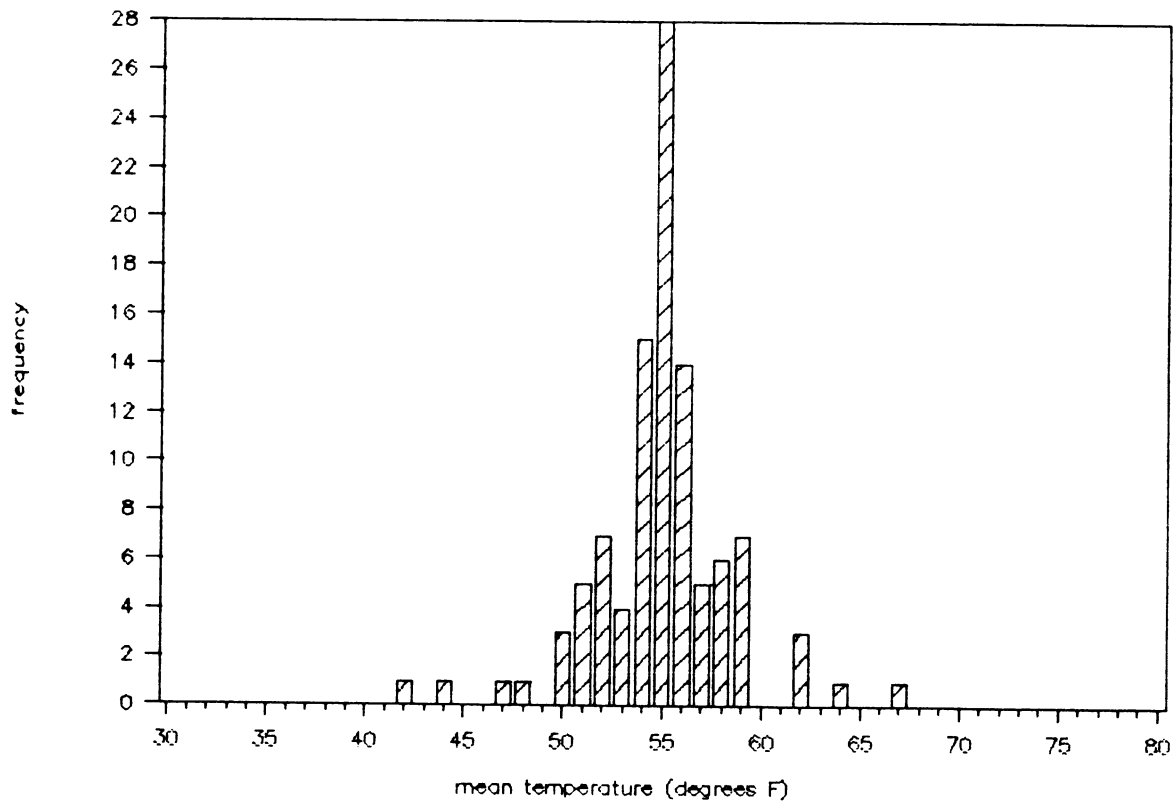


Figure 7.6b: Mean temperature distribution for the June 1 time slice using the theoretic distribution fitting generator.

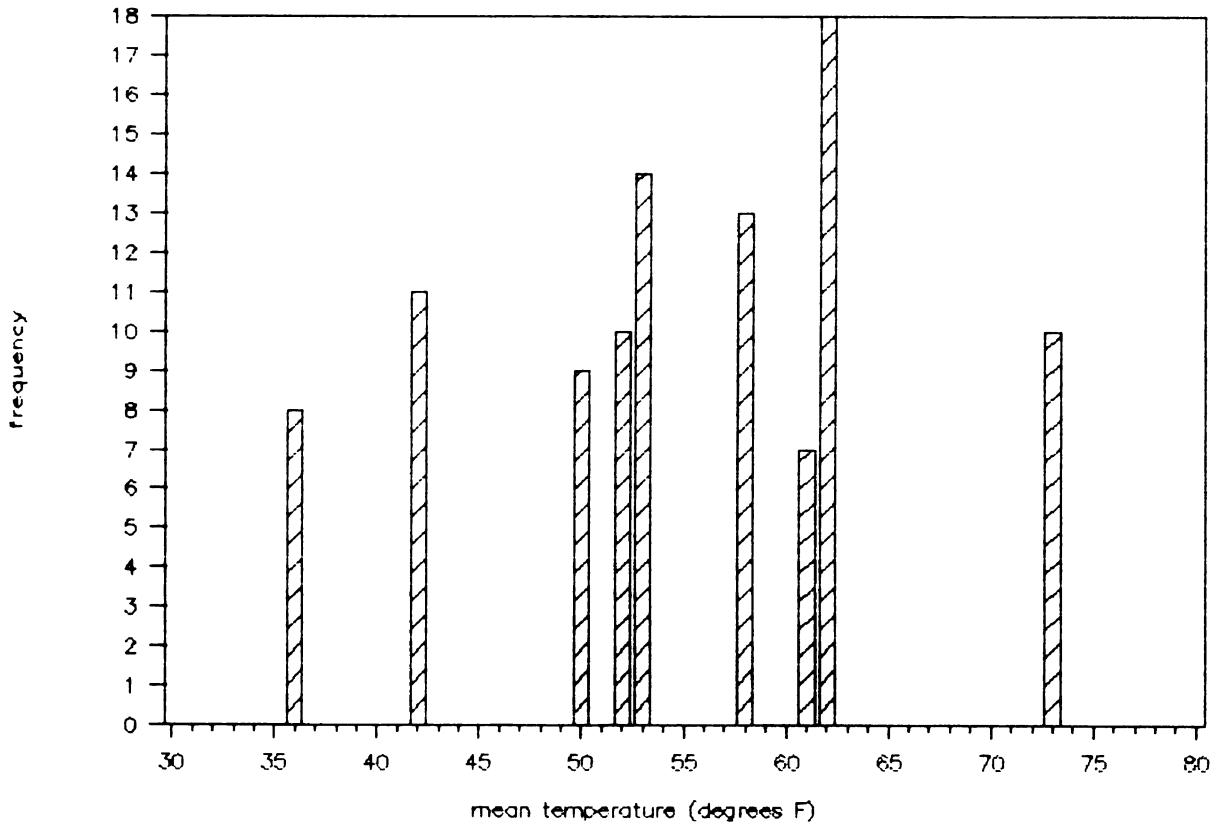


Figure 7.6c: **Mean temperature distribution for the June 1 time slice using the empirical distribution fitting generator.**

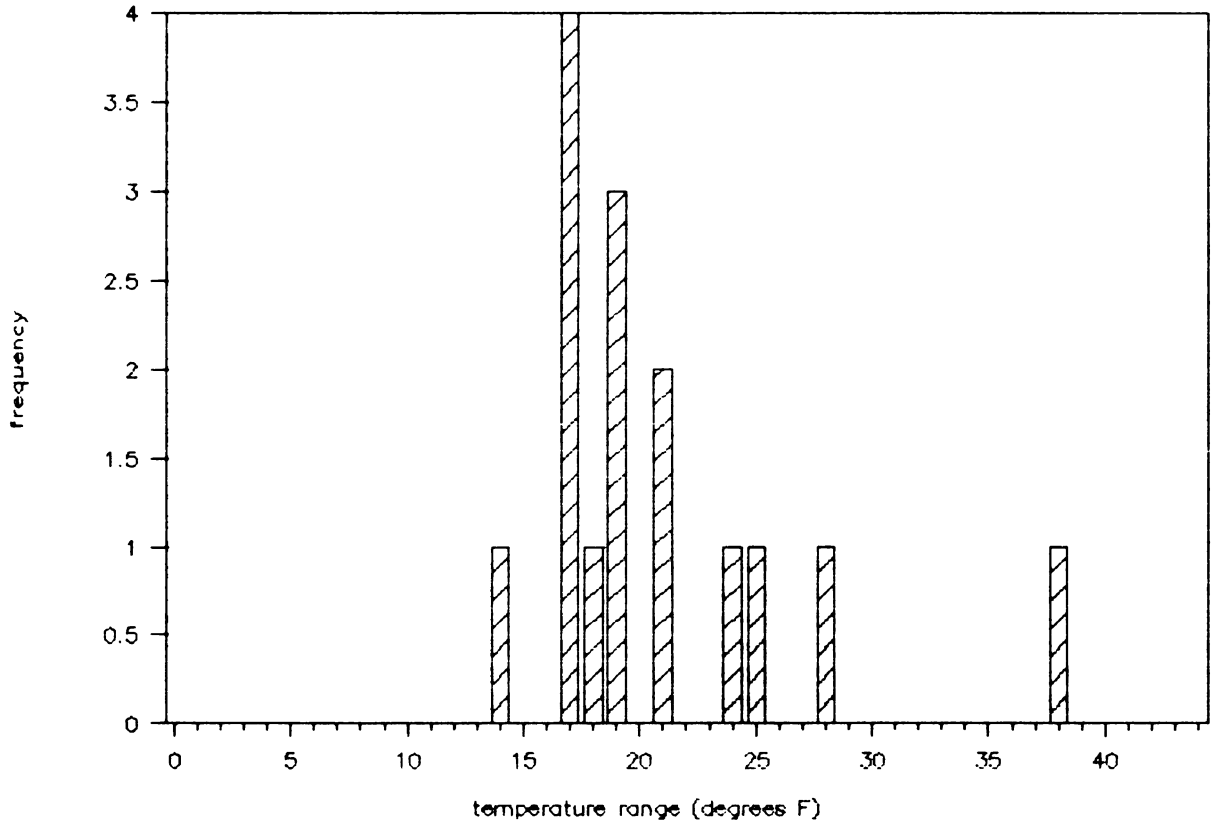


Figure 7.7a: Temperature range distribution for the June 1 time slice using the deterministic generator.

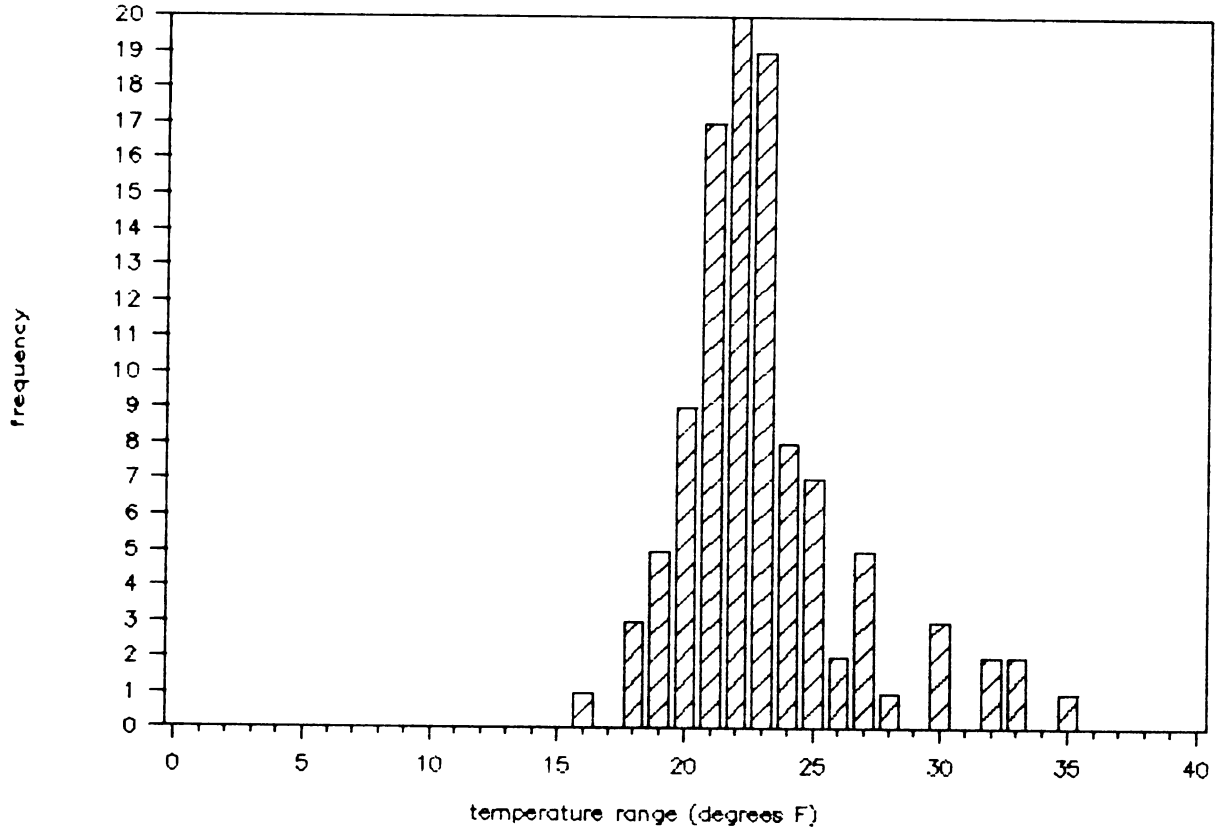


Figure 7.7b: Temperature range distribution for the June 1 time slice using the theoretic distribution fitting generator.

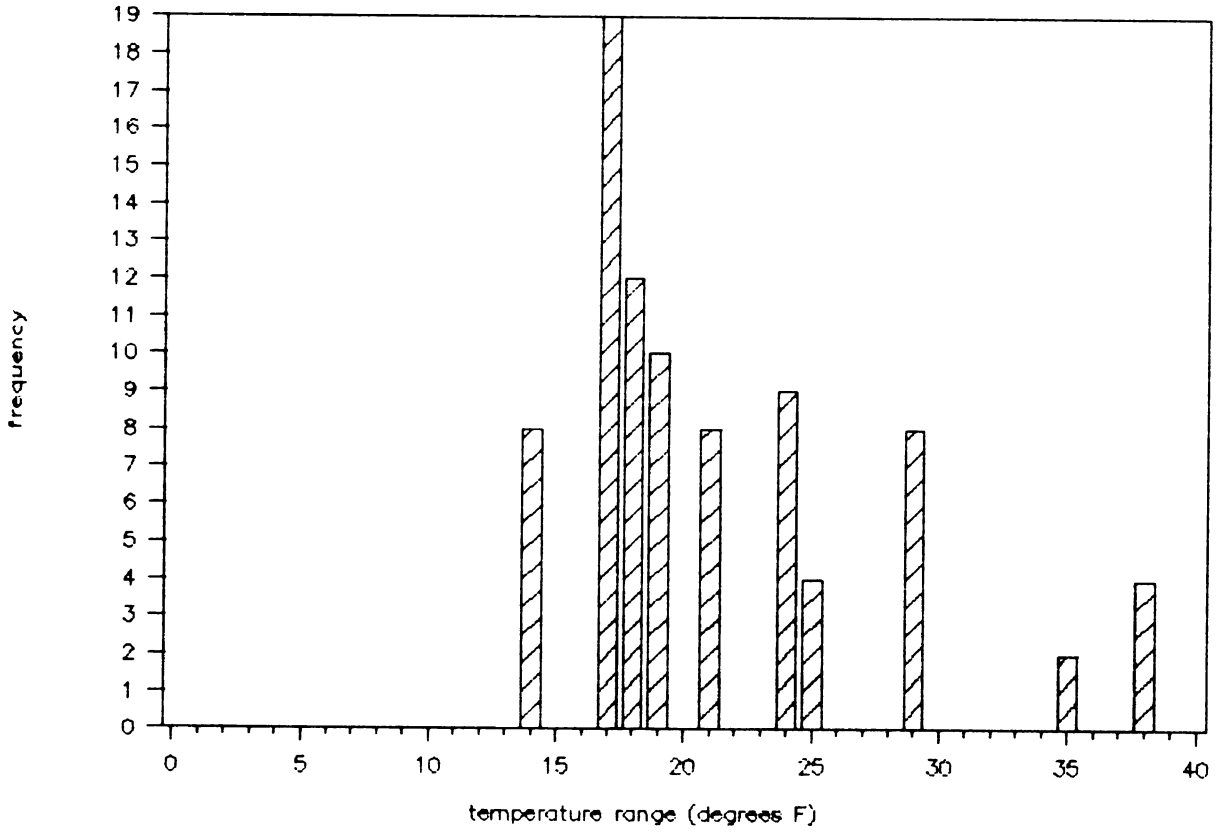


Figure 7.7c: Temperature range distribution for the June 1 time slice using the empirical distribution fitting generator.

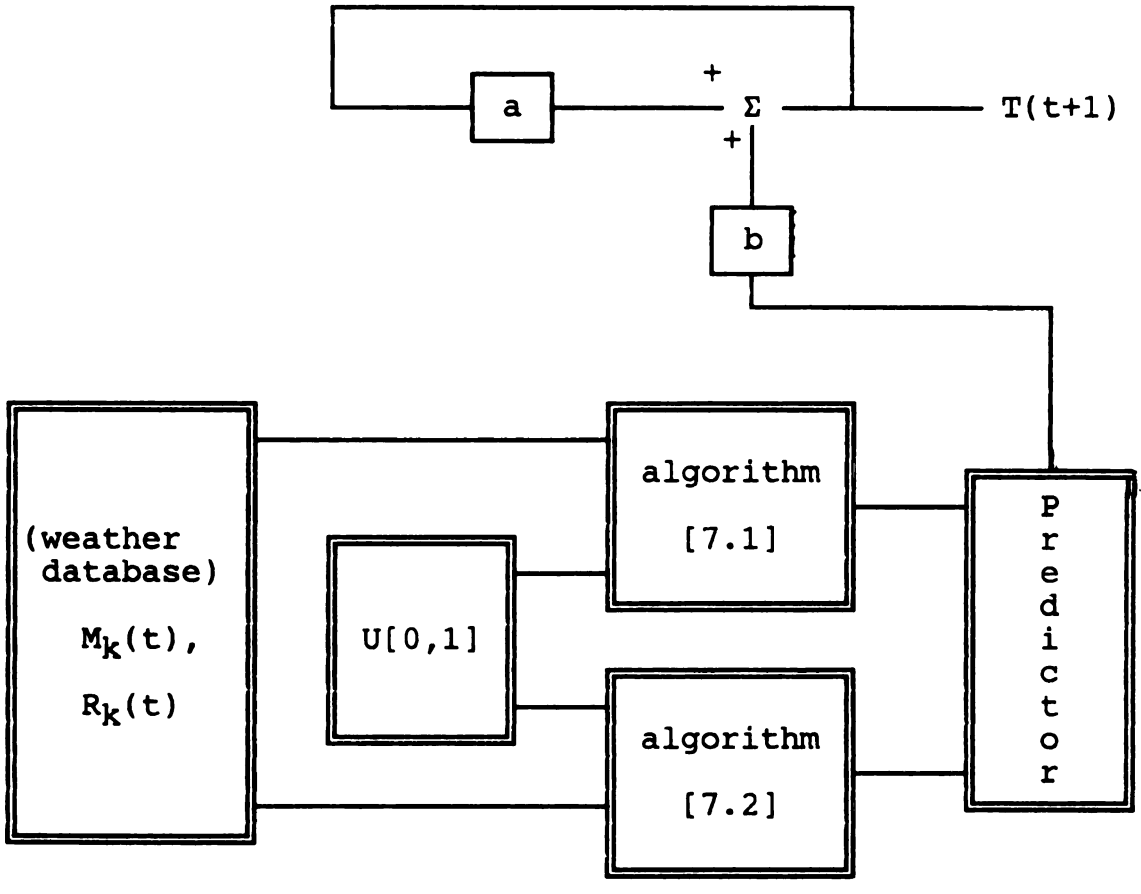


Figure 7.8: The Feedback Temperature Generator

EXAMPLE TWO

We again consider the temperature data for Lansing, Michigan. Using 1985 as the sample year, equation [7.13] was used to compute generated temperature T(t) and this was compared against the actual mean temperatures, M(t). For a values of 0 to 1 by steps of 0.01, the resulting graph plotting rms against a is shown in Figure 7.9.

At first glance it would appear that a relatively distinct minimum exists. However, noting the scale, not only is the minimum soft, but the curve itself rather flat. Even so, we can calculate a minimum by using a least squares quadratic approximation. An eyeball guess is probably just as sufficient and a = 2/3 is reasonable for this example.

Using the value a = 2/3, the means and standard deviations of 100 generator trials are tabulated in Table 7.3. It should be noted once again that these are not significantly different than the statistics calculated by other means. This distribution, however, has a more realistic autocorrelation property. Even so, the apparent softness of the minimum in Figure 7.9 gives reason not to stress this approach for this particular example.

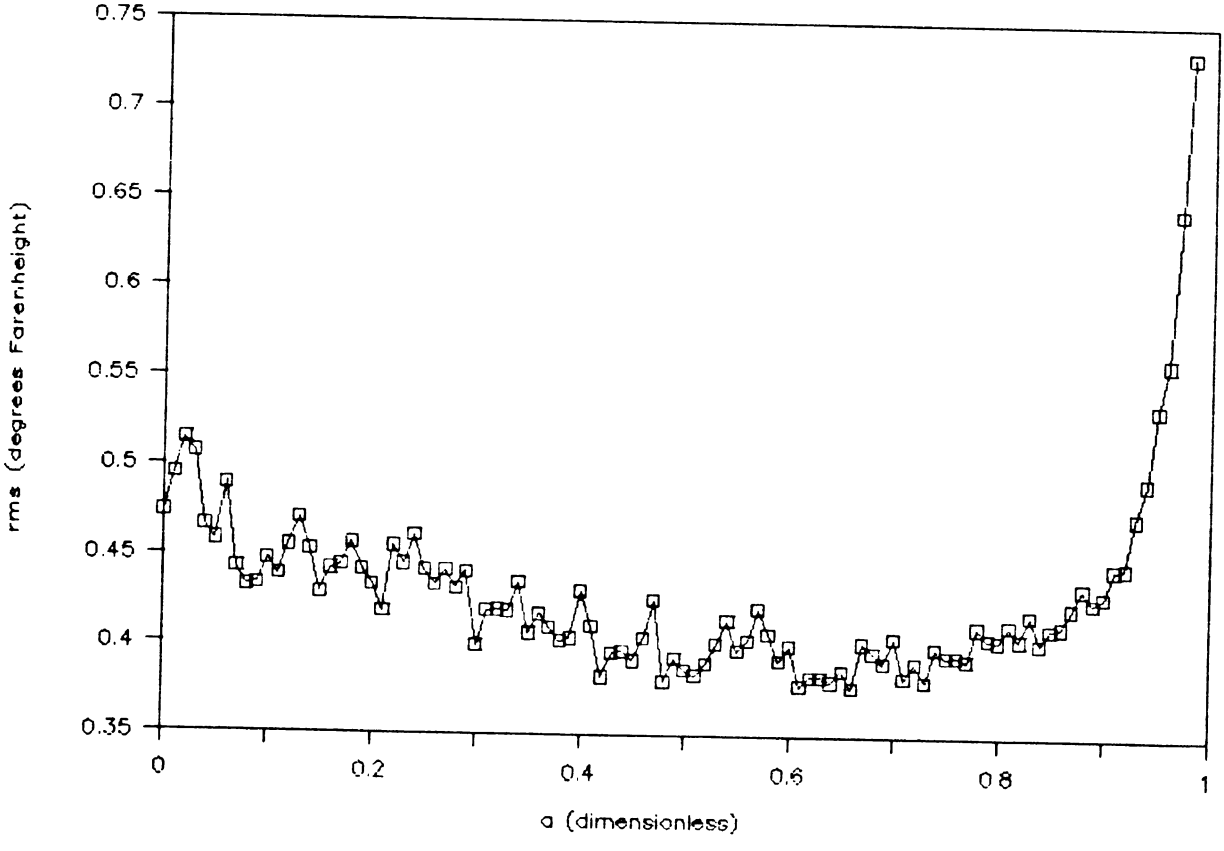


Figure 7.9: Plot of objective rms against parameter a for a feedback temperature generator. Lansing, Michigan (1985)

STATISTICAL PROPERTIES OF PHYSIOLOGICAL TIME

By the procedures outlined above, it is possible to generate a temperature profile $T(t)$ that will simulate any profile base by matching descriptive statistics, distributions, distribution parameters and time adjacent correlations. What is now important is to describe the statistical properties of the physiological time $\tau(t)$ and ultimately the population function $x(t)$. While the mean times and populations are useful, the variance of these quantities is the most important.

There are several significant correlations. The **means** and **variances** might depend on the various modelling parameters such as initial time t_0, temperature threshold T_0 and the final time horizon $t=t_f$. The results of several simulations illustrating this are summarized in Figure 7.10 where the mean physiological temperature is plotted against chronological time t for various initial times t_0 and thresholds T_0. Similarly, the standard deviations are plotted against time t in Figure 7.11.

It is readily seen that the means are close to linear in the center of the graph yet the distinctive sigmoid shape overall is quite apparent. Regardless of the initial time or threshold, all curves are roughly parallel.

On the other hand, Figures 7.8 illustrate that the standard deviations of each physiological time curve are very nearly identical. This is an important property. The standard deviation might vary with time t, but this is primarily a reflection of the temperature profile T(t). However, the variance of each curve is apparently **independent of threshold** for $T_0=40^o$, 45^o and 50^o.

Unfortunately this result is too good to be true for all thresholds T_0. The results of another series of simulations using different initial times t_0 in which the physiological time $\tau(t)$ is plotted against thresholds T_0 for $t_f=235$ days is given in Figures 7.10. We note that the mean number of heat units steadily decreases as would be expected. This is certainly a nice, smooth relationship, but not all that exciting! On the other hand, the variances as illustrated in Figures 7.11 are roughly constant over the threshold interval $35^o \leq T_0 \leq 56^o$. Coincidentally, realistic thresholds will be within this range. If, on the other hand, T_0 is greater than 56^o, the variance will vary significantly with threshold.

The significance of this result is as follows: Any simulation run in the $35^o \leq T_0 \leq 56^o$ range will demonstrate similar behavior in variance behavior. However, a higher threshold will lead to a much smaller statistical variance.

These thresholds will produce models which are less random and more deterministic.

Figures 7.10 and 7.11 illustrate the physiological time dependence on initial time t_0. Again, especially for reasonable starting times, the means will decrease with t_0 as expected, but the variances stay for the most part, constant.

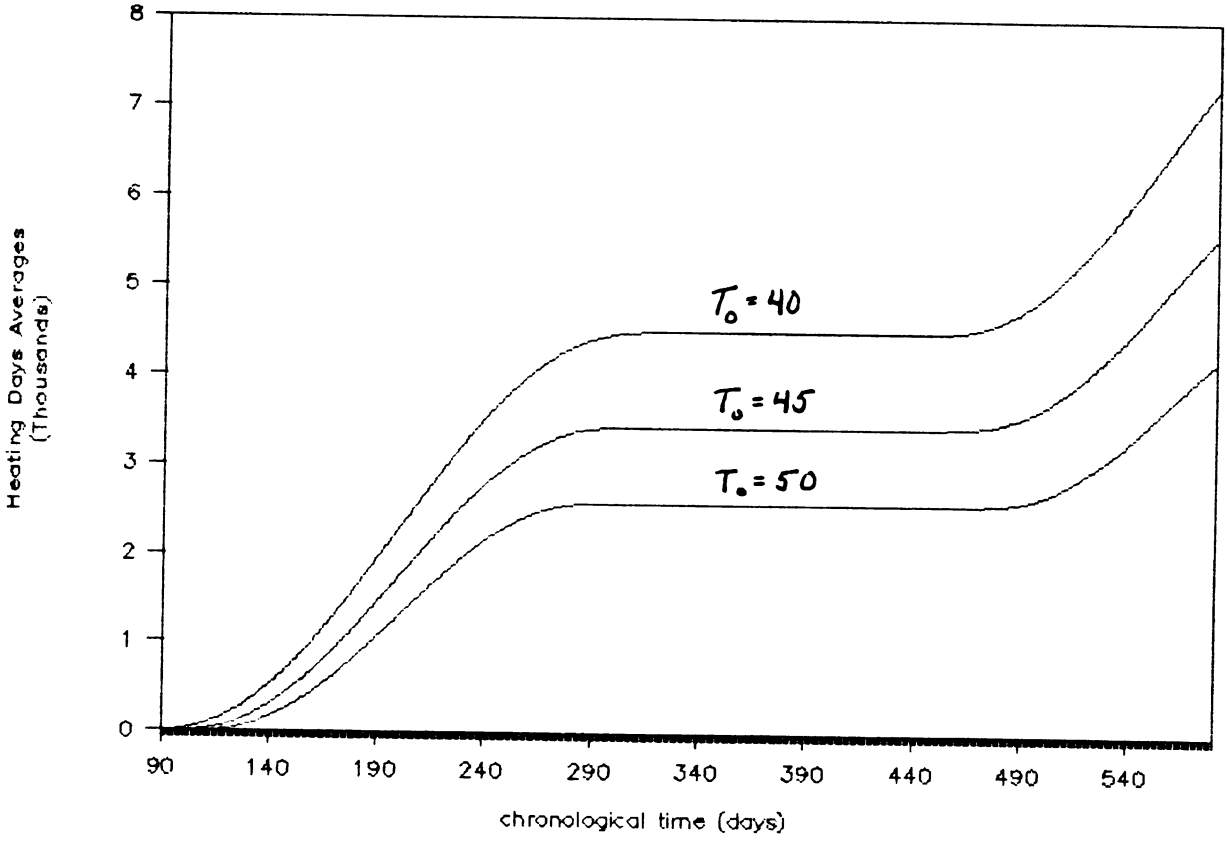


Figure 7.10a: Mean physiological time plotted against chronological time for various temperature thresholds.

t_0 = 90 days
T_0 = 40°, 45°, 50° F

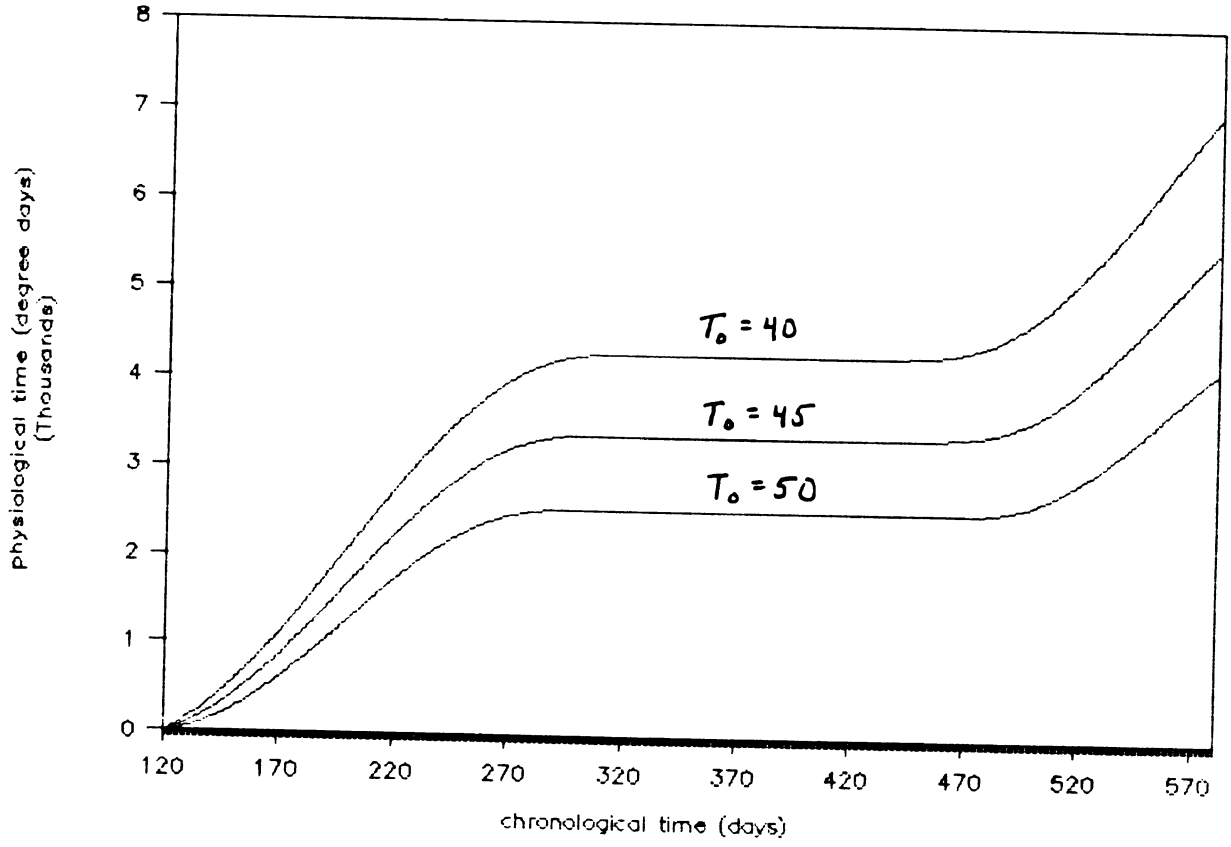


Figure 7.10b: Mean physiological time plotted against chronological time for various temperature thresholds.

t_0 = 120 days
T_0 = 40°, 45°, 50° F

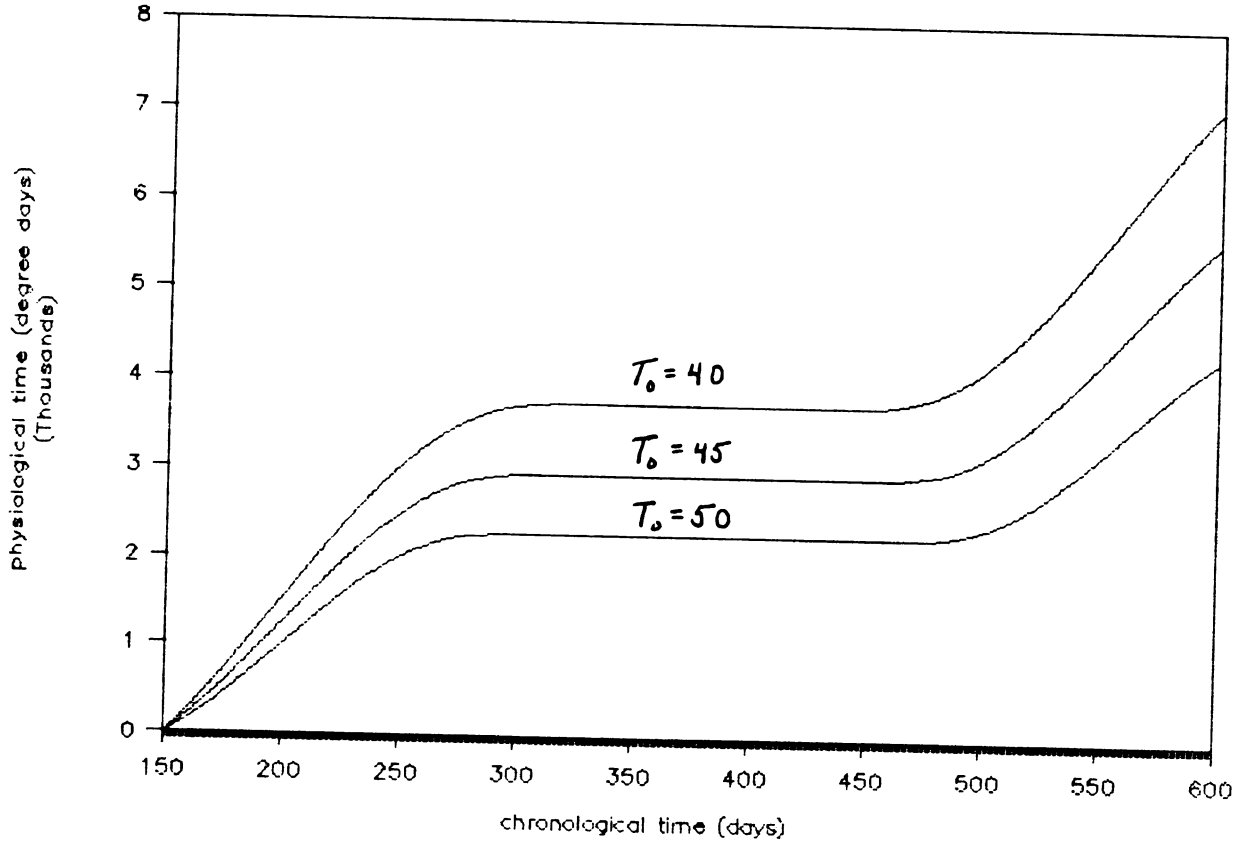


Figure 7.10c: Mean physiological time plotted against chronological time for various temperature thresholds.

t_0 = 150 days
T_0 = 40°, 45°, 50° F

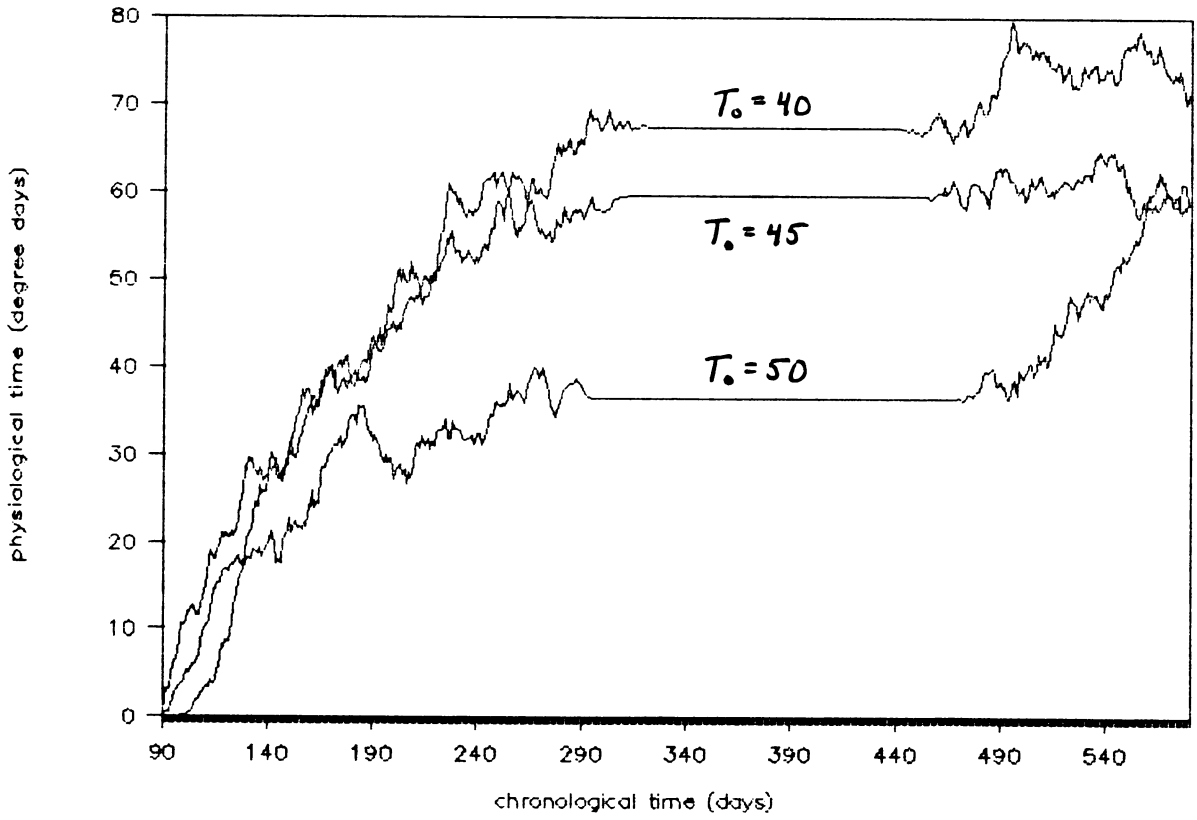


Figure 7.11a: Standard deviation of physiological time plotted against chronological time for various temperature thresholds.

t_0 = 90 days
T_0 = 40°, 45°, 50° F

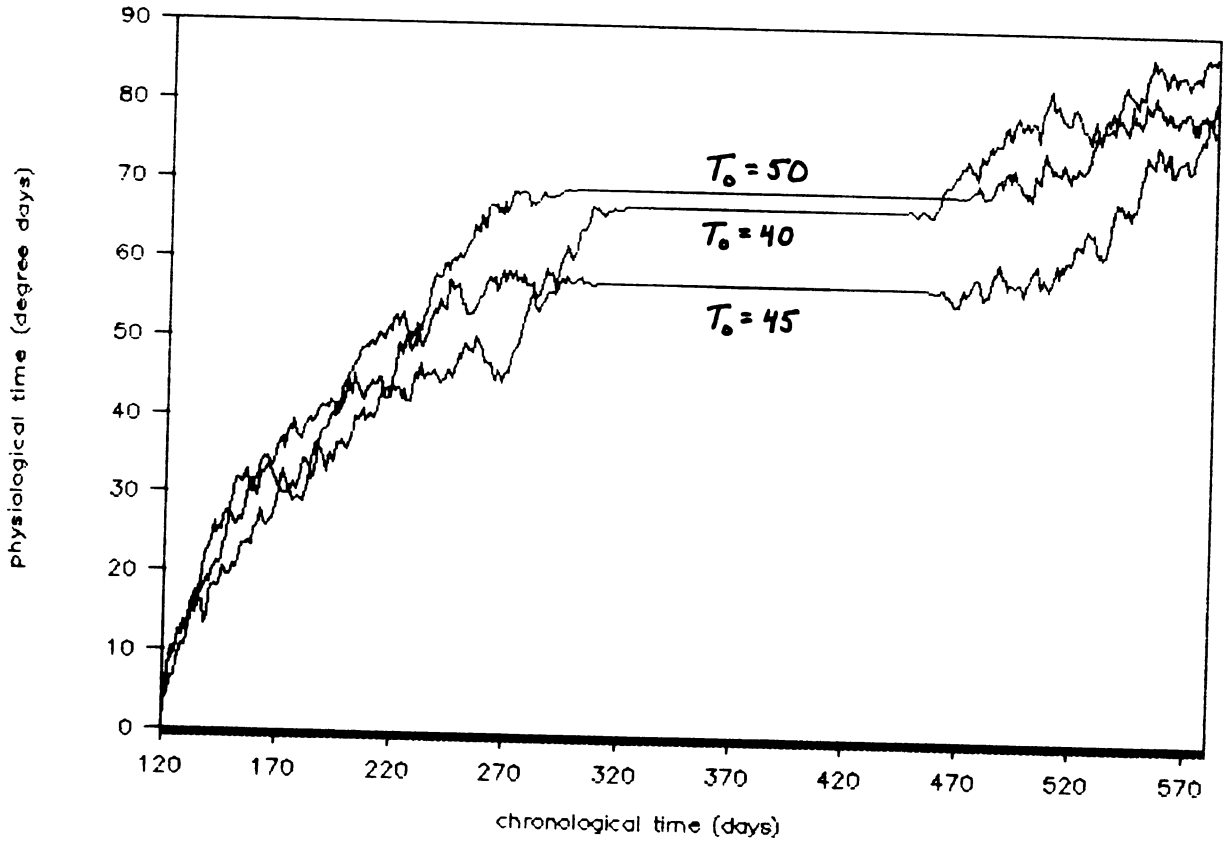


Figure 7.11b: Standard deviation of physiological time plotted against chronological time for various temperature thresholds.

t_0 = 120 days
T_0 = 40°, 45°, 50° F

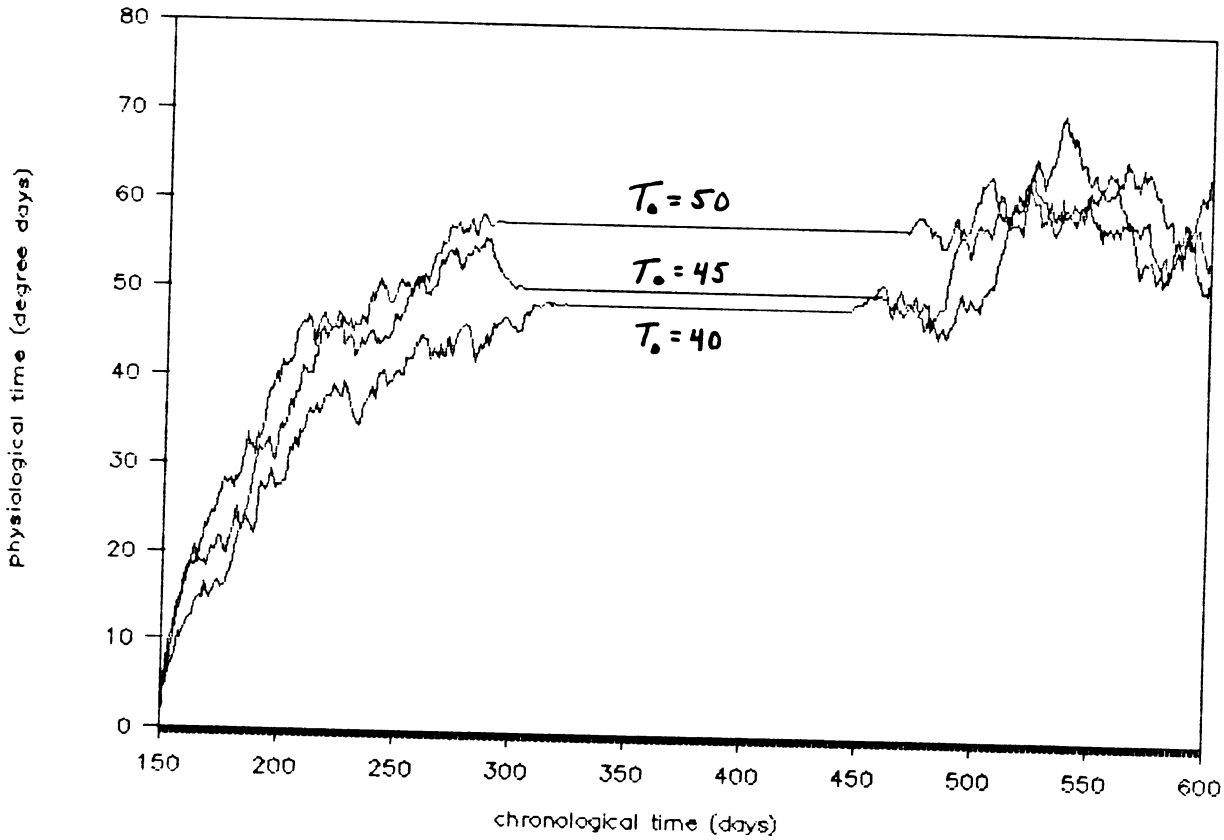


Figure 7.11c: Standard deviation of physiological time plotted against chronological time for various temperature thresholds.

t_0 = 150 days
T_0 = 40°, 45°, 50° F

STEADY STATE PHYSIOLOGICAL TIME PROPERTIES

The simulations of the previous section considered only a single season's growth. This is indeed a typical situation, especially when modeling agricultural production. However, it is often necessary to model over many year's time, and perhaps ultimately toward a steady-state ideal. We now consider the asymptotic behavior of $\tau(t)$ as $t \to \infty$.

The physiological time is a monotonically increasing function. However, the function

$$[7.16] \qquad \Omega(t) = \frac{\tau(t)}{\|\Gamma(t)\|}$$

$$= \frac{\int_{\Gamma(t)} [T(\mu)-T_0]d\mu}{\int_{\Gamma(t)} d\mu} \quad \text{for Type I models}$$

is clearly not. This function $\Omega(t)$ might be thought of as an **average temperature in excess of the lower threshold.** Therefore, we refer to this as the **effective ratio.** Over

the long run such a function has desirable statistical properties.

Consider the graph of the reasonably common situation shown in Figure [4.2]. Both functions $\tau(t)$ and $\|\Gamma(t)\|$ are monotonic increasing and both increase simultaneously only over the active region. In general the ratio $\Omega(t)$ will vary with time, but if $\tau(t)$ is not too erratic, the ratio $\Omega(t)$ will at least be bounded.

That this is so for any useful heat unit system should come as no surprise. The average excess temperature is well defined only for suitable temperature profiles $T(t)$, and certainly these must be bounded. However, we need to make an even stronger statement. If $\Omega(t)$ is not only bounded but also approaches a steady-state constant (independent of t) it will be possible to show that the **stochastic** variable obtained by randomizing $T(t)$ will be well behaved in the statistical sense outlined below.

As t increases, the randomized $T(t)$ will produce a stochastic $\tau(t)$. Since $\tau(t)$ is defined dynamically (as an integral), variances in $\tau(t)$ will produce compounded variances in $\tau(t+1)$. This will produce a "fan-out" effect so that $\tau(t)$ can ultimately (in principle) become any value desired. In order for this not to happen in practice we require that:

[7.17] $$\text{Prob}\{\tau(t)>E[\tau(t)]\} = \text{Prob}\{\tau(t+k)>E[\tau(t+k)]\}$$ for all times t and k.

That is to say, the likelihood an instance of the ensemble deviates from the mean is unchanged for all values of time. This can be controlled in the following manner:

> If $\text{Var}[\tau(t)]$ is bounded, there will be effectively (in the statistical sense defined by equation [7.17]) no unbounded spread in $\tau(t)$. In this case we refer to the random process $\tau(t)$ as **locally randomized.**

We proceed to show that $\tau(t)$ is locally randomized. This is shown in Theorem 7.1.

Theorem 7.1: Consider a temperature profile described by a sinusoid:

$$T(t) = A + B \cdot \cos(2\pi t/p + \theta) \qquad [7.18]$$

with period p. Also assume a Type I endogenous model with a single lower threshold T_0. Then for k cycles (with odd active intervals and even inactive intervals),

$$\tau(t) = \begin{cases} \dfrac{B \cdot p \cdot k}{\pi}\left[\sqrt{(1-z^2)} - z \cdot \cos^{-1}(z)\right] - B \cdot z\left[t - \sum_{i=1}^{k} t_{2i-1}\right] \\ \qquad \text{for even intervals: } k = \text{even} \\ \dfrac{B \cdot p}{2 \cdot \pi}\left[k \cdot \sqrt{(1-z^2)} + \sin(2\pi t/p + \theta)\right] - B \cdot z\left[t - \sum_{i=1}^{k} t_{2i-1}\right] \\ \qquad \text{for odd intervals: } k = \text{odd} \end{cases} \qquad [7.19]$$

$$\|\Gamma(t)\| = \begin{cases} \dfrac{p \cdot k}{\pi}\left[\pi - \cos^{-1}(z)\right] & \text{for } k = \text{even} \\ \dfrac{p}{2\pi}\left[(2\pi t/p + \theta) + k \cdot \cos^{-1}(z)\right] & \text{for } k = \text{odd} \end{cases} \qquad [7.20]$$

[7.21]
$$\Omega_k(t) = \begin{cases} B\cdot\left[\dfrac{\sqrt{(1-z^2)}}{\pi - \cos^{-1}(z)} - z\right] & \text{for even intervals} \\ B\cdot\left[\dfrac{\sqrt{(1-z^2)} + \sin(2\pi t/p+\Theta)}{(2\pi t/p+\Theta) - k\cdot\cos^{-1}(z)}\right] & \text{for odd} \end{cases}$$

Also, as $k\to\infty$, the steady-state function of $\Omega_k(t)$ asymptotically approaches the constant

$$\Omega = \Omega(t) = \lim_{k\to\infty} \frac{\tau(t)}{\|\Gamma(t)\|}$$

[7.23]
$$= B\cdot\left[\frac{\sqrt{(1-z^2)}}{\pi - \cos^{-1}(z)} - z\right]$$

[7.24]
$$\text{where } z = \frac{T_0 - A}{B}$$

Proof. Consider the sinusoid

$$T(t) = A + B\cdot\cos(2\pi t/p+\theta)$$

The critical points of the action set $\Gamma(t)$ are given by

[7.25] $$A + B\cdot\cos(2\pi t/p\theta) = T_0.$$

Defining $z = (T_0-A)/B$, the initial points of each interval of the action set are given by :

[7.26] $$t_{2k-1} = \frac{p}{2\pi}\,[-\theta + \cos^{-1}(z)] + (k-1)\cdot p$$

for k=1,2,...

Similarly, the terminal points for each interval are:

[7.27] $$t_{2k} = -\frac{p}{2\pi}\,[\,\theta + \cos^{-1}(z)] + k\cdot p$$

for k=1,2,...

Case 1: Suppose t is in an even interval, that is to say, on the inaction set:

$$I_{2n}:\ t_{2n} < t < t_{2n+1} \quad \text{for some n.}$$

The action set $\Gamma(t)$ contains n cycles, each with equal physiological time contributions:

$$\tau(t) = \sum_{k=1}^{n} \int_{t_{2k-1}}^{t_{2k}} [T(t) - T_0]dt$$

[7.28]

$$= n \cdot \int_{t_1}^{t_2} [T(t) - T_0]dt$$

$$= \frac{B \cdot p \cdot n}{\pi} \left[\sqrt{(1-z^2)} - z \cdot [\pi - \cos^{-1}(z)]\right]$$

Also, the measure of $\Gamma(t)$ is given by:

$$\|\Gamma(t)\| = \sum_{k=1}^{n} (t_{2k} - t_{2k-1})$$

[7.29]

$$= n \cdot (t_2 - t_1)$$

$$= \frac{p \cdot n}{\pi} [\pi - \cos^{-1}(z)]$$

Therefore, the effective ratio is:

$$\Omega_n(t) = \frac{\frac{B \cdot p \cdot n}{\pi}\left[\sqrt{(1-z^2)} - z \cdot [\pi - \cos^{-1}(z)]\right]}{\frac{p \cdot n}{\pi}[\pi - \cos^{-1}(z)]}$$

[7.30]

$$= B \cdot \left[\frac{\sqrt{(1-z^2)}}{\pi - \cos^{-1}(z)} - z\right]$$

Case 2: Suppose t is in an odd interval, that is to say, on the action set:

$$I_{2n-1}: \quad t_{2n-1} < t < t_{2n}$$

Here the action set contains (n-1) completed cycles and a partial n^{th} cycle. The physiological time for the partial cycle is:

$$\tau(t) = \int_{t_{2n-1}}^{t} [T(t) - T_0]dt$$

[7.31]

$$= \frac{B \cdot p}{2 \cdot \pi}[\sin(2\pi t/p + \Theta) + \sqrt{(1-z^2)}] - B \cdot z \cdot (t - t_{2n-1})$$

Combining this with the (n-1) completed cycles, the total physiological time is:

$$\tau(t) = \sum_{k=1}^{n-1} \int_{t_{2k-1}}^{2k} [T(t) - T_0]dt + \int_{t_{2n-1}}^{t} [T(t) - T_0]dt$$

[3.32]

$$= \frac{B \cdot p \cdot n}{\pi} \left[\sqrt{(1-z^2)} - z \cdot [\pi - \cos^{-1}(z)]\right]$$

$$+ \frac{B \cdot p}{2 \cdot \pi} [\sin(2\pi t/p+\theta)+\sqrt{(1-z^2)}]-B \cdot z \cdot (t-t_{2n-1})$$

$$= \frac{B \cdot p}{2 \cdot \pi} \left[n \cdot \sqrt{(1-z^2)} + \sin(2\pi t/p+\theta)\right] - B \cdot z \left[t - \sum_{i=1}^{2n-1} t_{2i-1}\right]$$

Also, the measure of $\Gamma(t)$ is given by:

$$\|\Gamma(t)\| = t + \sum_{k=1}^{n} (t_{2k} - t_{2k-1})$$

[7.33]

$$= t + \sum_{k=1}^{2n-1} (-1)^k \cdot t_k$$

$$= \frac{p}{2 \cdot \pi} [(2\pi t/p+\theta) + k \cdot \cos^{-1}(z)]$$

Therefore, the effective ratio is:

$$\Omega_n(t) = \frac{\frac{B\cdot p}{2\cdot\pi}\left[n\cdot\sqrt{(1-z^2)}+\sin(2\pi t/p+\theta)\right]}{\frac{p}{2\cdot\pi}\left[(2\pi t/p+\theta)+k\cdot\cos^{-1}(z)\right]} - \frac{B\cdot z\left[t - \sum_{i=1}^{2n-1} t_{2i-1}\right]}{t + \sum_{k=1}^{2n-1}(-1)^k\cdot t_k}$$

[7.34]

$$= B\cdot\left[\frac{\sqrt{(1-z^2)} + \sin(2\pi t/p+\theta)}{(2\pi t/p+\theta) - k\cdot\cos^{-1}(z)}\right]$$

In summary,

[7.35]

$$\Omega_k(t) = \begin{cases} B\cdot\left[\dfrac{\sqrt{(1-z^2)}}{\pi - \cos^{-1}(z)} - z\right] & k=\text{even} \\ B\cdot\left[\dfrac{\sqrt{(1-z^2)} + \sin(2\pi t/p+\theta)}{(2\pi t/p+\theta) - k\cdot\cos^{-1}(z)}\right] & k=\text{odd} \end{cases}$$

It remains but to consider the steady-state behavior of $\Omega_k(t)$. To do this requires a minor adjustment for $\|\Gamma(t)\|$ on the odd interval.

$$\|\Gamma(t)\| = \frac{p}{2\cdot\pi}\,[(2\pi t/p+\theta) + k\cdot\cos^{-1}(z)]$$

[7.36]

$$= \frac{p}{2\pi}\left[\frac{2\pi}{p}(t-t_k)+\theta) + \frac{2\pi}{p}(t_k) + k\cdot\cos^{-1}(z)\right]$$

$$= \frac{p}{2\pi}\left[(2\pi t/p+\theta)-(2\pi t_k/p+\theta)+\theta-[-\theta+\cos^{-1}(z)\right.$$

$$\left. + \frac{2\pi}{p}\,[(k-1)\cdot p/2]+k\cdot\cos^{-1}(z)\right]$$

$$= \frac{p}{2\pi}\left[(2\pi t/p+\theta) - (2\pi t_k/p+\theta)\right.$$

$$\left. + (k-1)\cdot[\pi - \cos^{-1}(z)]\right]$$

for odd k.

Therefore,

[7.37] $$\Omega_k(t) = B\cdot\left[\frac{\sqrt{(1-z^2)}+\sin(2\pi t/p+\theta)}{[2\pi(t-t_k)/p+\theta]+(k-1)[\pi-\cos^{-1}(z)]} - z\right]$$

Since the difference $(t-t_k)$ is bounded (while t and t_k individually are not), the steady-state value $\Omega(t)$ is given by equation [7.38].

$$\Omega(t) = \lim_{k\to\infty} \Omega_k(t)$$

[7.38]

$$= B\cdot\left[\frac{\sqrt{(1-z^2)}}{\pi - \cos^{-1}(z)} - z\right] \qquad \text{for all } k.$$

which is independent of t and, thus, constant.

We write:

$$\Omega = \Omega_{SS}(t)$$

[7.39]

$$= B\cdot\left[\frac{\sqrt{(1-z^2)}}{\pi - \cos^{-1}(z)} - z\right]$$

where $z = (T_0-A)/B$.

Q.E.D.

In practice we can apply this theorem by allowing the following substitutions:

A = average [annual] mean temperature

[7.40]

$$= \frac{1}{p} \cdot \sum_{t=1}^{p} M(t)$$

B = average [annual] temperature range

[7.41]

$$= \frac{1}{p} \cdot \sum_{t=1}^{p} R(t)$$

as a course [sinusoidal] model.

Formula [7.23] is not as practical as might be first supposed. Yes, the physiological time is asymptotically proportional to the norm of the action set, but the action set itself is not easily found.

As a verification, the results of several simulations are shown in Figures 7.10 and 7.11. Even though the simulations were run using the empirical Lansing data, which is not truly sinusoid, the results are seen to be quite good! As in chapter three we are encouraged to use these sinusoid results even for arbitrary, but periodic stochastic models.

Aside from the aesthetically pleasing results derived in Theorem 7.1, this steady-state formula immediately proves the physiological time to be a local randomization. Since the steady-state value of $\Omega(t)$ exists (thus is a constant), the expectation is bounded by continuity. Theorem 7.2 states this fact assuming the amplitude to be a uniform distribution. Even though earlier results show that a uniform distribution is unrealistic for temperature profiles, it is certainly no worse than assuming the sinusoid. Simulations validate this approach.

Theorem 7.2: Suppose the Temperature Profile is modeled as a sinusoid:

$$T(t) = A + B\cdot\cos(2\pi t/p+\theta)$$

with period p and B being a uniformly distributed random variable on the interval $\mu-\frac{1}{2}r<\beta<\mu+\frac{1}{2}r$. Also assume a Type I endogenous model. Then for k cycles (with active intervals odd, inactive even), T(t) is a random process with:

$$\begin{aligned} \text{Exp}[\Omega(t)] &= \lim_{k\to\infty} [\Omega_k(t)] \\ &= (1-e)\cdot\mu + (1+e)\cdot(A+T_0) \end{aligned} \qquad [7.42]$$

$$\begin{aligned} \text{Var}[\Omega(t)] &= \lim_{k\to\infty} [\Omega_k(t)] \\ &= \text{Exp}[\Omega^2(t)] - \text{Exp}^2[\Omega(t)] \\ &= (1+e)\cdot[A\cdot(2e-1) + T_0\cdot(2e-3)]\cdot\mu \\ &\quad + [r^2\cdot(1-e)/12 + (1+e)^2\cdot(A+T_0)^2 \end{aligned} \qquad [7.43]$$

where e is chosen on the interval $0<e<\frac{1}{2}$.

Proof: Consider the function:

$$f(z) = \begin{cases} \dfrac{\sqrt{(1-z^2)}}{\pi - \cos^{-1}(z)}, & -1<z\leq 1 \\ 1 & z=-1 \end{cases} \qquad [7.44]$$

This function is well defined on $-1 \leq z \leq 1$.

Also, by l'Hospital's rule, f is also

continuous on this interval. A quick sketch of the graph yields (Figure 7.14) three convenient properties.

i. $0 \le f(z) \le 1$

ii. $f(-1)=1$ and $f(1)=0$

iii. f is almost linear:

$$f(z) \approx 1 - e \cdot (z+1) \equiv f'(z)$$

where e is chosen such that $0 \le e \le \frac{1}{2}$.

For example, if $e=0$, $f'(z)$ gives a simple upper bound of $f(z)$.

If $e=\frac{1}{2}$, $f'(z)=\frac{1}{2}(1-z)$ gives a lower bound for $f(z)$.

If $e \approx 0.35$, $f'(z)=(13-7 \cdot z)/20$ gives a good approximation on the interval $-1 \le z \le \frac{1}{2}$.

We use this approximation to establish both upper bounds and reasonable approximations

for the expected value [mean] and variance [standard deviation] of the steady-state value of the effective ratio $\Omega_{SS}(t)$.

Case 1: The expected value.

Since the distribution of A is taken to be uniform on the interval $\mu-\frac{1}{2}r \le \beta \le \mu+\frac{1}{2}r$,

$$[7.45] \qquad \mathrm{Exp}[\Omega_{SS}(t)] = \frac{1}{r}\int_{\mu-\frac{1}{2}r}^{\mu+\frac{1}{2}r}\left[\frac{\sqrt{(1-z^2)}}{\pi - \cos^{-1}(z)}\right]\cdot\beta d\beta$$

where $z = (T_0-A)/\beta$

Using the linear approximation

$$[7.46] \qquad \frac{\sqrt{(1-z^2)}}{\pi - \cos^{-1}(z)} \approx 1 - e\cdot(1-z)$$

$$\mathrm{Exp}[\Omega_{SS}(t)] \approx \frac{1}{r}\int_{\mu-\frac{1}{2}r}^{\mu+\frac{1}{2}r}[1-e\cdot(z+1)-z]\cdot\beta d\beta$$

$$[7.47] \qquad = \frac{1}{r}\int_{\mu+\frac{1}{2}r}^{\mu+\frac{1}{2}r}\left[1-e\cdot\left[1+\frac{T_0-A}{\beta}\right]-\frac{T_0-A}{\beta}\right]\cdot\beta d\beta$$

$$= (1-e)\cdot\mu + (1-e)\cdot(A+T_0)$$

Case 2: Variance.

Similarly, we approximate the second moment.

$$\mathrm{Exp}[\Omega_{SS}{}^2(t)] \approx \frac{1}{r}\int_{\mu-\frac{1}{2}r}^{\mu+\frac{1}{2}r} [1-e\cdot(z+1)-z]\cdot\beta^2 d\beta$$

[7.48]

$$= \frac{1}{r}\int_{\mu-\frac{1}{2}r}^{\mu+\frac{1}{2}r} \left[1-e\cdot\left[1+\frac{T_0-A}{\beta}\right]-\frac{T_0-A}{\beta}\right]\cdot\beta^2 d\beta$$

$$= (1-e)\cdot[\mu^2 + r^2/12] + \mu\cdot(A-T_0)\cdot(1+e)$$

The variance is found by...

$$\mathrm{Var}[\Omega_{SS}(t)] = \mathrm{Exp}[\Omega^2(t)] - \mathrm{Exp}^2[\Omega(t)]$$

[7.49]

$$= (1+e)\cdot[A\cdot(2e-1) + T_0\cdot(2e-3)]\cdot\mu + [r^2\cdot(1-e)/12 + (1+e)^2\cdot(A+T_0)^2$$

Q.E.D.

Similar formulas can be derived considering A to be a uniformly distributed random variable on the interval:

$$\mu-\tfrac{1}{2}r \le \alpha \le \mu+\tfrac{1}{2}r$$

They are as follows:

[7.51]

$$\mathbf{Exp}[\Omega_{SS}(t)] \approx [1 - e - \frac{T_0 \cdot (1+e)}{B}] \cdot \mu + \frac{1+e}{B} [\mu^2 + r^2/12]$$

[7.52]

$$\mathbf{Exp}[\Omega_{SS}{}^2(t)] \approx [1 - e - \frac{T_0 \cdot (1+e)}{B}] \cdot [\mu^2 + r^2/12]$$

$$+ \frac{1+e}{B} [2 \cdot \mu + r] \cdot [4\mu^2 + r^2]$$

The variance can be computed using equation [7.49].

The results of this theorem indicate that the mean, and especially the standard deviation of any long range view of physiological time are bounded as compared to the norm $\|\Gamma(t)\|$. Thus, at no time do we anticipate runaway variance and we can expect only local randomization effects.

Perhaps the best illustration of this point is to observe the simulation results of Figure 7.15. This graph is a plot of the effective ratio from 100 runs over a 5 year period using various thresholds. We note that as predicted, the family members are roughly "parallel" and the asymptotic limits are well defined, finite values.

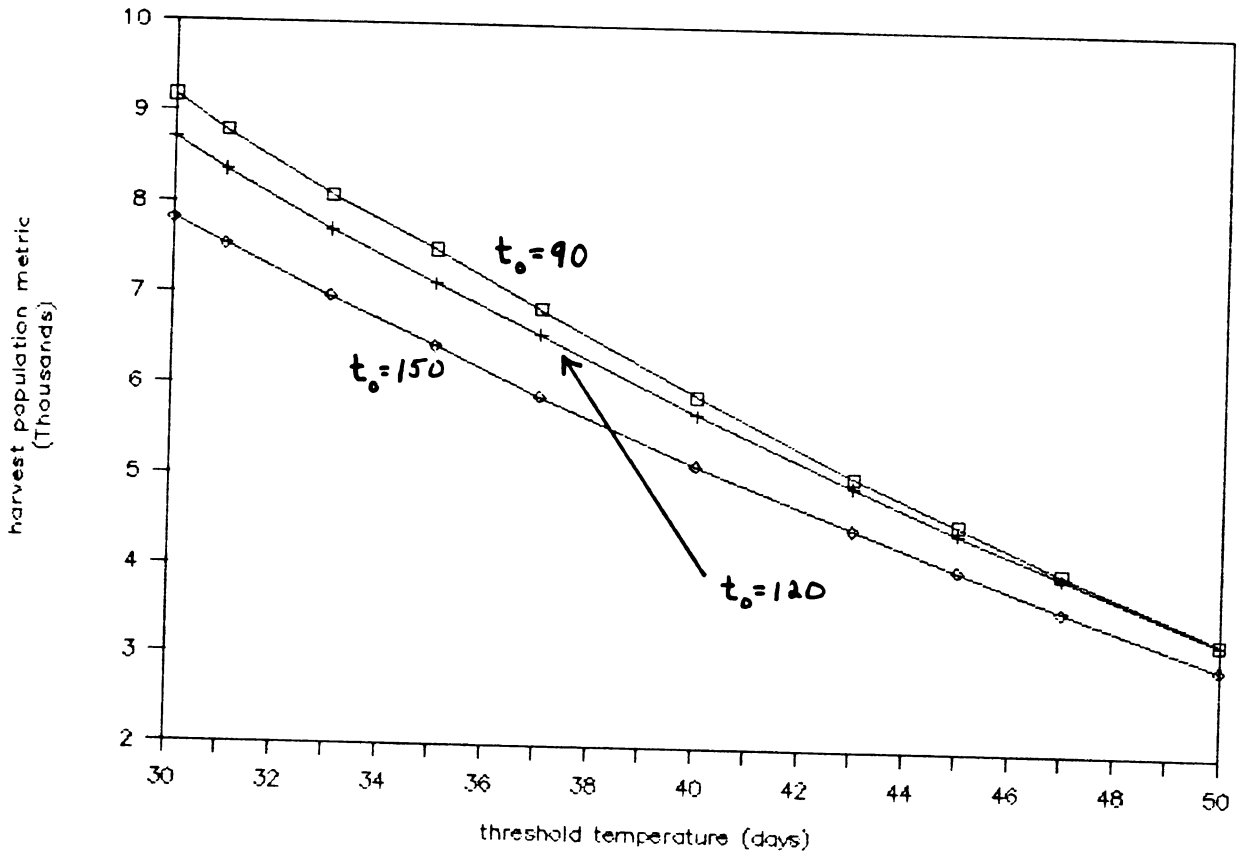


Figure 7.12: Mean final physiological time as a function of threshold for various initial times. T=40° F.

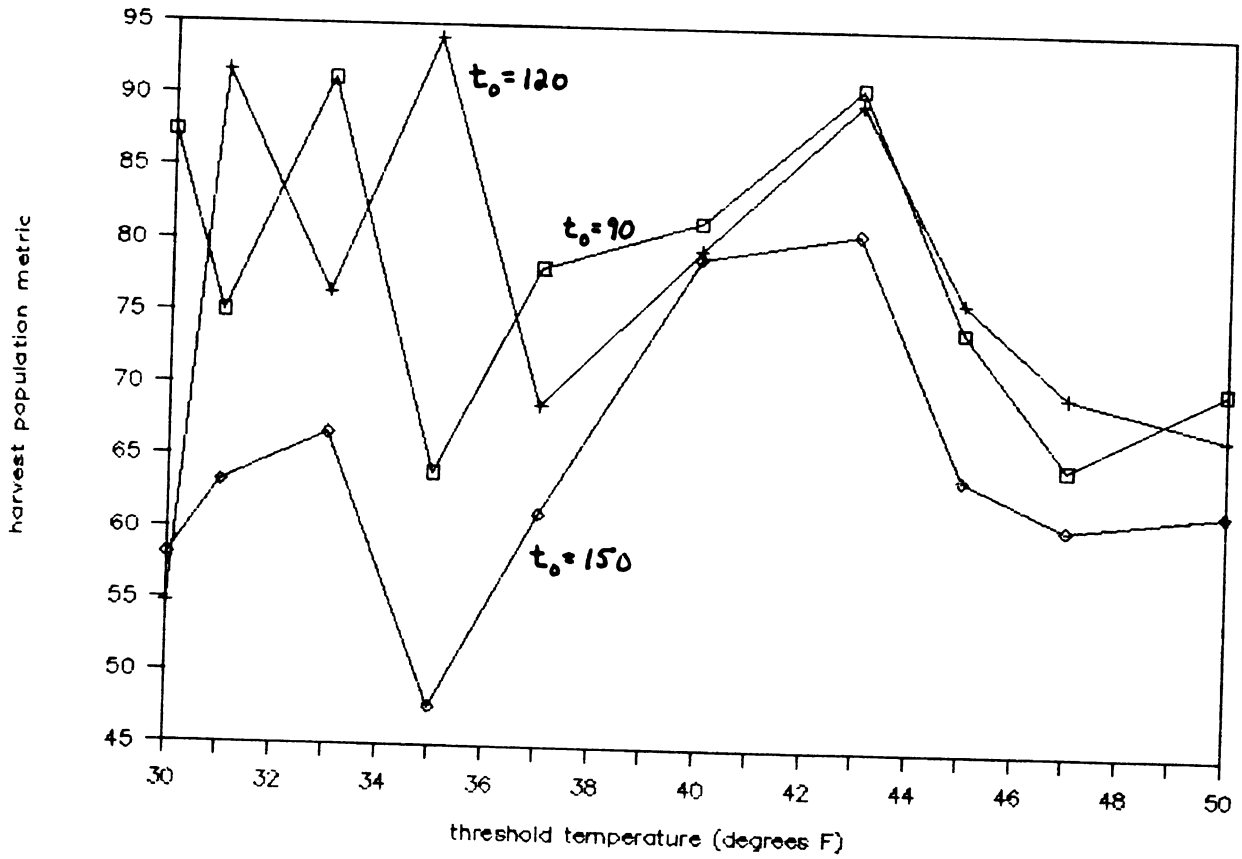


Figure 7.13: Standard deviation of final physiological time as a function of threshold for various initial times. T=40° F.

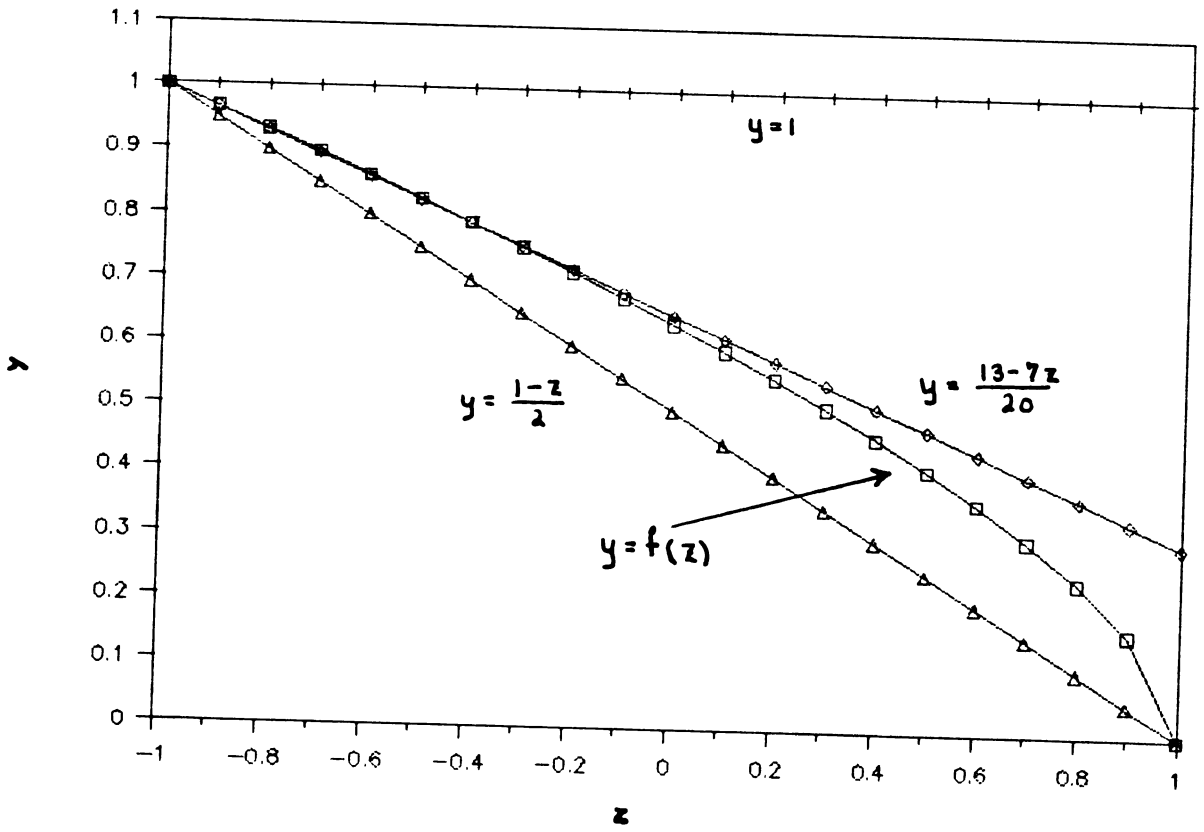


Figure 7.14: Graph of the function defined in equation 7.44 along with various approximations to f(z).

$$f(z) = \begin{cases} \dfrac{\sqrt{(1-z^2)}}{\pi - \cos^{-1}(z)}, & -1<z\leq 1 \\ 1 & z=-1 \end{cases}$$

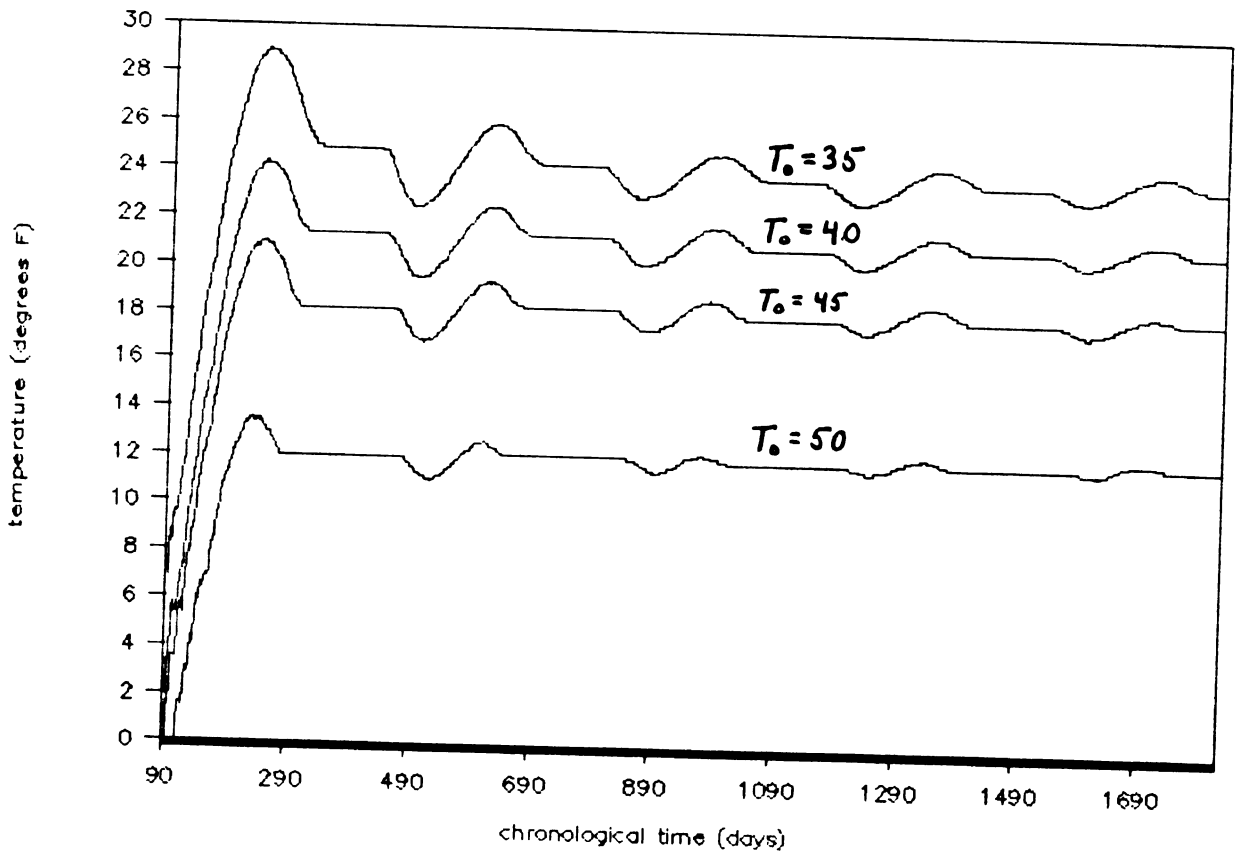


Figure 7.15: The effective ratio over k=5 cycles using various temperature thresholds.

STATISTICAL PROPERTIES OF POPULATION SIZE

Thus far it has been demonstrated that both the temperature profile $T(t)$ and physiological time $\tau(t)$ possess very tolerable statistical properties. The physiological time does indeed behave as "time" should.

It is, therefore, no surprise $s(t)$ is also dependent only on local randomization effects. That is to say, as time increases, the solution $x(t)$ to a well posed stochastic modeling problem will not possess compounding randomizing effects. Said another way, the variance of $x(t)$ does not "grow" over time and we consider any random effects therefore, to be strictly local phenomena. This does not mean that extreme values are not possible. The distribution does, in fact, diverge from the mean. However, the **standard deviation** is invariant over time. The solution is bounded in that sense. Therefore, we think of this as demonstrating **local** randomization effects as opposed to **compounding** or **global** randomization effects.

EXAMPLE THREE

This local randomization effect can be demonstrated for a logistic system. Consider a logistic (single exogenous parameter) population model with:

$r = 0.001$ $t_0 = 120$ days $x_m = 100$

$T_0 = 45^o$ $x_0 = 1$

Lansing 1985 weather data.

The simulation for this model was run 100 times with the results as shown in Figures 7.16 and 7.17. We observe that although the mean population increases toward the system capacity x_m=100, the statistical variance of the population ensemble remains bounded. Thus, as expected, randomization effects are strictly local.

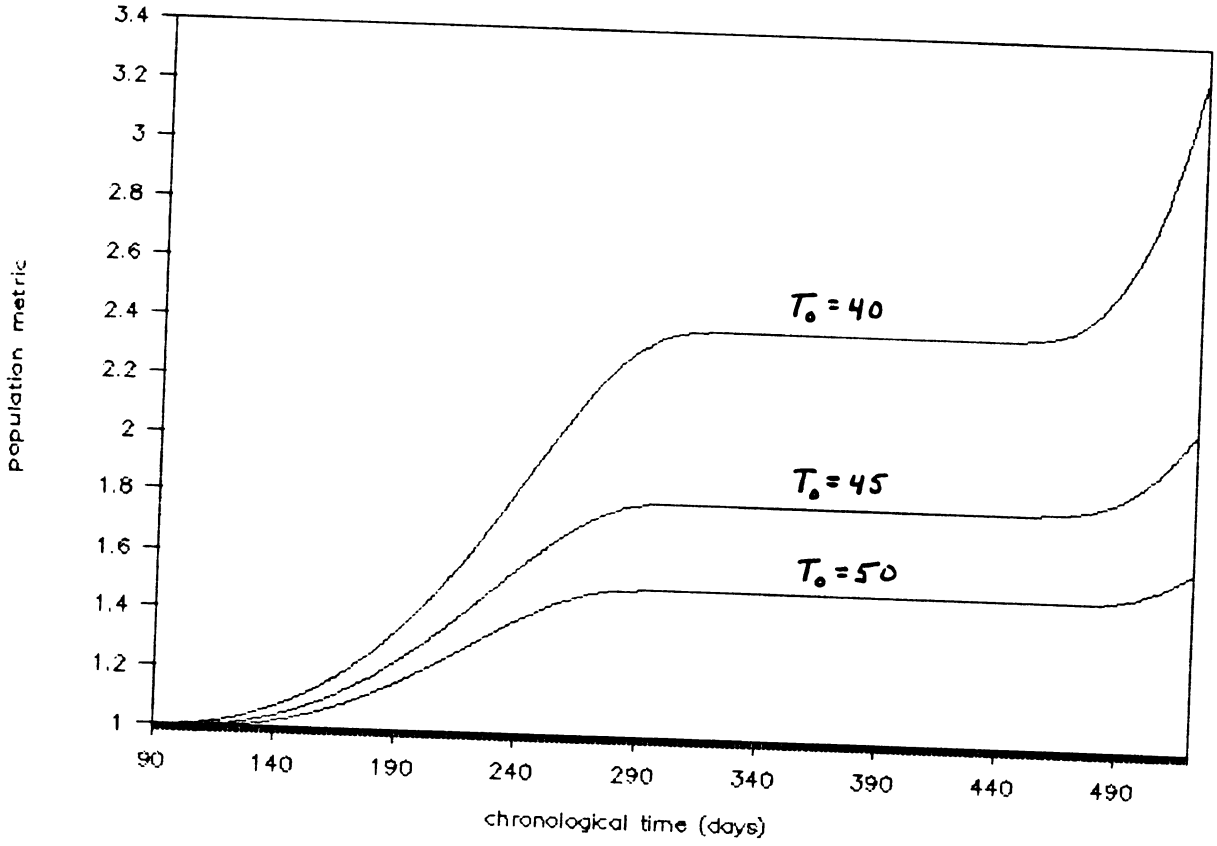


Figure 7.16a: Mean population metric as a function of time for various thresholds.

t_0 = 90 days.
T_0 = 40°, 45°, 50° F.

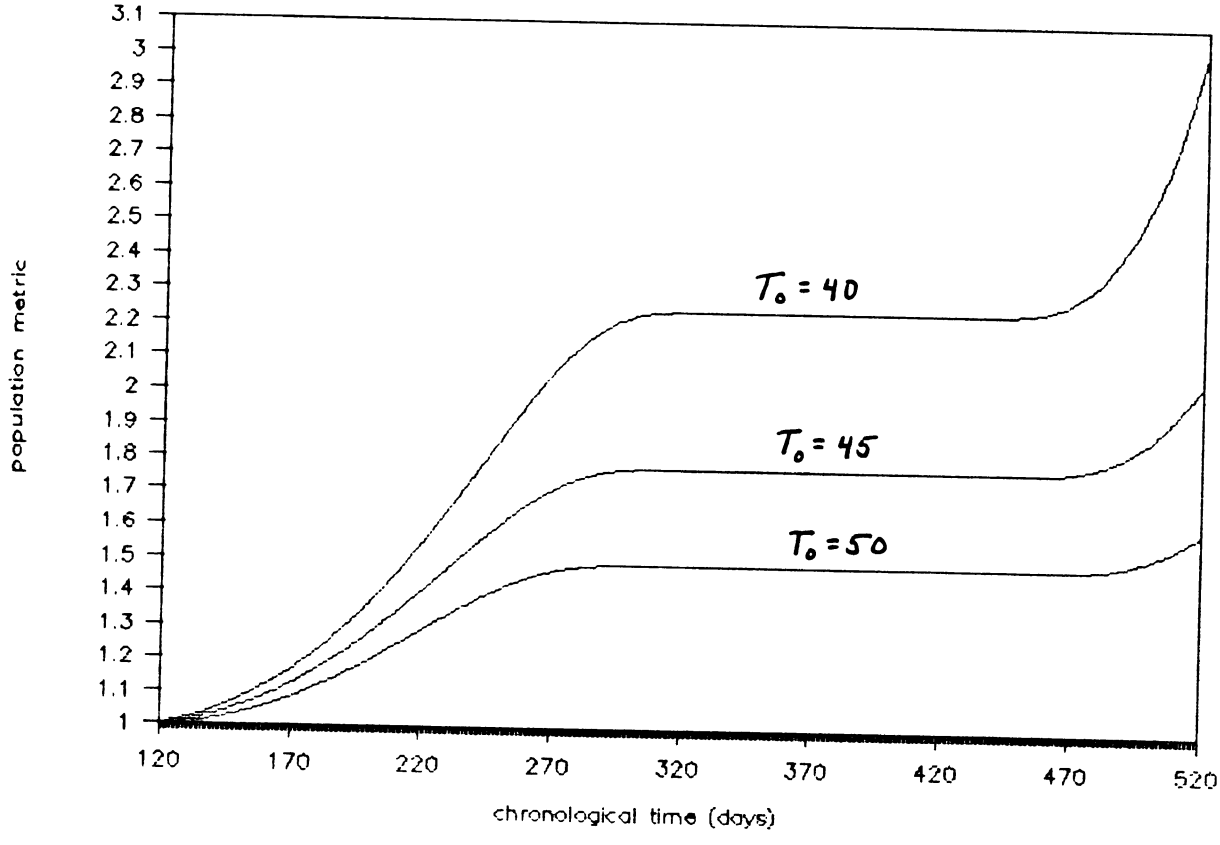


Figure 7.16b: Mean population metric as a function of time for various thresholds.

t_0 = 120 days.
T_0 = 40°, 45°, 50° F.

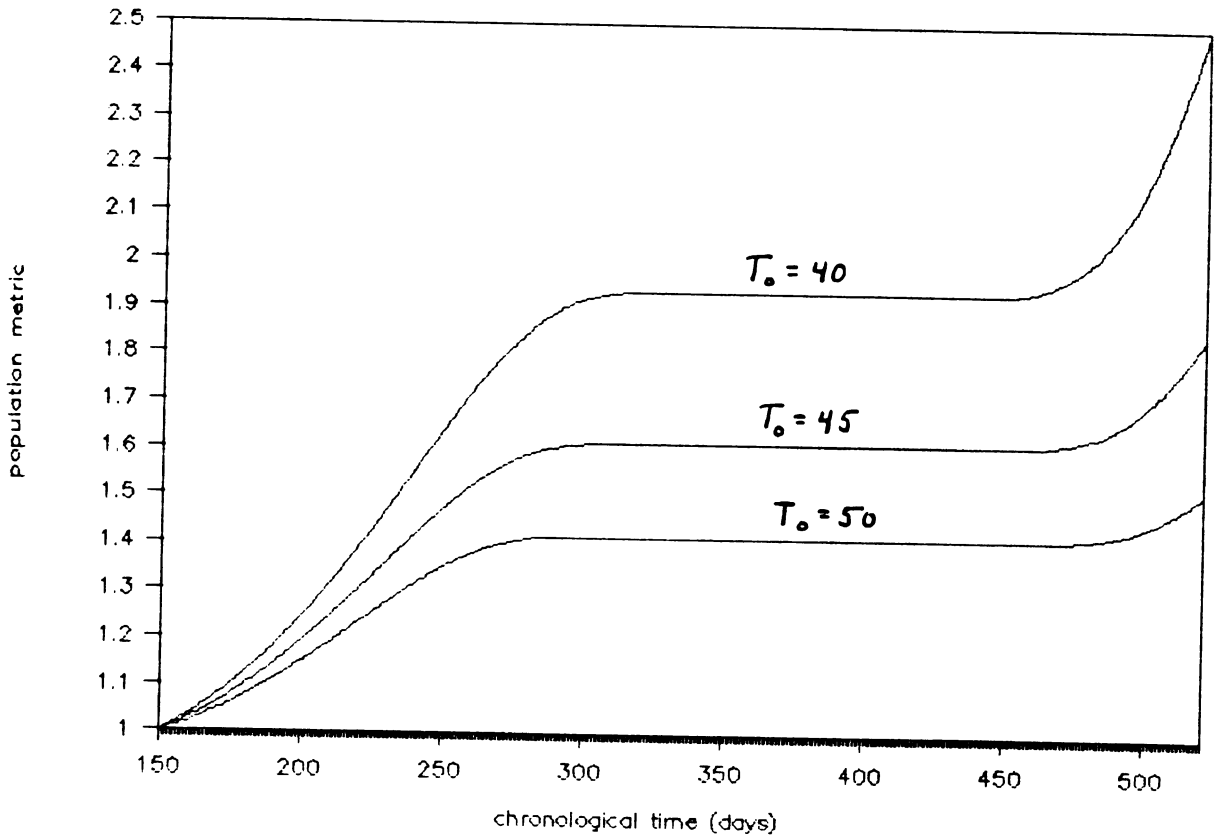


Figure 7.16c: Mean population metric as a function of time for various thresholds.

t_0 = 150 days.
T_0 = 40°, 45°, 50° F.

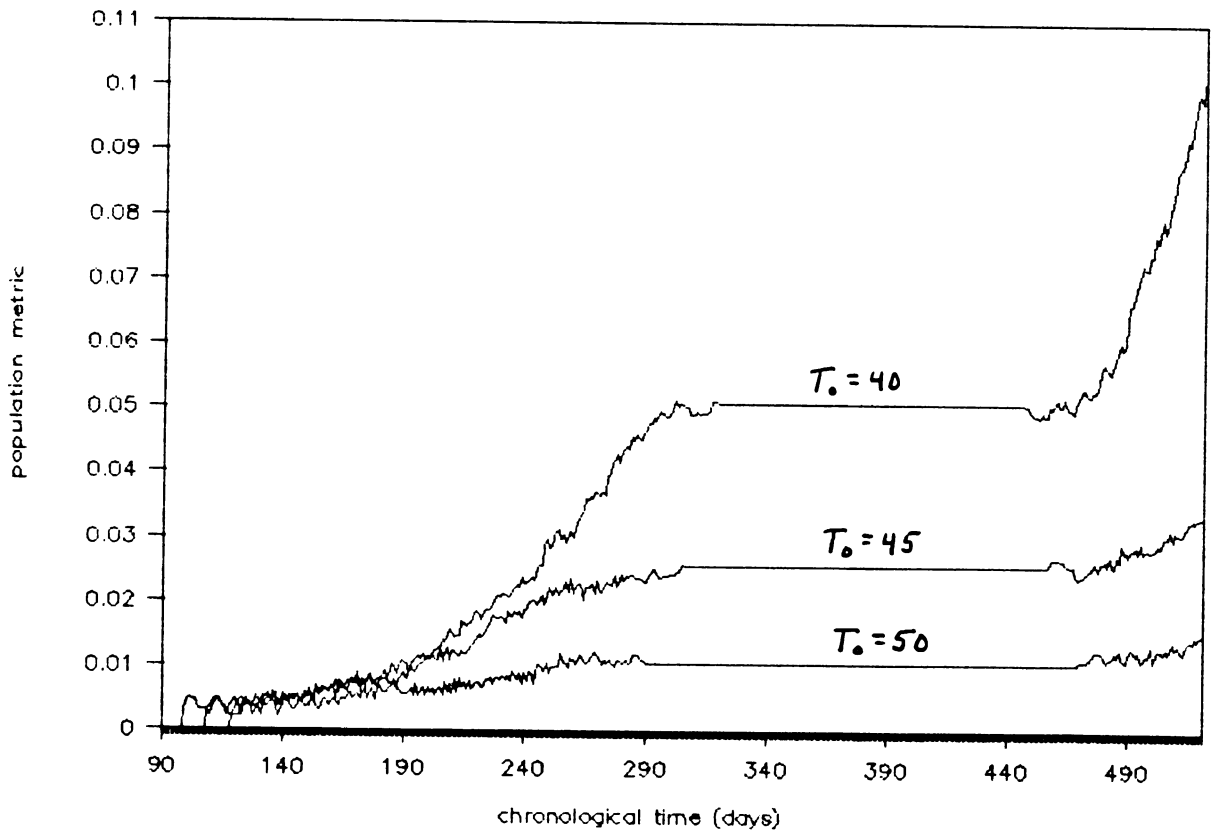


Figure 7.17a: Standard deviation of the population metric as a function of time for various thresholds.

t_0 = 90 days.
T_0 = 40°, 45°, 50° F.

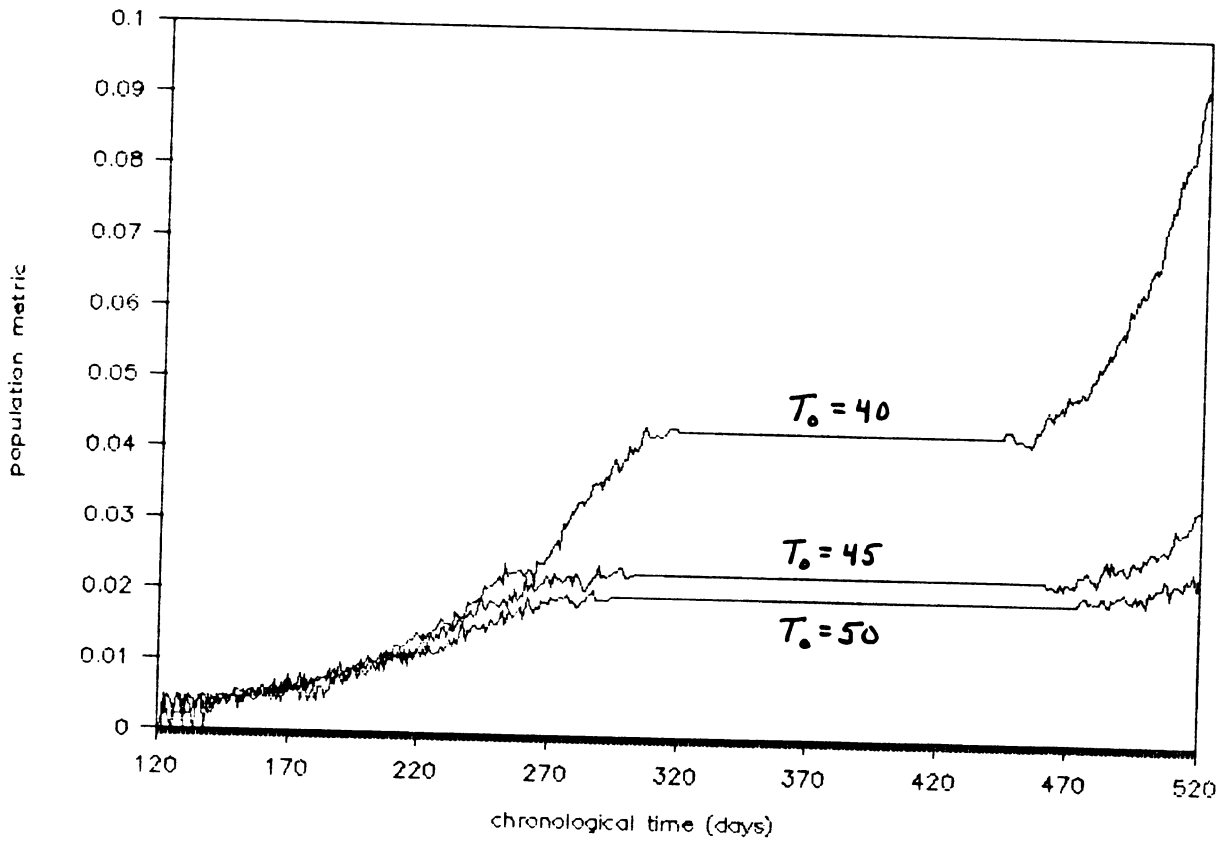


Figure 7.17b: Standard deviation of the population metric as a function of time for various thresholds.

t_0 = 120 days.
T_0 = 40°, 45°, 50° F.

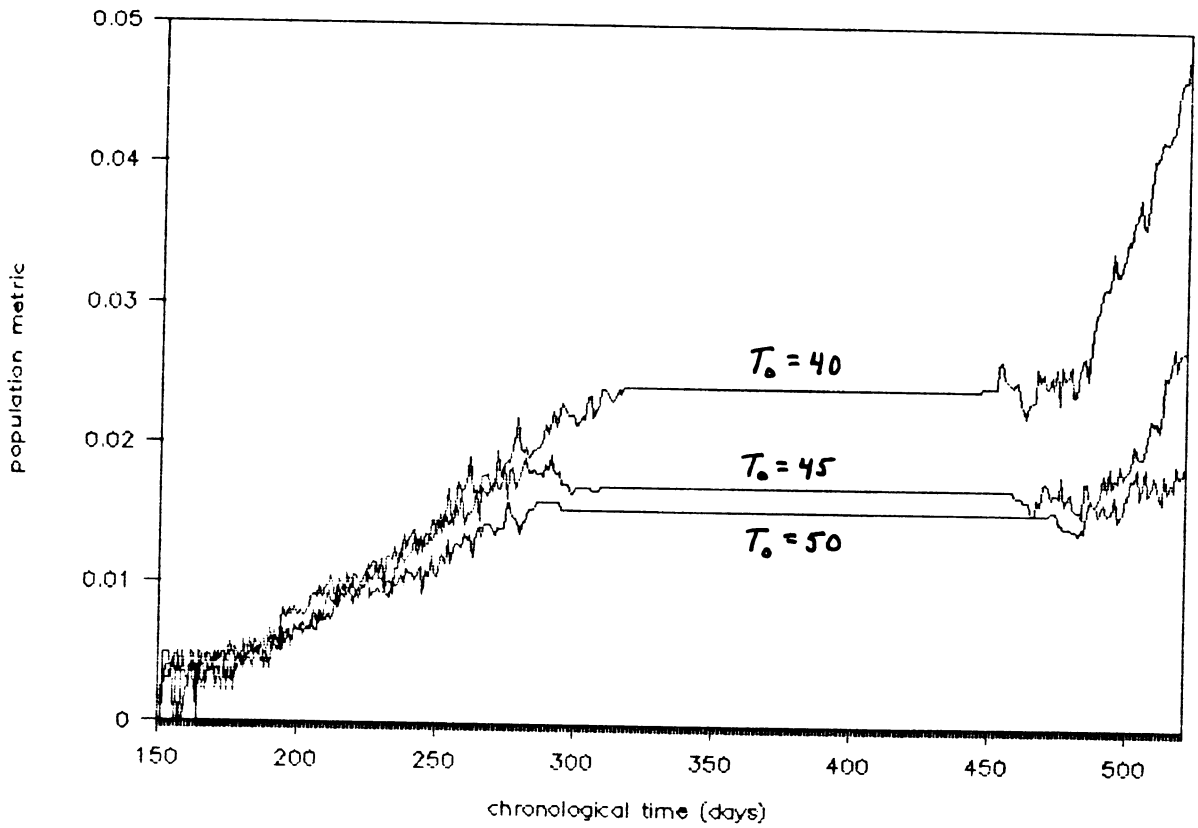


Figure 7.17c: Standard deviation of the population metric as a function of time for various thresholds.

t_0 = 150 days.
T_0 = 40°, 45°, 50° F.

CONCLUSION

The purpose of this chapter has been to introduce a systematic stochastic approach into heat-unit models. We have developed differing degrees of statistical realism into the temperature models. Algorithms were derived and justification was made by several types of simulations of the temperature profile. These models include:

i. The Deterministic Model.

While this is currently used often in practice, this model produces no stochastic effect. However, it is useful to compare this model against other, higher order, models.

ii. The Theoretic Distribution Fitting Model.

By identifying a hypothetical distribution and its parameters, it is possible to generate stochastic temperature data. In addition to mimicking the theoretic distribution, the usual descriptive statistics (in particular the mean and

standard deviation), match that of the profile basis.

iii. The Empirical Distribution Fitting Model

It is also possible to match the generated data distribution to that of the actual historical data itself. This approach is probably only valid if the profile base could be accurately approximated by a large database and the distribution is truly unique. However, as above, the descriptive statistics and distribution do match as desired.

iv. The Feedback Model

Regardless of the distribution fitting technique used, it is possible to maximize the correlation between successive daily temperatures. Indeed, we can maximize the correlation between a temperature and any other data!. While this has a strong intuitive flavor, it was discovered that there is only a weak relationship in

actuality. Even so, there is a relationship and the feedback model allows this to be exploited.

Algorithms were presented and their results illustrated for each model.

The temperature profile is directly related to the physiological time base. This is the fundamental heat-unit hypothesis! Therefore, it was necessary to demonstrate the effect that a randomized temperature profile has on physiological time. This will, in turn, affect the population dynamics of a biological system.

The essential requirement to be met is that randomization effects are local. This was proved for sinusoidal temperature profiles and was also simulated for non-sinusoidal profiles. Using empirical temperature data, strikingly agreeable results were obtained. In fact, the requirement that

$$\mathrm{Var}\left[\frac{\tau(t)}{\|\Gamma(t)\|}\right]$$

be bounded for all thresholds T_0 was more than satisfied. For agriculturally realistic thresholds, the variance was found to be not only bounded but essentially constant. Since the physiological time possesses the local randomizing effects property, it follows that the species population must also possess this property. While this was not explicitly proven, it was aptly demonstrated using simulation.

We are, therefore, not only justified but encouraged to use each of the above temperature generation models as applied to the heat unit approaches outlined in chapters five and six. In chapter nine we demonstrate this approach to population models generally.

CHAPTER EIGHT

CONTROL SYSTEMS

INTRODUCTION

Thus far we have considered the modeling and identification problems for multiple time-based heat-unit systems. Our approach has been to imbed the heat-unit hypothesis into an already existing exogenous (classical) model. Upon identification of the endogenous and exogenous parameters the system was driven by either deterministic or stochastic input. This is illustrated in Figure 8.1.

Specifically, the temperature database block was described in detail in chapters three and seven while the physiological time model was described in chapter six. A method for evaluating the [endogenous and exogenous] parameters was discussed in chapter six.

On the other hand, the population model block can take on many forms. Thus far we have confined our attention to the exponential and logistic models. In this chapter we will use the Lotka-Volterra model almost exclusively. This is hardly to say there are not other possibilities. These

macro models assume either homogenous or effectively homogenous systems. In particular, the Compartmental and Leslie models are useful and can certainly be effectively used with the heat-unit hypothesis as their basis.

Nor are these classical models the only ones that can be applied. It is recognized that the majority of predictive models currently used in practice are on the micro level and go to great lengths to exploit the finer system dynamics. Even so, these models also will (and currently do) use a heat-unit basis for their time scales.

There should be no loss of generality in using only the classical models as exemplifying the properties of the heat-unit basis. Toward this notion, we review the details of two delightful classical population models: the Compartmental and the Lotka-Volterra.

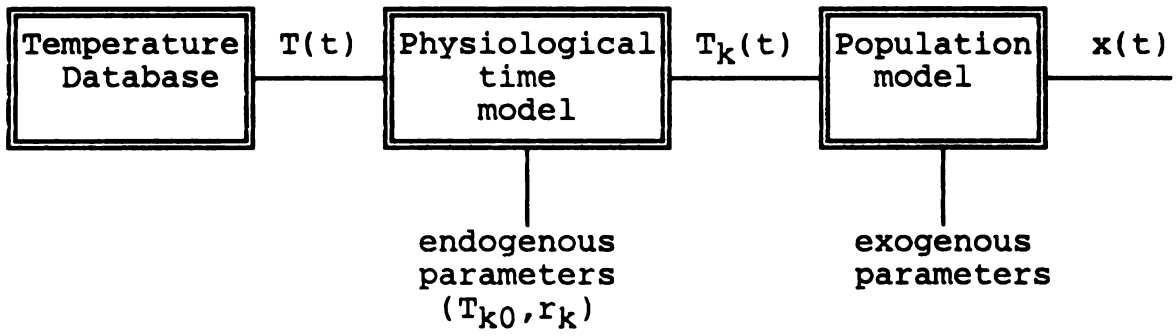


Figure 8.1: Block diagram of the heat-unit population model.

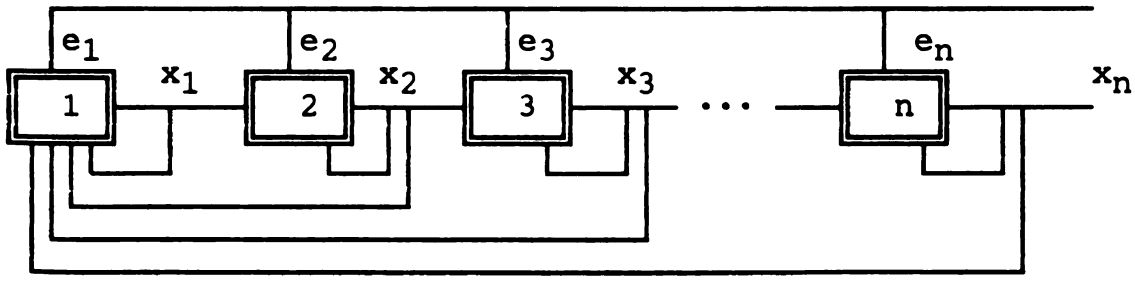


Figure 8.2: Block diagram of the compartmental model.

THE COMPARTMENTAL MODEL

The Logistic, Exponential and Lotka-Volterra models attempt to describe system dynamics on a very course level. It assumes a homogenous species population. This is to say that all individuals behave similarly, young members feed, reproduce and die at the same rates as adult entities. While this may be true in amoebas and other lower levels of life, it most certainly is not the case in more developed organisms.

Nor is it useful to assume these differences will all merge into a single effective species entity. For instance, the very practical problem of modeling a single growing season of a specific agricultural crop is directly contrary to this notion. Since this crop is annual, the growth stage at any time will be essentially uniform throughout the entire population. The entire group will be introduced, germinate, grow and mature in virtual unison. In other words, the indigenous properties and behavior of the species vary significantly with time.

Thus we evolve a hierarchy of population models. At the macro level we assume fixed population properties throughout. The only dynamic is taken to be the inter-species behavior. Therefore, this very course model is referred to as a macro model. Examples include the

Exponential, Logistic and Lotka-Volterra representations. In order to represent the time varying behavior of systems it is useful to break the individual's life history into stages where each stage has its own unique behavioral characteristics. Therefore, we consider these to be micro models.

In its most general form, the compartmental model deals with this by partitioning the species lifetime into n stages. For example: larvae, pupae, adult, etc. as illustrated in Figure 8.2.

The output of each stage includes x_k and e_k which are the size of the live population and number of deaths incurred by stage k. Notice that the population size x_k feeds back to the first (k=1: birth) stage. Each stage produces deaths and new population for the new stage. This may be described mathematically as follows:

$$x_k(t+\delta) = f_k(x_k(t),x_{k-1}(t),t)$$

[8.1] $$x_1(t+\delta) = f_1(x_1(t),\cdots,x_n(t),t), \quad k=2,3,\cdots,n$$

$$e_k(t+\delta) = g_k(x_k(t),t), \qquad k=1,2,3,\cdots,n$$

As defined above, this model has a great deal of generality. In practice the functions f_k and g_k are specified a'priori and the parameters introduced evaluated.

One particular case of the compartmental model is the Leslie model. It is usually based on the following assumptions:

i. Functions f_k and g_k are linear.

ii. Functions f_k depend only on x_{k-1} (and independent of x_k) for $k=2,3,\cdots,n$.

iii. Functions f_k and g_k are explicitly independent of time.

iv. Functions f_k and g_k are often constant.

In this case the only feedback is to the birth stage. Equations [8.1] can now be expressed as follows:

$$x_1(t+\delta) = \sum_{k=2}^{n} a_k(t)\cdot x_k(t)$$

[8.2] $$x_k(t+\delta) = b_k(t)\cdot x_{k-1}(t), \qquad k=2,3,\cdots,n$$

$$e_k(t+\delta) = c_k(t)\cdot x_{k-1}(t), \qquad k=1,2,\cdots,n$$

This is illustrated by the block diagram in Figure 8.3.

If $a_k(t)$, $b_k(t)$ and $c_k(t)$ are known functions, the identification problem can be handled as in chapter seven. The more general problem of handling unknown f_k, g_k for $k=1,2,\cdots,n$ is well beyond the scope of this dissertation. However, the author knows of no work toward this problem and the author feels this would be the genesis of some excellent research.

Consider once again the basic logistic model as defined in chapter six. This model considers the population as a total entity:

$$[8.3] \qquad x(t) = \sum_{k=1}^{n} x_k(t)$$

One can now consider the compartmental model as a subsystem of the larger logistic system where feedback goes only to the birth stage, but other stages act as delays. These delays may be proportional or distributed. In the later case, the dynamics are described as differential equations. Again, this is beyond the scope of the present work.

Only in the case of the simple Leslie model (simple meaning assumption (iv.) is taken) can the parameters of a larger system with a Leslie subsystem be handled as in chapter six.

It should come as no surprise that by introducing the heat-unit hypothesis, a significant complication is added. This too will not be handled here, even though it seems to be an attractive path to take. Rather, let us pursue another amalgamated species macro model: the Lotka-Volterra.

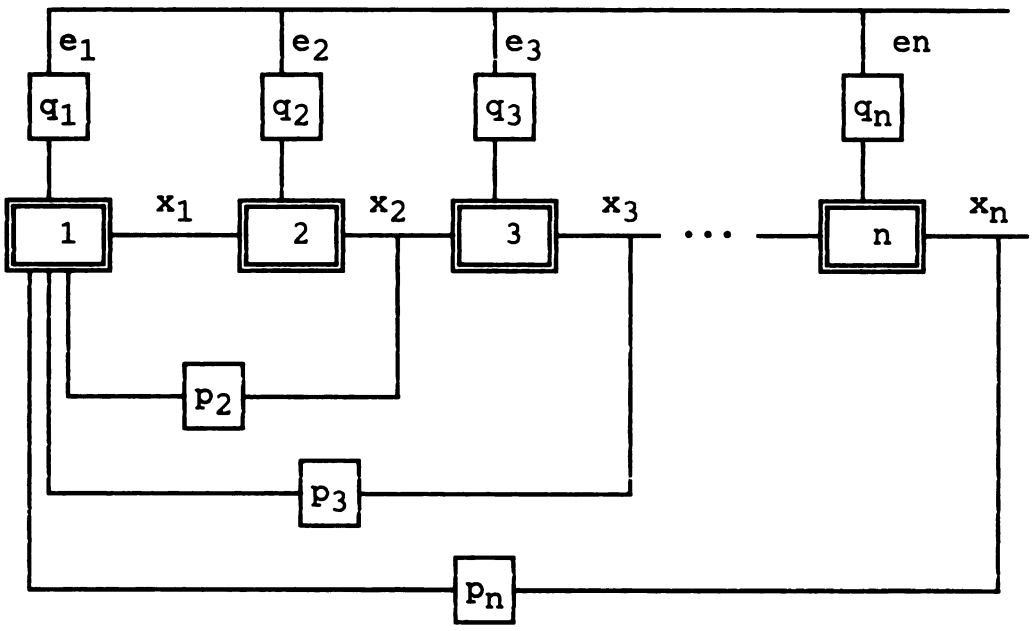


Figure 8.3: Block diagram of the Leslie model.

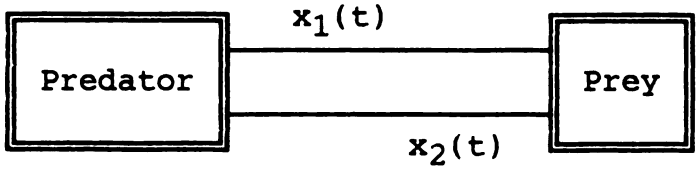


Figure 8.4: Block diagram of the predator/prey interaction model.

THE LOTKA-VOLTERRA MODEL

Thus far we have considered only single species populations. By introducing a second species there are three possibilities:

i. The species interact as predator and prey.

ii. The species do not interact, but do compete for the same food source.

iii. The species neither interact nor compete.

Lotka-Volterra addresses case (i.). It is illustrated in Figure 8.4 where each block is represented mathematically as:

$$\frac{d}{dt} x_1(t) = r_1 \cdot x_1(t) \cdot \left[1 - \frac{x_2(t)}{x_{2m}} \right]$$

[8.4]

$$\frac{d}{dt} x_2(t) = r_2 \cdot x_2(t) \cdot \left[1 - \frac{x_1(t)}{x_{1m}} \right]$$

where r_1, r_2 are (endogenous) growth parameters and x_{1m}, x_{2m} are exogenous system parameters.

Equations [8.4] represent an obviously nonlinear system. Indeed, Lotka-Volterra is a classical example showing all the limit cycles and asymptotic stability features normally associated with such systems. Using the results of Theorem 4.2 in chapter four, it is easily converted to a heat-unit based system.

If $\tau_1(t)$ and $\tau_2(t)$ are the physiological time bases of x_1 and x_2 respectively, equation [8.4] reduces to the following system:

$$\frac{d}{d\tau_1} x_1(\tau_1) = r_1 \cdot x_1(\tau_1) \cdot \left[1 - \frac{x_2(\tau_2)}{x_{2m}} \right]$$

[8.5]

$$\frac{d}{d\tau_2} x_2(\tau_2) = -r_2 \cdot x_2(\tau_2) \cdot \left[1 - \frac{x_1(\tau_1)}{x_{1m}} \right]$$

In practice, each of the system blocks in Figure 8.4 can be subdivided into compartmental subsystems. However, we consider only the courser model here in which each population is considered to be homogenous.

Using Eulerian integration with $\delta=1$ day we can numerically integrate in a straight-forward manner:

$$x_1(\tau_1(t+1)) = \left[1 + r_1 \cdot [\tau_1(t+1) - \tau_1(t)] \cdot \left[1 - \frac{x_2(\tau_2(t))}{x_{2m}} \right] \right] \cdot x_1(\tau_1(t))$$

[8.6]

$$x_2(\tau_2(t+1)) = \left[1 + r_1 \cdot [\tau_2(t+1) - \tau_2(t)] \cdot \left[1 - \frac{x_1(\tau_1(t))}{x_{1m}} \right] \right] \cdot x_2(\tau_2(t))$$

Subject to...

$$x_1(\tau_1(t_{10})) = x_{10}$$

[8.7]

$$x_2(\tau_2(t_{20})) = x_{20}$$

Such systems can be solved and analyzed using this model applying all the techniques illustrated in earlier chapters. However, it is the goal of this concluding chapter to use Lotka-Volterra in order to illustrate the application of control systems on heat-based biological models.

Applying system controls to mathematical models is not new. However, in order to stimulate a systematic development along more mathematical lines, several control problems will be defined and solved using numerical methods.

CONTROL PROBLEMS

We proceed to describe a typical biological control application. The model will be intentionally simple: Lotka-Volterra. This has been done intentionally so that the reader can easily extend this to any mathematical model, most likely of the compartmental type.

Consider an agricultural system in which there exist a single predator (population x_2) and a prey (a crop: population x_1). Again, although this is an intentionally simple system, the extensions are straight forward. Each species has its own physiological time base and it is desired to maximize crop yield.

Actually, the objective can be made much more realistic. If there are pesticides, herbicides (in general: Xicides) or fertilizers applied, the amount of each will need to be minimized as well as the number of applications for economic considerations. Since the "price" associated with any of these items is usually given a "dollars per unit", this amounts to a linear objective function.

We consider the system revenue as the objective which is to be maximized. Assuming p_0 to be the price received per unit crop at harvest time, the revenue (Rev) is given by:

[8.8] $$\text{Rev} = p_0 \cdot x_1(t_f), \qquad p_0>0$$

where t_f is the time of harvest for the crop x_1.

The effect of a pesticide or herbicide is that of killing a number of the prey species. In practice, it is reasonable to assume that the more effective the Xicide, the greater the population that will be lost. Thus we assume that The effectiveness is inversely proportional to the fraction of individuals remaining at the end of the Xicide application. We also assume the cost to be directly proportional to the effectiveness. Thus, for each application of the Xicide, the species population and cost are given by:

[8.9]
$$x_2(t+\delta) = f_2 \cdot x_2(t), \qquad 0<f_2<1$$
$$\text{Rev} = -c_2/f_2, \qquad c_2>0$$

where f_2 is the fraction of individuals left after an application and c_2 is the cost per unit applied.

Similarly, there is a cost associated with each Xicide application. Assuming each application costs the same,

[8.10] $$\text{Rev} = -c_3 \cdot n_2, \qquad c_3>0$$

where c_3 is the cost per application and n_2 is the number of applications.

Applications of fertilizer have the opposite effect. In principle this will assist only the prey population, although in practice some predators will likely benefit too. Again, the cost is taken to be inversely proportional to the fertilizer effectiveness. This is summarized below:

$$x_1(t+\delta) = f_1 \cdot x_1(t), \qquad f_1>1$$

[8.11]

$$\text{Rev} = -c_1/f_1, \qquad c_1>0$$

Also, the cost per application is given by

[8.12] $$\text{Rev} = -c_4 \cdot n_1, \qquad c_4>0$$

where c_4 is the cost per application.

For this simplified model the total revenue is given by the combination of equations [8.9], [8.10],[8.11] and [8.12].

[8.13] $$\text{Rev} = p_0 \cdot x(t_f) - c_1/f_1 - c_2/f_2 - c_3 \cdot n_1 - c_4 \cdot n_2$$

where for each Xicide application,

$$x_2(t+\delta) = f_2 \cdot x_2(t)$$

[8.14]

$$n_2 = n_2 + 1$$

and each fertilizer application,

$$x_1(t+\delta) = f_1 \cdot x_1(t)$$

[8.15]

$$n_1 = n_1 + 1$$

It is required that this objective function be maximized. The variables that will effect this are the Xicide and fertilizer proportions, the protocols for their application and the population dynamics.

Biological controls may be thought of as being either active or passive. Active controls are those in which an external agent is introduced in order to change the values of the state variables outside the parameters of the population dynamic. The most obvious of these are chemicals and include both Xicides (effecting the pest) and fertilizers (effecting the crop). The effect of a chemical introduction is usually thought of as fractionating and is represented by equations [8.9] and [8.11] above. They have no effect on the system dynamic, but only upon the system state at the moment of application.

It is also possible to introduce a new biological species to the system. The effect of this active control is to alter the population dynamic itself. Needless to say, this will often have a long range effect and history is replead with experiments in which control was lost!

Active control systems can be either based on open loop or closed loop (feedback) principles. In the case of heat-unit based systems, open loop is not clearly inferior as in many physical systems. Since temperature histories are so regular, controls applied by the calendar have traditionally been used and found effective. However, other important entities, such as precipitation, give rise to a real need for feedback.

Passive controls are also a possibility. These controls attempt to modulate system parameters to the advantage of the crop. For instance, by varying the time of crop introduction (t_{10}) one can strike a balance between frozen seedlings (too early) and a well developed and hungry prey (too late). Recently the developing technology of genetic alteration has led to an adaptive potential to alter the actual exogenous parameters of the species too.

We now consider several specific control problems with which the above principles can be applied.

THE CROP INTRODUCTION TIME PROBLEM

Consider a passive control problem in a simple single predator - single prey Lotka-Volterra model. As the growing season develops, the predator population (in the form of animal pests or competing vegetation) will become increasingly robust with time and a newly introduced crop will be hard pressed to hold its own against a well established foe. In order to counter this, an early introduction of the crop would be especially helpful. By introducing the crop early, the number of degree-days experienced will be maximal and the harvest will also be maximized.

Early introduction of the crop has its hazards too. Spring frosts will produce a killing effect on germinating plants. Unfortunately, this basic theory has not allowed for this possibility! The heat-unit hypothesis as was put forward in preceding chapters allows only for dormancy under the temperature threshold. This modeling problem is easily overcome by defining an appropriate biological activity function as below.

Let the biological activity be defined using two temperature thresholds:

T_0: beneath which the crop population is fractionated with rate r_0 depending on the threshold excess,

T_1: the threshold of dormancy.

This activity is mathematically defined as follows:

$$[8.16] \qquad \beta(T) = \begin{cases} r_0 \cdot [T(t)-T_0], & T(t)<T_0 \\ 0, & T_0 \le T(t) \le T_1 \\ r_1 \cdot [T(t)-T_1], & T(t)>T_1 \end{cases}$$

Since with Type II and IV systems it is not possible to find a single explicit formula for the physiological time. This is so because the activity function cannot be decomposed into factors, one of which is strictly time dependent. Even so, this problem is still tractable.

The decomposition proceeds as in chapter six:

$$[8.17] \qquad \int_{\Gamma(t)} \beta(T(\mu))d\mu = r_0 \cdot \tau_0(t) + r_1 \cdot \tau_1(t)$$

$$[8.18] \qquad \text{where} \quad \tau_0(t) = \int_{\Gamma_0(t)} [T(\mu)-T_0]\, d\mu$$

$$\tau_1(t) = \int_{\Gamma_1(t)} [T(\mu)-T_0]\, d\mu$$

$$\Gamma_0(t) = \{\ t>t_0 : T(t)<T_0\ \}$$

[8.19]

$$\Gamma_0(t) = \{\ t>t_0 : T(t)<T_0\ \}$$

By the results of chapter four, this integral can be accurately integrated by using the numerical Eulerian approximation with $\delta=1$ day. The evaluation of this integral is presented as Algorithm 8.1, which allows us to solve and ultimately analyze the crop introduction time problem discussed above.

Let us be more precise. Suppose the prey population at time t is represented by $x_1(t)$ and the its biological activity be represented as in the following equation:

[8.20]
$$\beta_1(T) = \begin{cases} r_0 \cdot [T(t)-T_{10}], & T(t)<T_{10} \\ 0, & T_{10}\le T(t)\le T_{11} \\ r_1 \cdot [T(t)-T_{11}], & T_{11}\le T(t)\le T_{12} \\ 0, & T(t)\ge T_{10} \end{cases}$$

Also suppose the prey has biological activity based on a Type II model as below:

$$[8.21] \qquad \beta_2(t) = \begin{cases} 0, & T(t) \leq T_{20} \\ r_2 \cdot [T(t) - T_{20}], & T_{20} < T(t) < T_{21} \\ 0, & T(t) \geq T_{21} \end{cases}$$

Graphs of each of these activity functions are shown in Figures 8.6 and 8.7.

By the procedures outlined in chapter seven, it is possible to determine the endogenous parameters r_0, r_1, r_2 and T_{10}, T_{11}, T_{12} and T_{20}, T_{21} if a season's history is available. However, if this is not the case, it is still desirable to make a qualitative determination of any possible results. Toward this end we will look at illustrative examples, put forward tentative conjectures and plan for further work using these control methods.

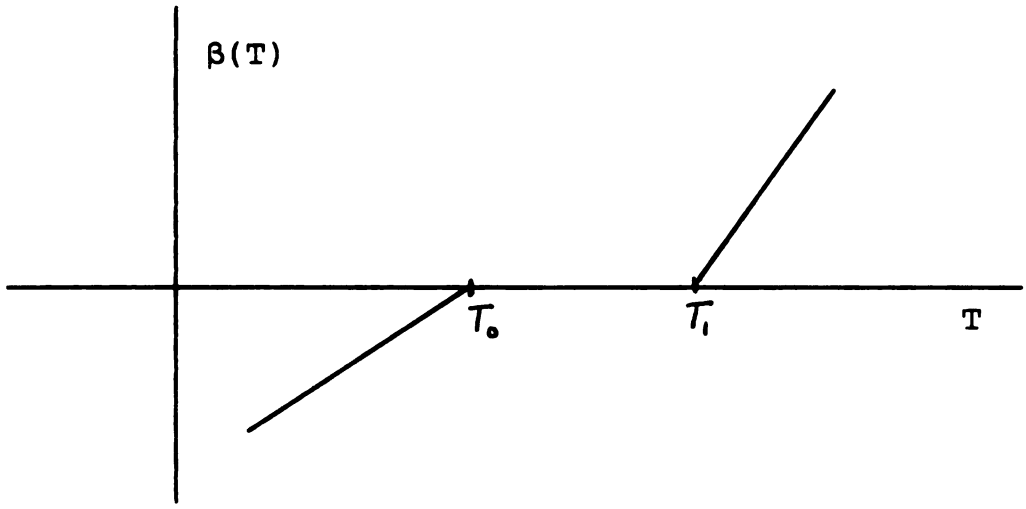


Figure 8.5: The biological activity function defined by equation [8.16].

```
DO t=1,t0   { τ(t) = 0

DO t=t0+1,p { T(t)<T0 { τ(t) = τ(t-1)+ r0·[T(t)-T0]
            { T(t)≥T1 { τ(t) = τ(t-1)+ r1·[T(t)-T1]
```

Algorithm 8.1: Computing the physiological time for the multiple growth rate system of equation [8.16].

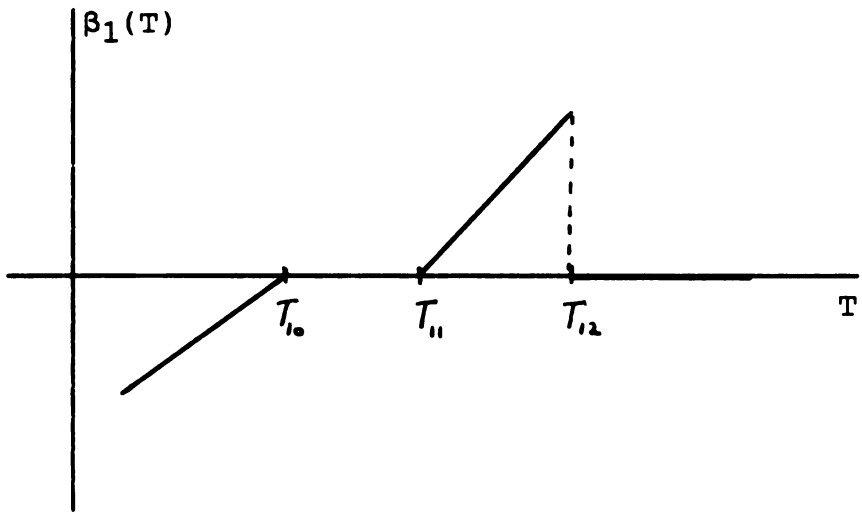


Figure 8.6: Graph of the Biological Activity $\beta_1(T)$ of equation [8.20].

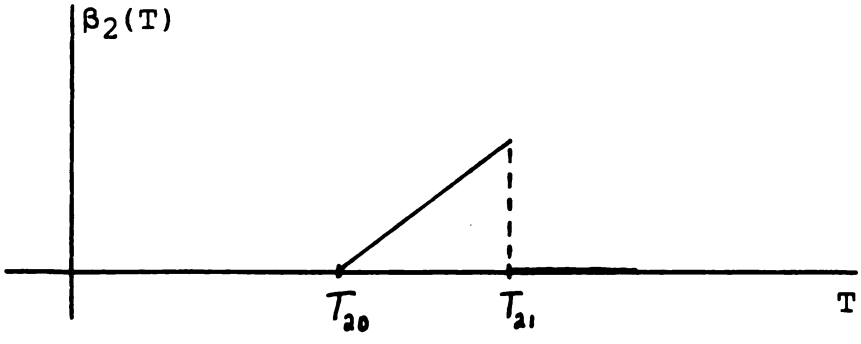


Figure 8.7: Graph of the Biological Activity $\beta_2(T)$ of equation [8.21].

A CONTROL EXAMPLE

We consider here the Lotka-Volterra model described in the previous section. We assume the growth rates r_0, r_1, r_2, the predator temperature thresholds T_{20}, T_{21} and upper dormant threshold for the prey, T_{12}, are known a'priori. This is summarized along with values in Table 8.1.

It is desired to find the optimal time for introducing the prey (crop) population into the system. In order to do this we will consider two situations:

i. The predator population is controlled (by a feedback fractionating method to be described later).

ii. The predator population is uncontrolled.

In either case we wish to observe how the results vary with threshold temperatures of the crop. This will be done using the theoretic distribution preserving generator described in chapter seven.

As an initial simulation the deterministic Lotka-Volterra model was implemented and the population metrics of both

predator and prey are plotted over the year's time. This is illustrated in Figures 8.8 and 8.9 using $T_{10}=15^{o}$, $T_{11}{}^{o}$ F and $T_{10}=25^{o}$, $T_{11}=40^{o}$ F respectively. It is clearly shown that the responses to each are different, but the point of interest is really relative population metrics at the time horizon (harvest). We choose this, also somewhat arbitrarily, at $t_f=270$ days.

In order to determine an optimal initial time t_{10} (planting time) it would be most useful to vary this parameter and check the corresponding harvest. Thus, what we wish is a plot of $x_1(t_f)$ against t_{10}. Since it is desired to check the results against both controlled and uncontrolled predator models, it would also be helpful to plot $x_2(t_f)$ against t_{10}. This is done and the results of this simulation are shown in Figures 8.10, 8.11 and 8.12 showing average uncontrolled results. Figures 8.21, 8.22 and 8.23 show average controlled predator results.

The curves are rather smooth and an absolute maximum does exist. It would be of minimal effort to implement a line search algorithm such as the golden section method implemented in chapter seven. This would determine the optimal T_{10} for any set of parameters. Certainly this should (and was) done. However tempting it would be to conclude the analysis here, further work can give us more.

Rather that a **optimal point,** each figure seems to indicate an **optimal interval.** Figures 8.10a, 8.10b and 8.10c (with the death threshold at 15^{o}F) are optimal over roughly the same interval (approximately 35 to 100 days) while those in Figures 8.11 are similar, and also these in Figures 8.12. Indeed, the major difference between these figures is simply the magnitude of harvest population. Interestingly, the more tolerant species (in a lower death threshold sense) are more productive, even with a relatively late introduction time.

To better understand this relationship, the comparative prey populations are plotted in Figures 8.13. We note that the magnitudes are clearly nested for early introduction times and tend to almost identical harvest figures for later introduction times. The point at which these become essentially equal is later for the lowest death threshold.

This relationship is to be contrasted with that found in Figures 8.14 where the families have been grouped by dormancy threshold. The interval of optimality is no longer the same throughout the family and it is observed that a more clear-cut optimal point is achieved for the lower death threshold species. Each curve starts and finishes essentially the same. Only for moderate starting times do they significantly differ.

Since this simulation was performed using a stochastic driver, it is most important that we also study variation in harvest. Graphs of the standard deviation in both the predator and prey populations, corresponding to the means shown in Figures 8.10, 8.11 and 8.12, are given in Figures 8.19, 8.20 and 8.21. It should be noticed that in each case the standard deviations of both predator and prey are essentially the same and moderate, but as the initial time becomes later, the prey harvest population can vary significantly (≈ 0.7 standard units). However, the prey remains or becomes moderate in each case.

As the death threshold increases the variance of (both) populations becomes ever more significant. For predator populations this means that the standard deviation curves become more "U" shaped and only for moderate ranged introduction times are the variances small enough to give relatively deterministic results in the man harvest predictor. For prey (crops) this will be true for moderate and late introduction times.

In summary, it is evident that there is structure in both the mean and standard deviation. Any useful simulation must look at both in order to be a useful predictor.

Table 8.1 Values of parameters in the introduction time control problem.

Endogenous Parameters	Exogenous Parameters	Initial Conditions
$r_0 = 0.01$ $r_1 = 0.0005$ $r_2 = 0.0003$ $T_{12} = 80$ $T_{20} = 45$ $T_{21} = 75$	$x_{1m} = 10$ $x_{2m} = 15$	$x_{10} = 5$ $x_{20} = 10$ $t_{20} = 100$

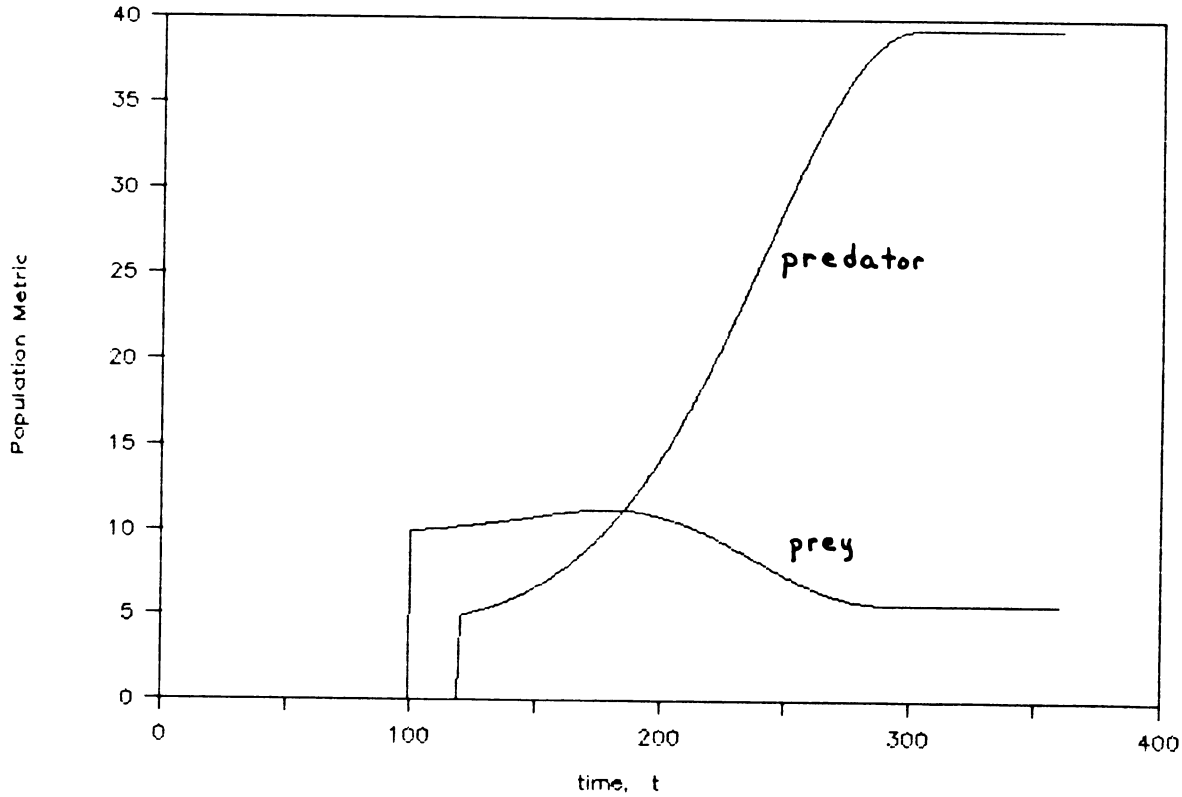


Figure 8.8: Population metric over time with uncontrolled predator, uncontrolled prey. $T_{10}=15^{o}$, $T_{11}=40^{o}$ F.

▪ prey

\+ predator

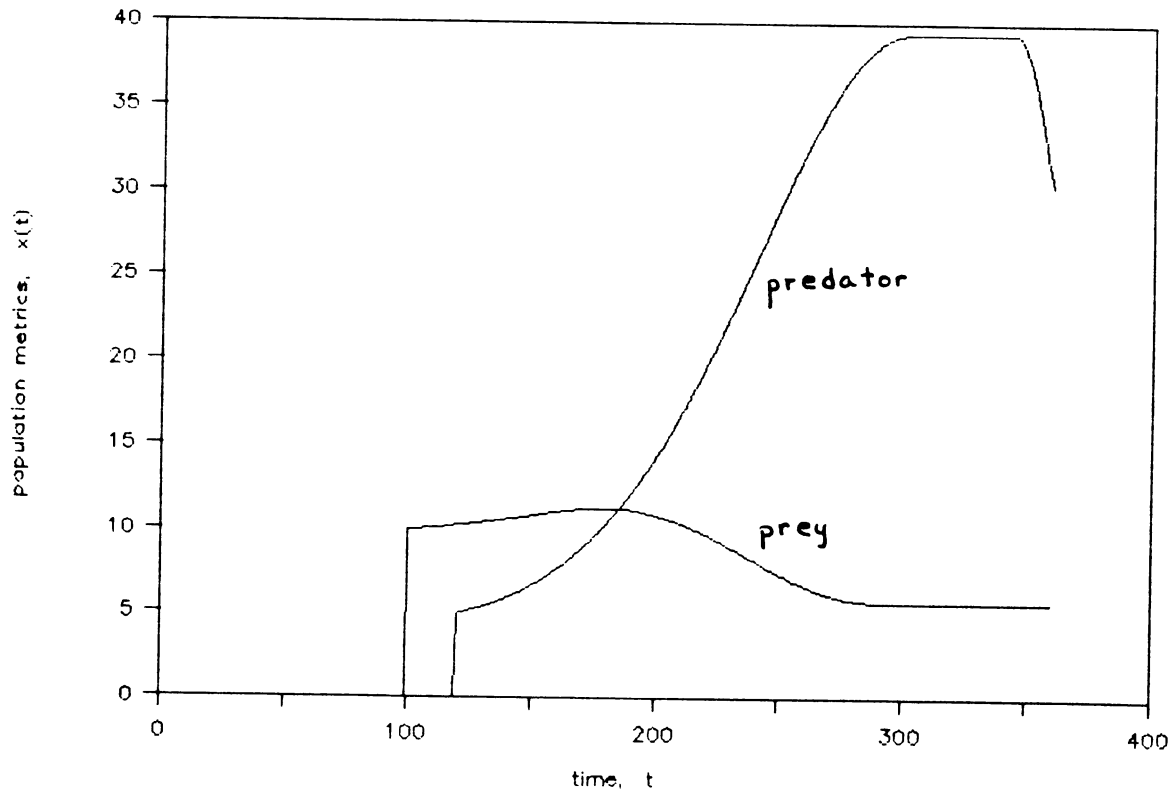


Figure 8.9: Population metric over time with uncontrolled predator, uncontrolled prey. $T_{10}=25^{o}$, $T_{11}=40^{o}$ F.

- • prey
- + predator

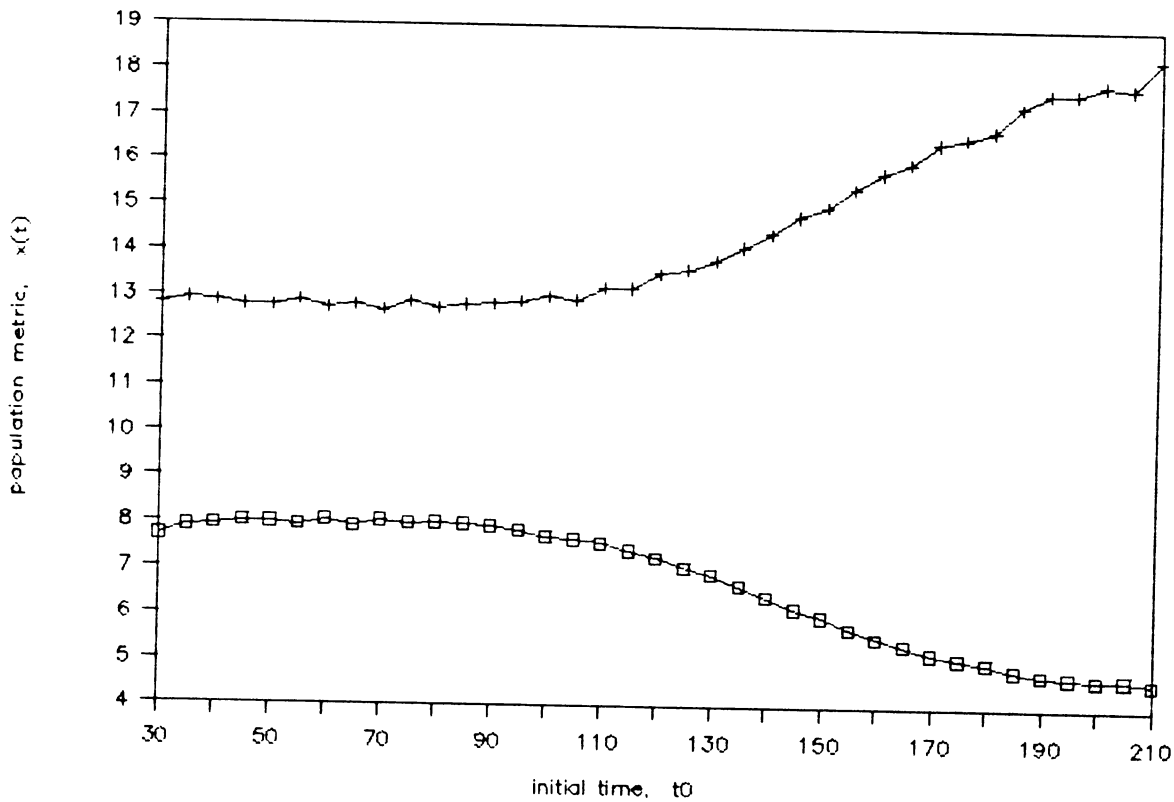


Figure 8.10a Harvest predator/prey mean population metrics as a function of introduction time with an uncontrolled predator. $T_0=15^o$, $T_1=40^o$ F.

▪ prey

\+ predator

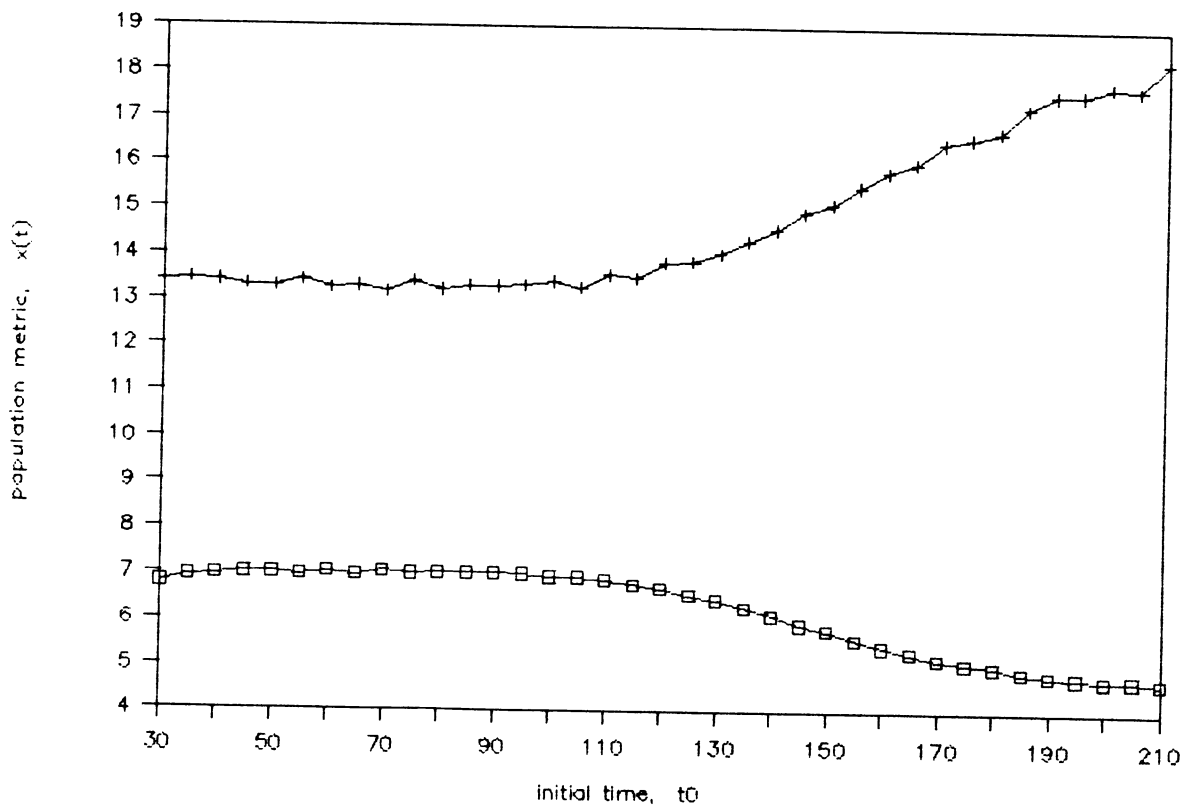


Figure 8.10b Harvest predator/prey mean population metrics as a function of introduction time with an uncontrolled predator. $T_0=15^{o}$, $T_1=45^{o}$ F.

• prey

\+ predator

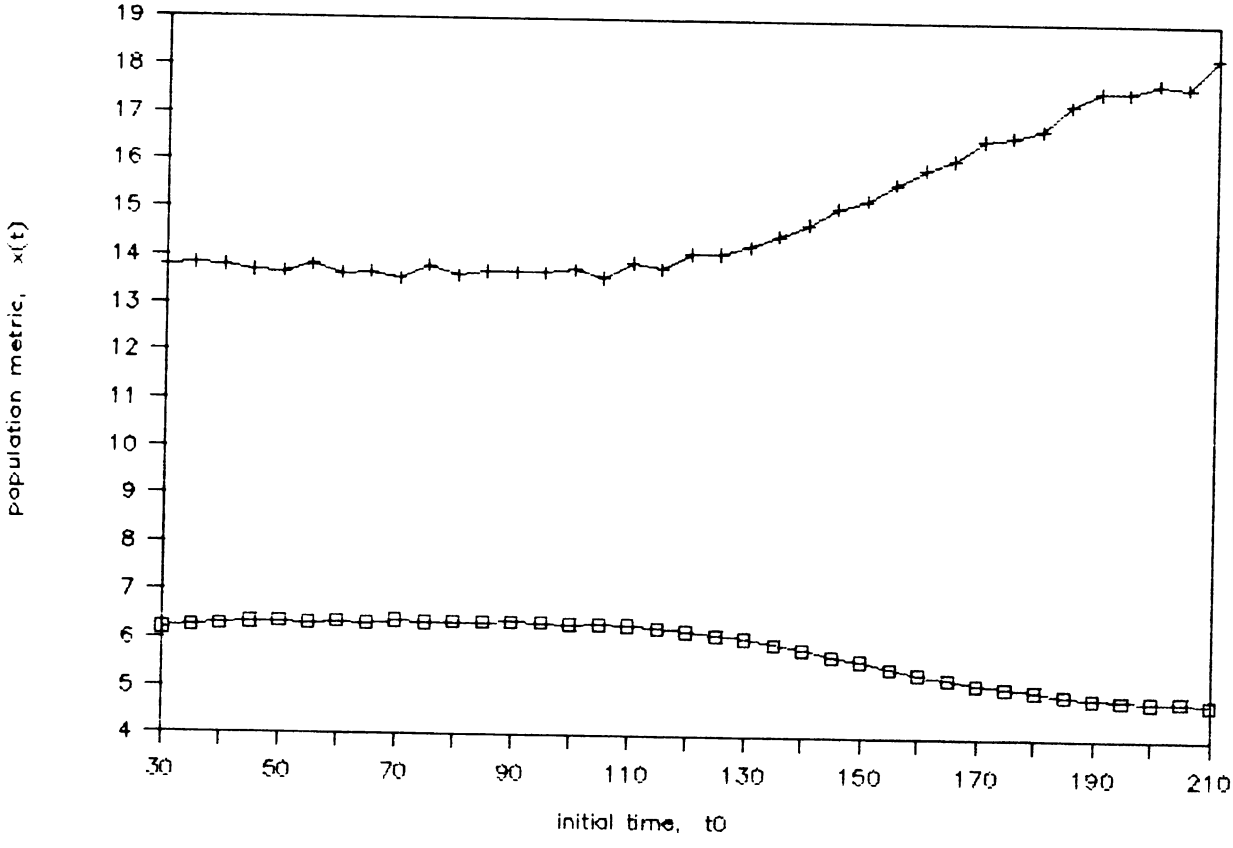


Figure 8.10c Harvest predator/prey mean population metrics as a function of introduction time with an uncontrolled predator. $T_0=15^{o}$, $T_1=50^{o}$ F.

* prey

\+ predator

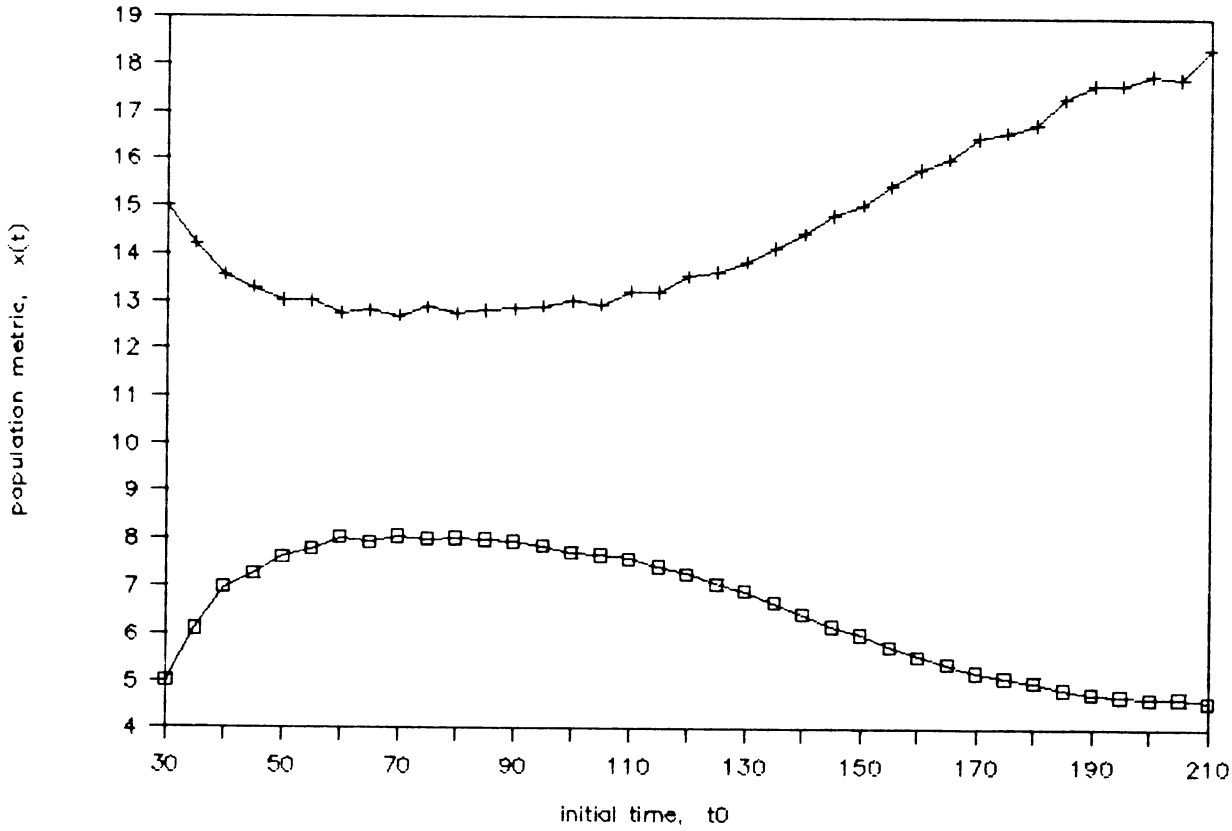


Figure 8.11a Harvest predator/prey mean population metrics as a function of introduction time with an uncontrolled predator. $T_0=20^o$, $T_1=40^o$ F.

▪ prey

\+ predator

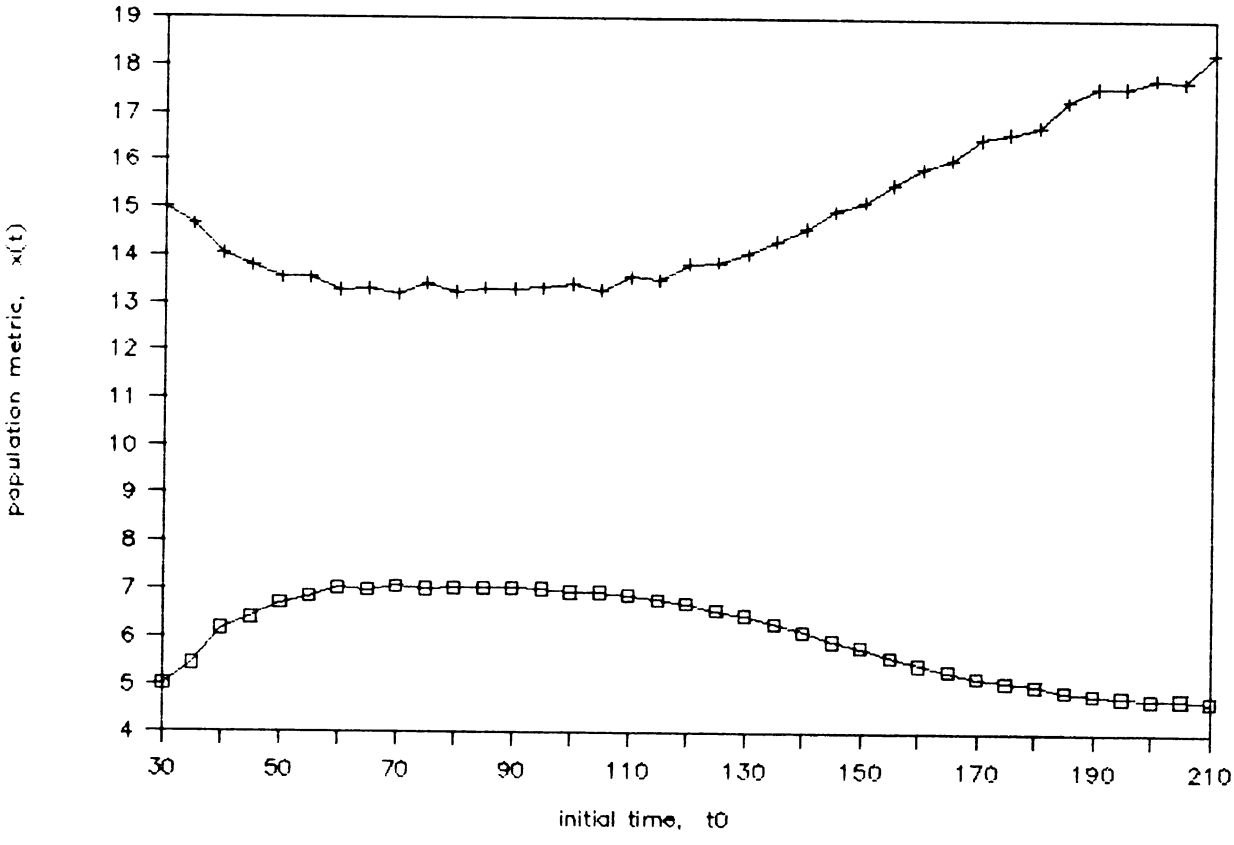


Figure 8.11b: Harvest predator/prey mean population metrics as a function of introduction time with an uncontrolled predator. $T_0=20^{\circ}$, $T_1=45^{\circ}$ F.

• prey

\+ predator

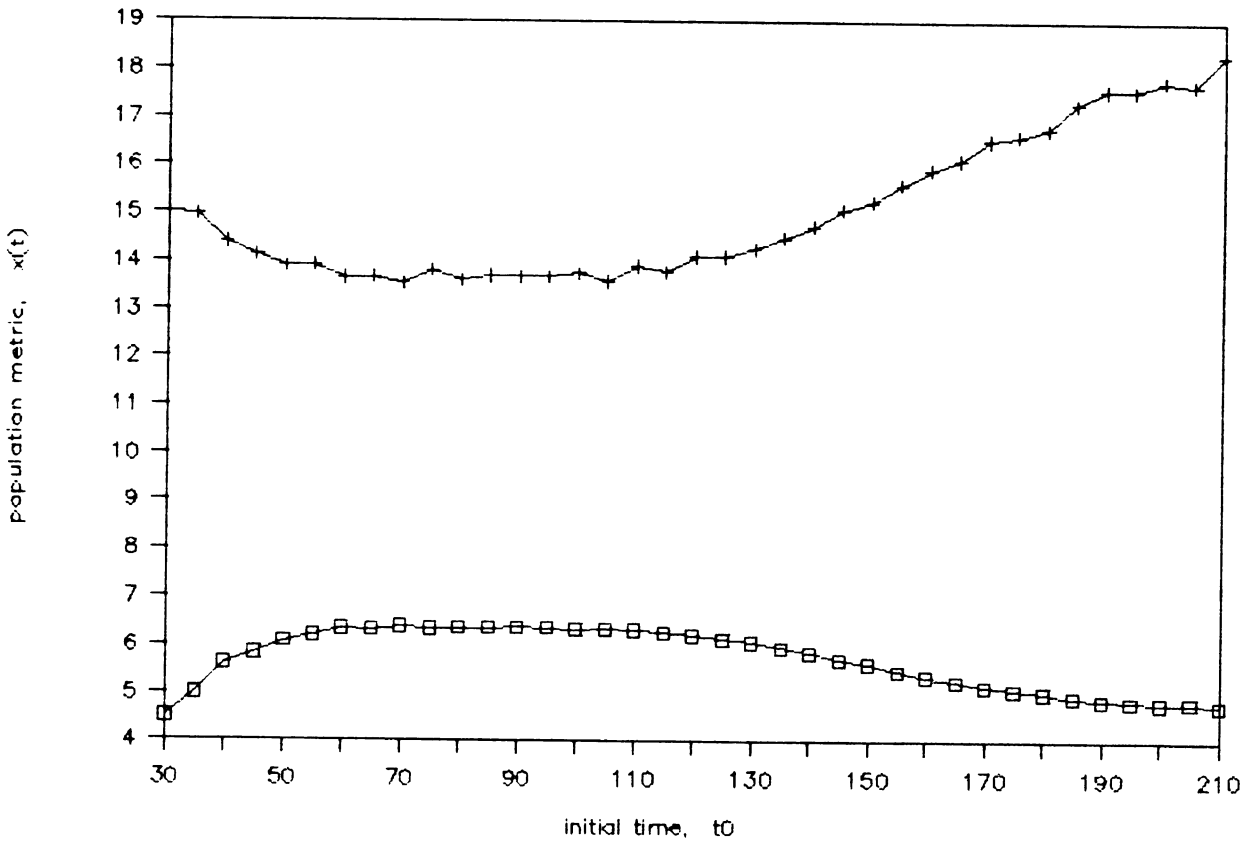


Figure 8.11c Harvest predator/prey mean population metrics as a function of introduction time with an uncontrolled predator. $T_0=20^o$, $T_1=50^o$ F.

▪ prey

\+ predator

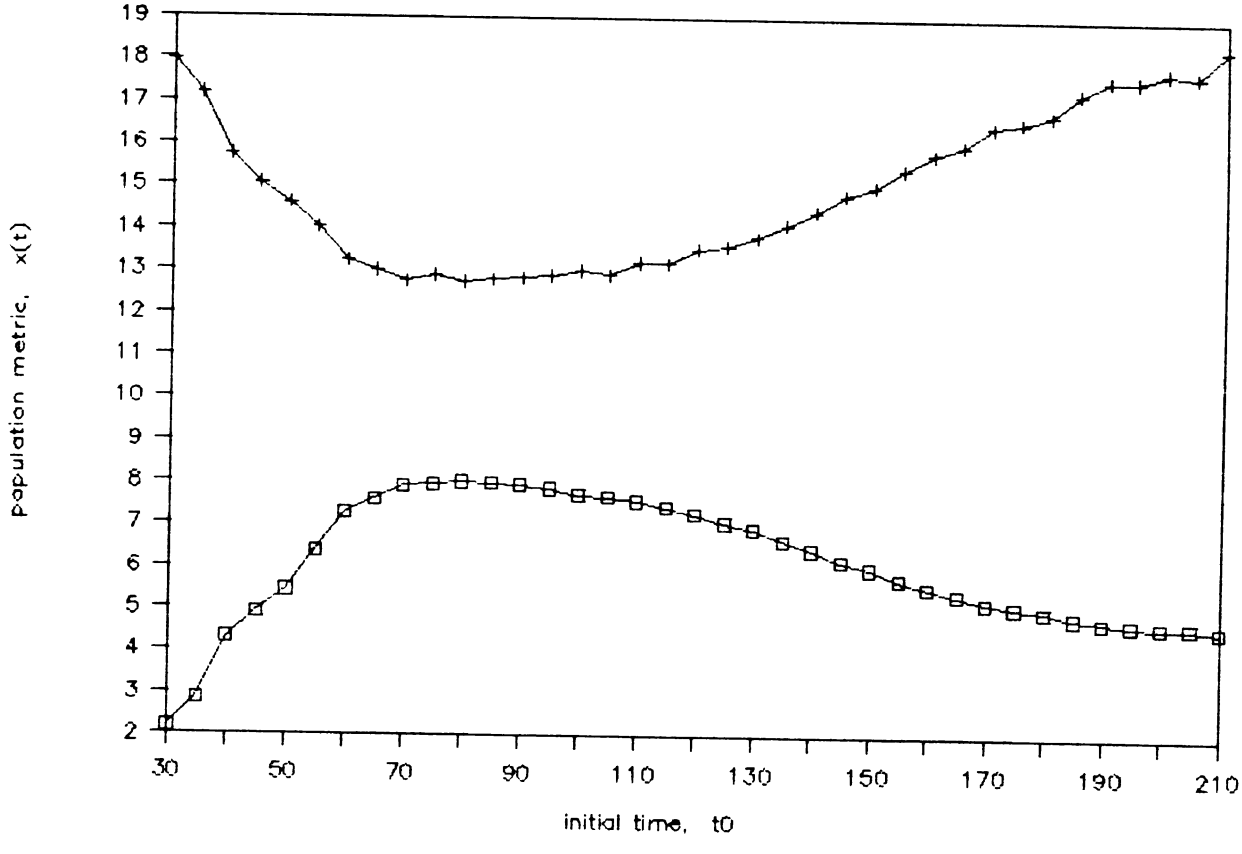


Figure 8.12a Harvest predator/prey mean population metrics as a function of introduction time with an uncontrolled predator. $T_0=25^{\circ}$, $T_1=40^{\circ}$ F.

▪ prey

\+ predator

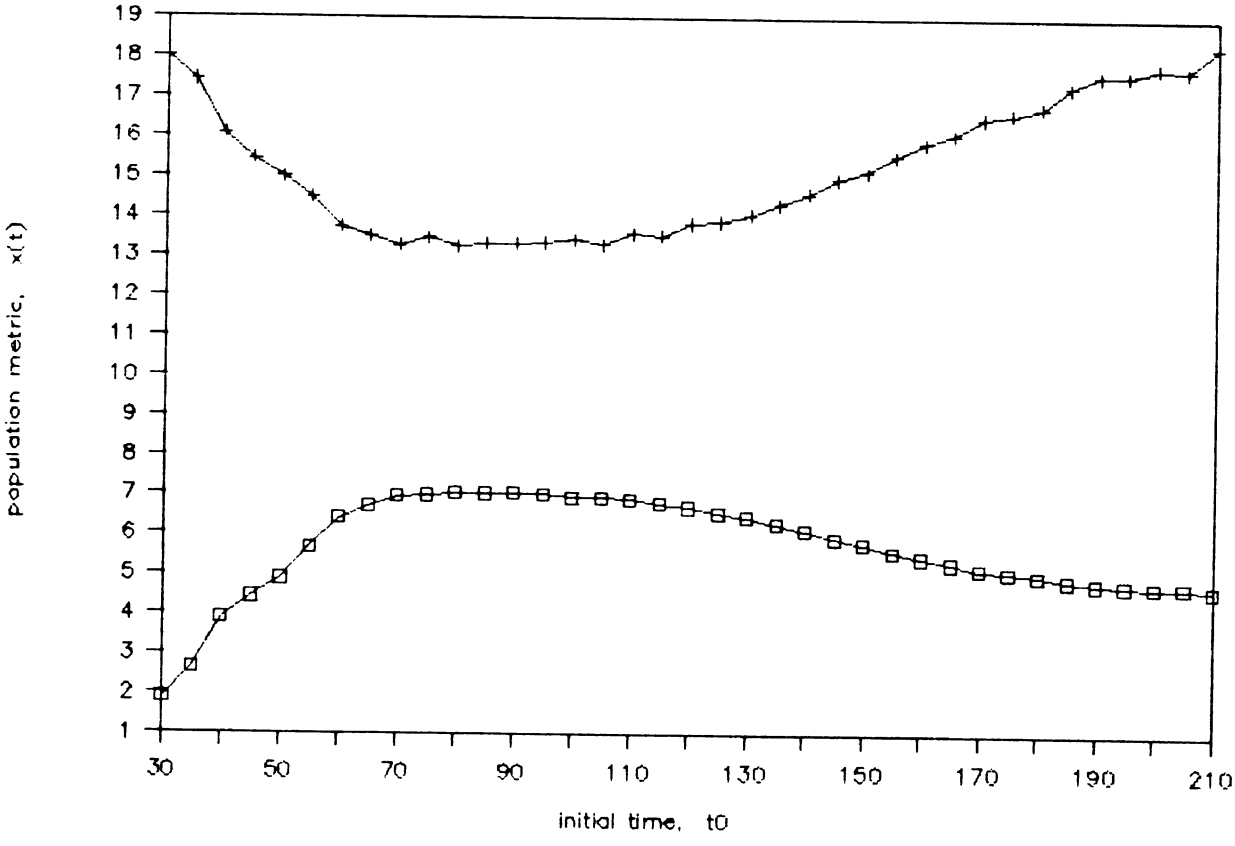


Figure 8.12b Harvest predator/prey mean population metrics as a function of introduction time with an uncontrolled predator. $T_0=25^o$, $T_1=45^o$ F.

• prey

+ predator

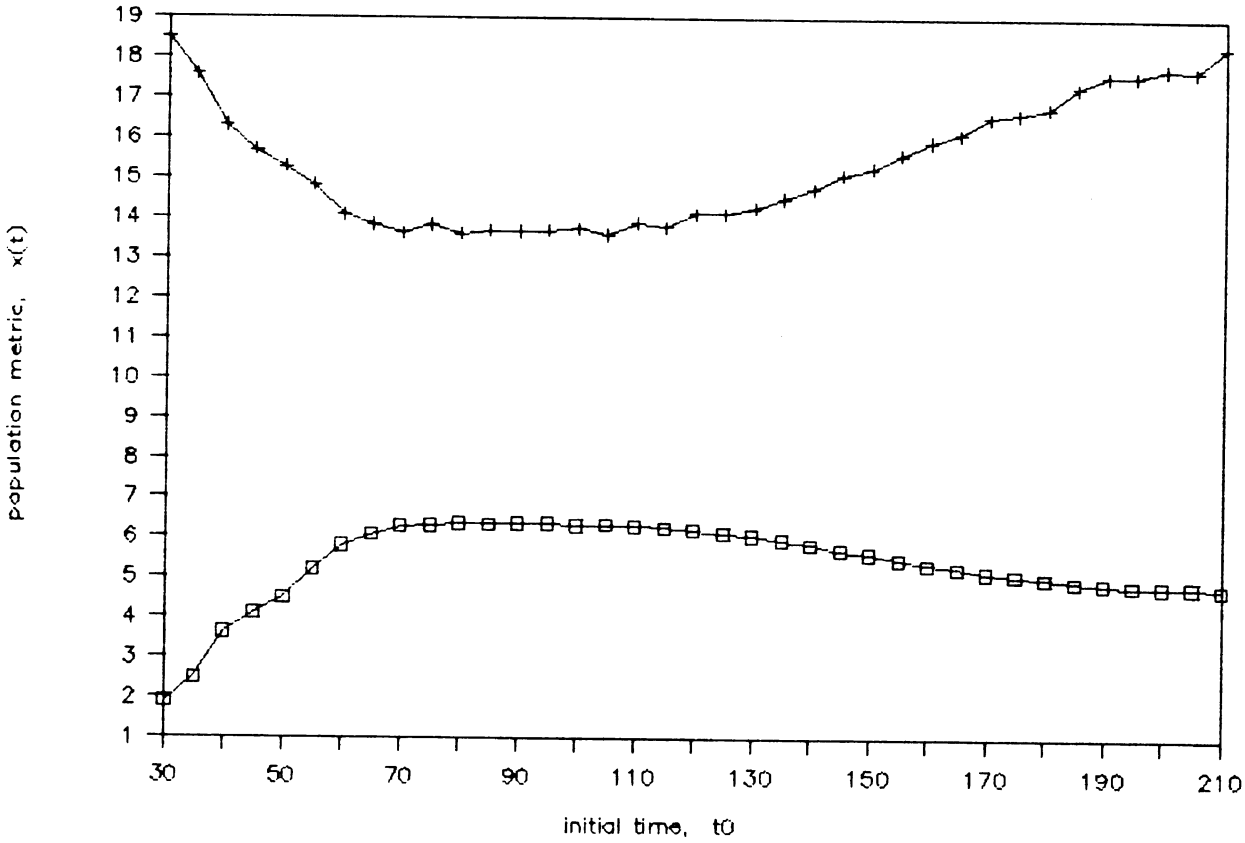


Figure 8.12c Harvest predator/prey mean population metrics as a function of introduction time with an uncontrolled predator. $T_0=25^{o}$, $T_1=50^{o}$ F.

▪ prey

\+ predator

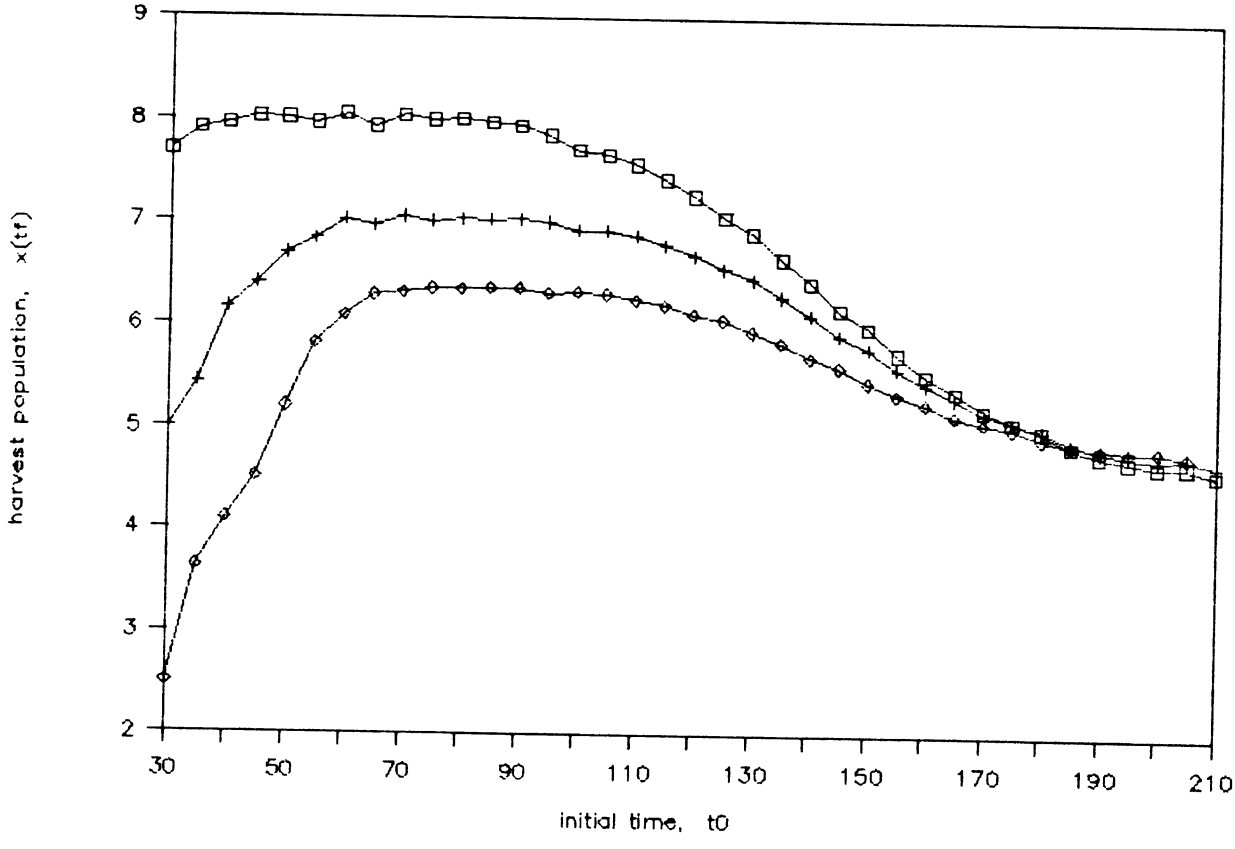


Figure 8.13a Comparison predator/prey mean population metrics summarizing Figures 8.10. $T_{10}=15^{o}$ F.

- ▪ $T_{11}=40^{o}$ F
- \+ $T_{11}=45^{o}$ F
- ♦ $T_{11}=50^{o}$ F

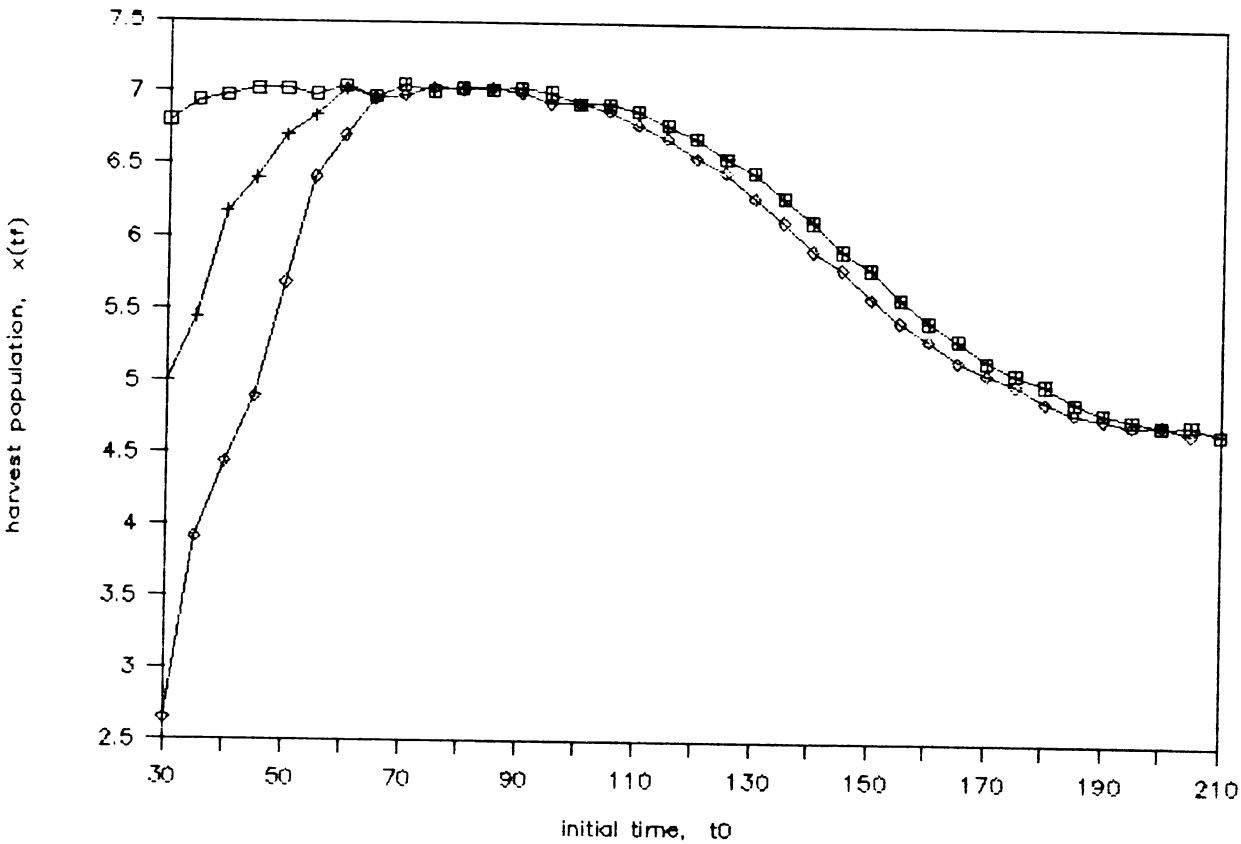


Figure 8.13b Comparison predator/prey mean population metrics summarizing Figures 8.11. $T_{10}=20^{o}$ F.

- ▪ $T_{11}=40^{o}$ F
- \+ $T_{11}=45^{o}$ F
- ♦ $T_{11}=50^{o}$ F

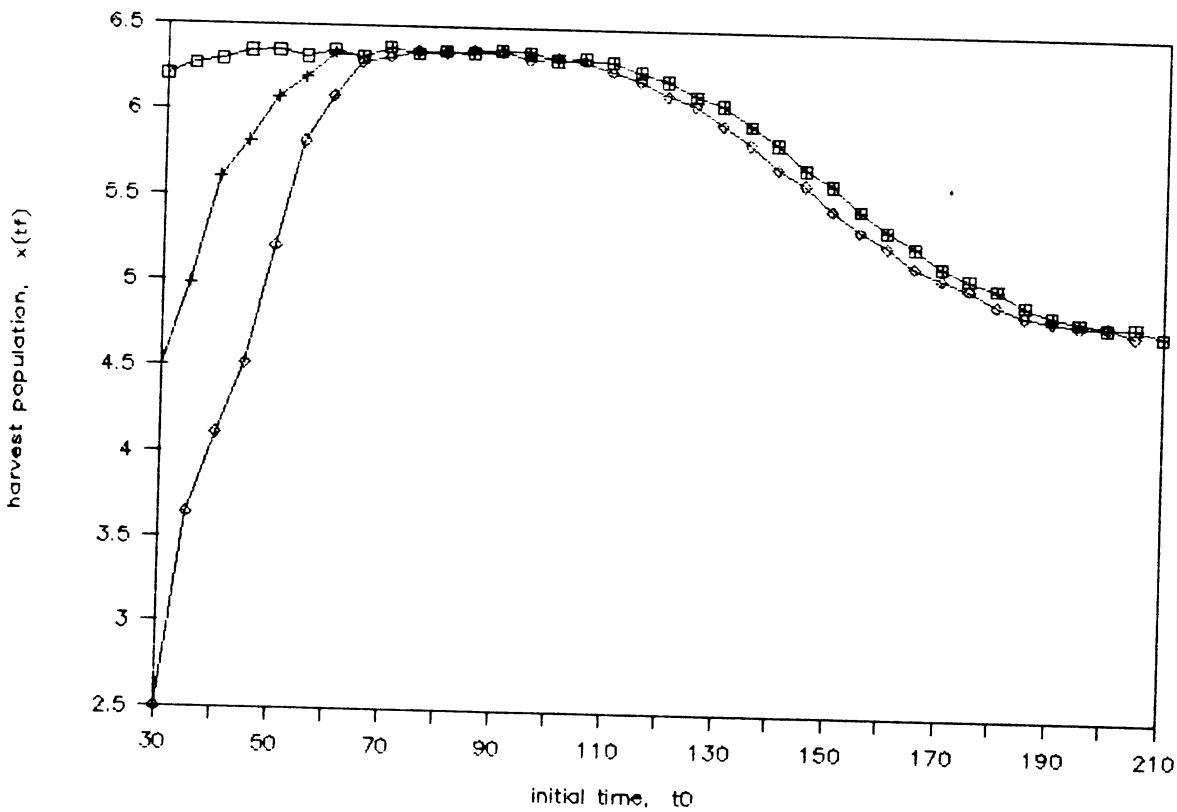


Figure 8.13c Comparison predator/prey mean population metrics summarizing Figures 8.12. $T_{10}=25^{o}$ F.

- ▪ $T_{11}=40^{o}$ F
- \+ $T_{11}=45^{o}$ F
- ♦ $T_{11}=50^{o}$ F

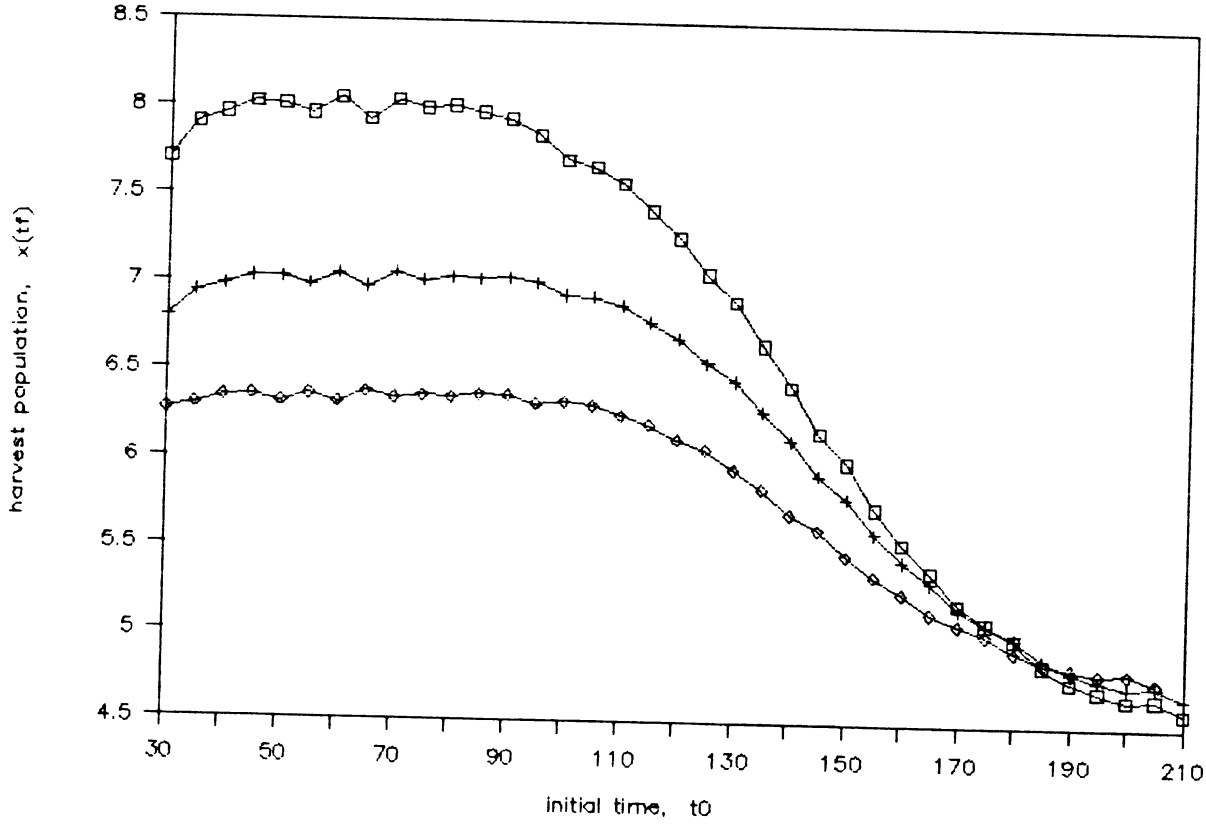


Figure 8.14a Comparison predator/prey mean population metrics summarizing Figures 8.10a, 8.11a, 8.12a. $T_{11}=40^{o}$ F.

- ▪ $T_{10}=15^{o}$ F
- \+ $T_{10}=20^{o}$ F
- ♦ $T_{10}=25^{o}$ F

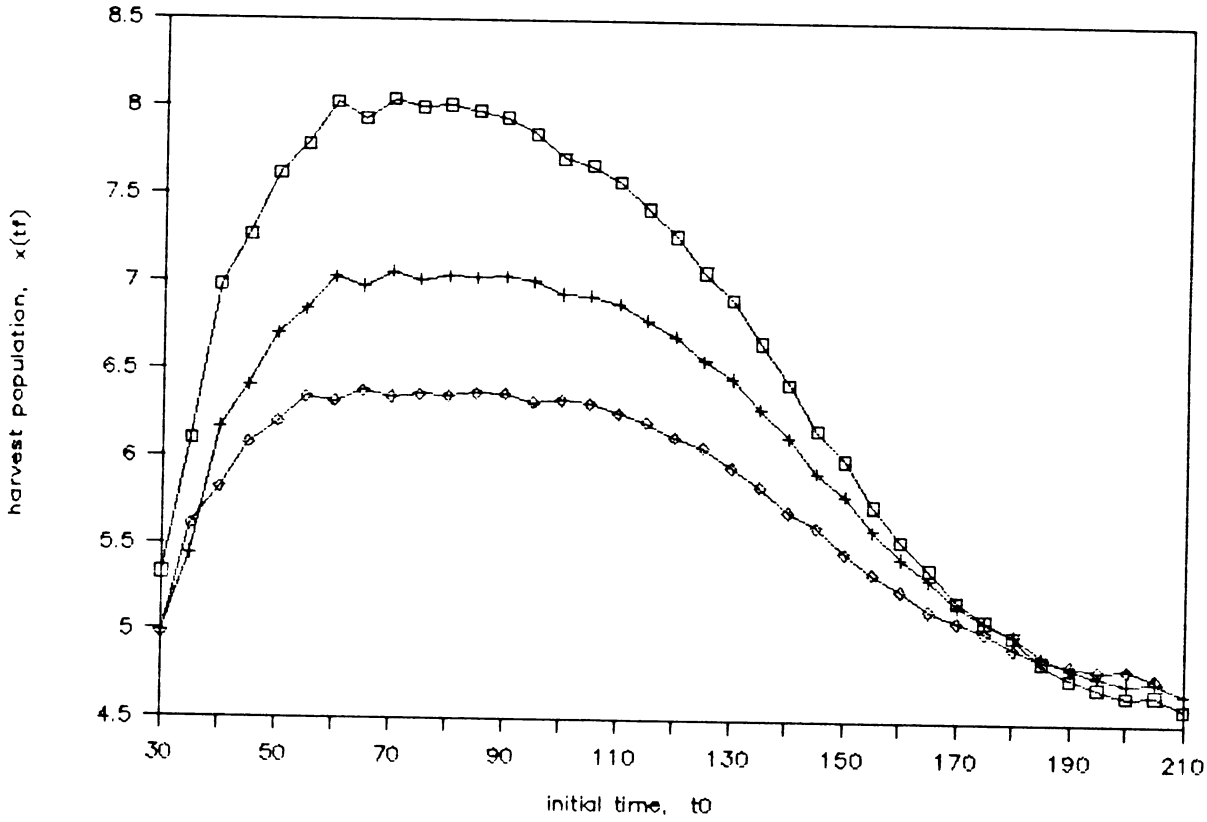


Figure 8.14b Comparison predator/prey mean population metrics summarizing Figures 8.10b, 8.11b, 8.12b. $T_{11}=45^{o}$ F.

- • $T_{10}=15^{o}$ F
- + $T_{10}=20^{o}$ F
- ♦ $T_{10}=25^{o}$ F

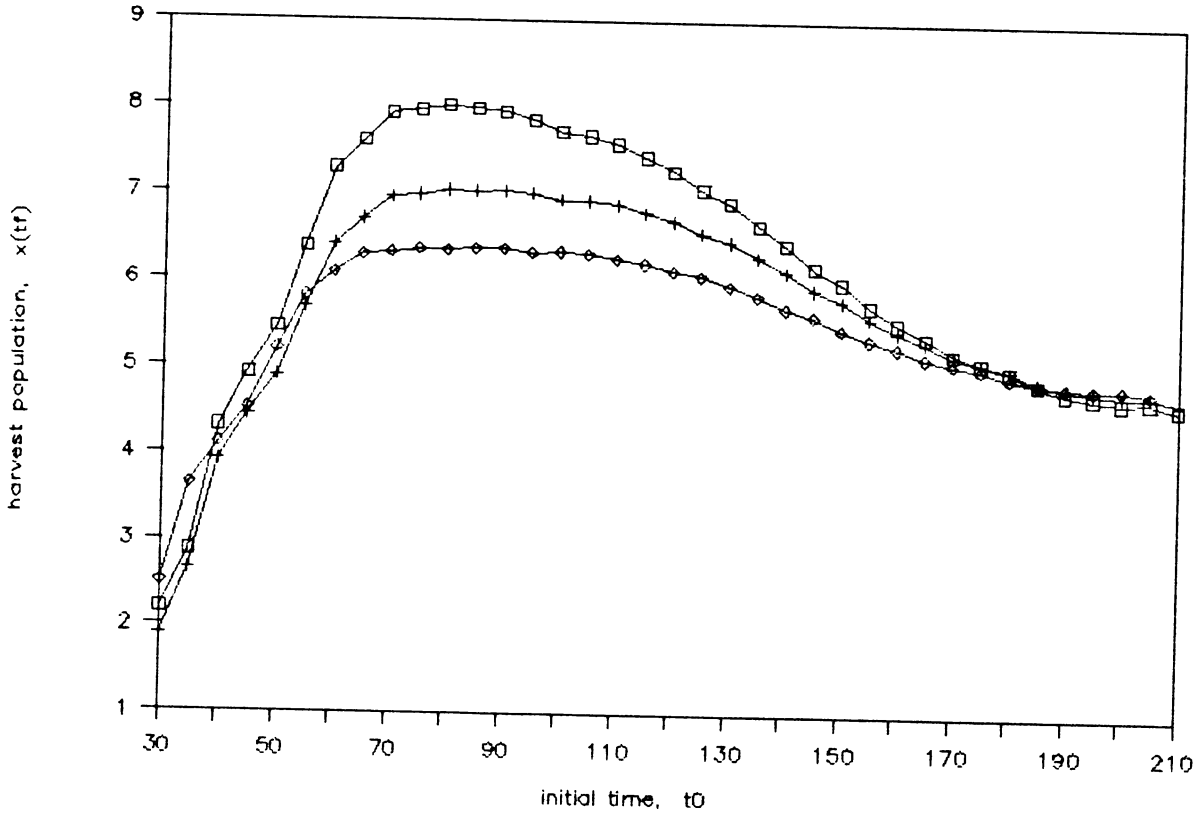


Figure 8.14c Comparison predator/prey mean population metrics summarizing Figures 8.10c, 8.11c, 8.12c. $T_{11}=50^{o}$ F.

- ▪ $T_{10}=15^{o}$ F
- \+ $T_{10}=20^{o}$ F
- ♦ $T_{10}=25^{o}$ F

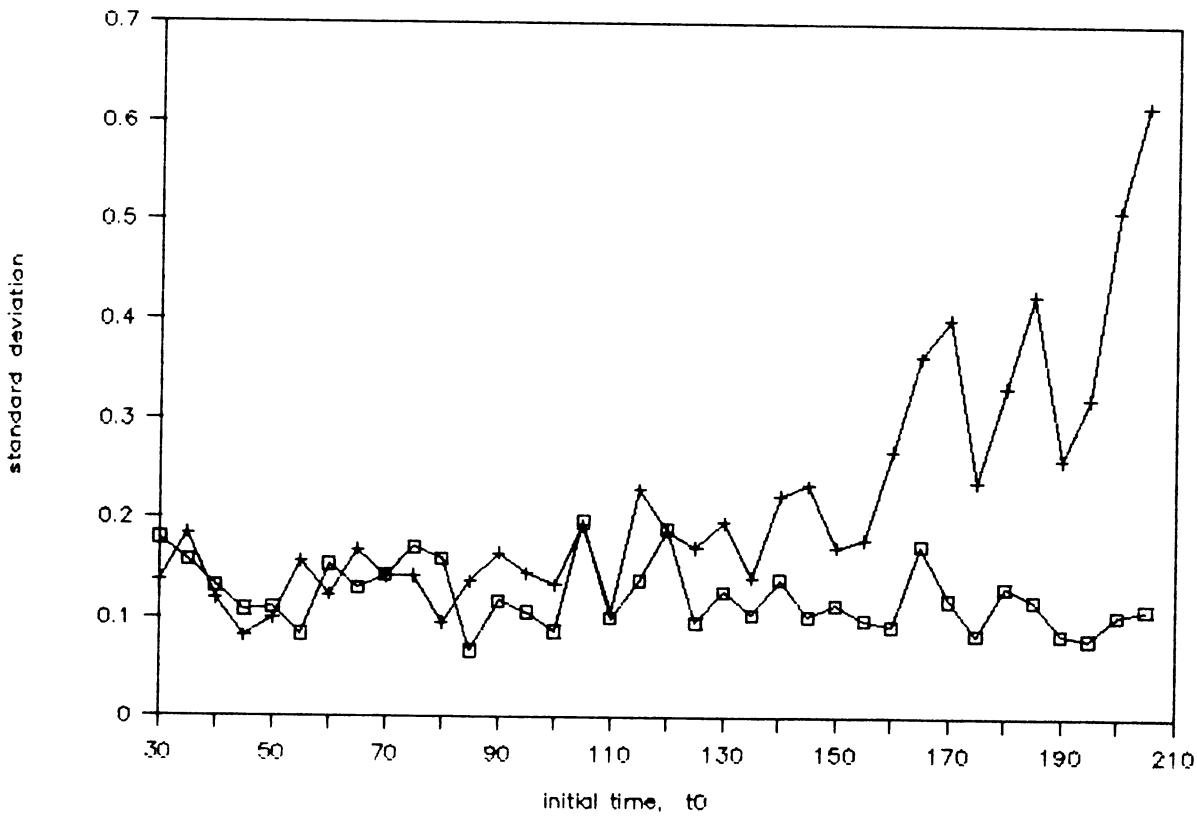


Figure 8.15a Harvest predator/prey variation in population metrics as a function of introduction time with an uncontrolled predator. $T_{10}=15^{o}$, $T_{11}=40^{o}$ F.

▪ prey

\+ predator

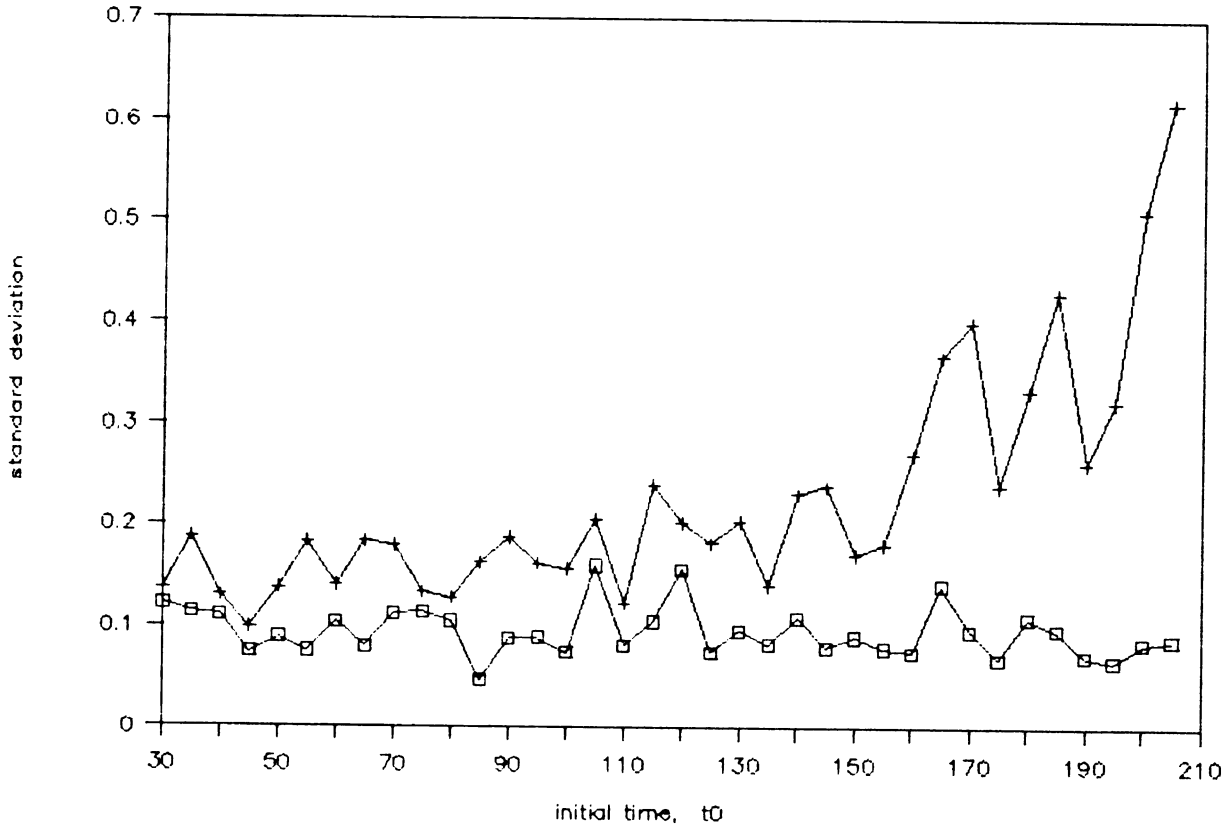


Figure 8.15b Harvest predator/prey variation in population metrics as a function of introduction time with an uncontrolled predator. $T_{10}=15^{o}$, $T_{11}=45^{o}$ F.

▪ prey

\+ predator

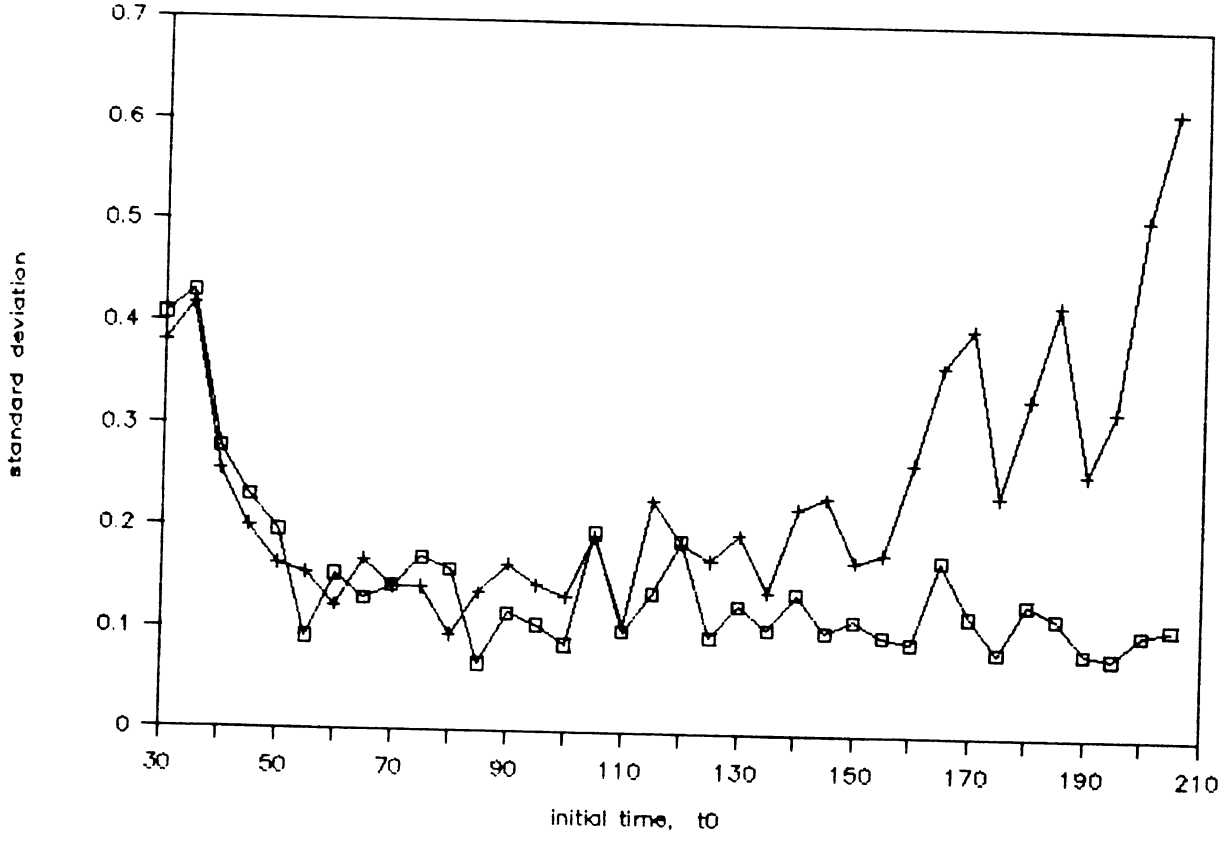


Figure 8.16a Harvest predator/prey variation in population metrics as a function of introduction time with an uncontrolled predator. $T_{10}=20^{o}$, $T_{11}=40^{o}$ F.

▪ prey

\+ predator

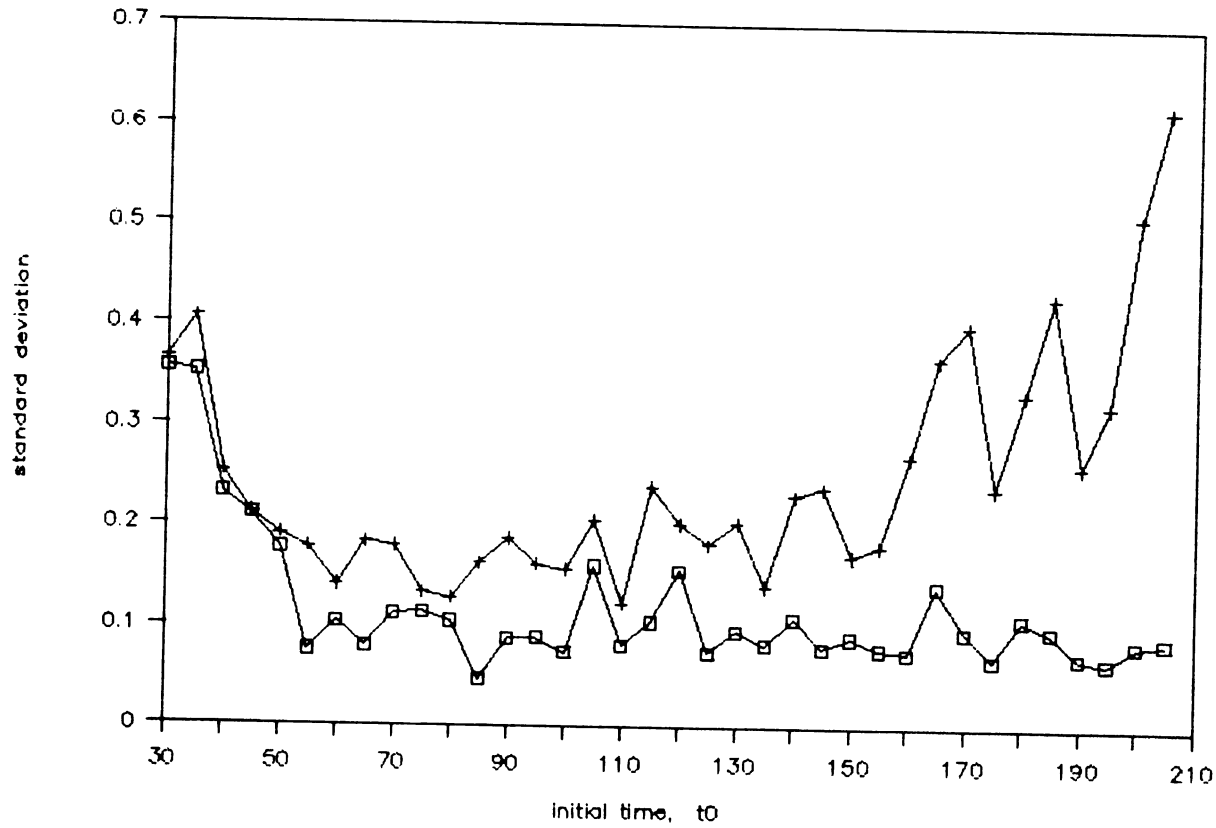


Figure 8.16b Harvest predator/prey variation in population metrics as a function of introduction time with an uncontrolled predator. $T_{10}=20^{o}$, $T_{11}=45^{o}$ F.

▪ prey

\+ predator

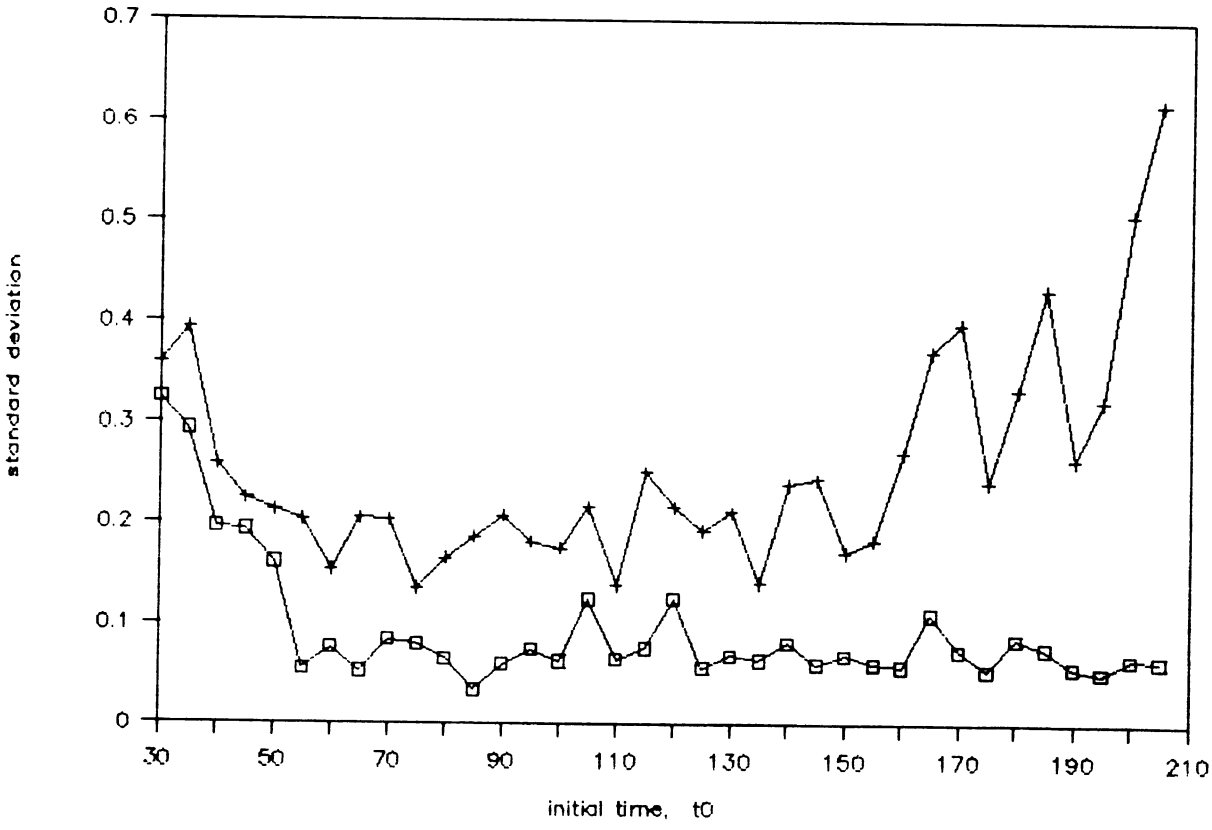


Figure 8.16c Harvest predator/prey variation in population metrics as a function of introduction time with an uncontrolled predator. $T_{10}=20^{o}$, $T_{11}=50^{o}$ F.

▪ prey

\+ predator

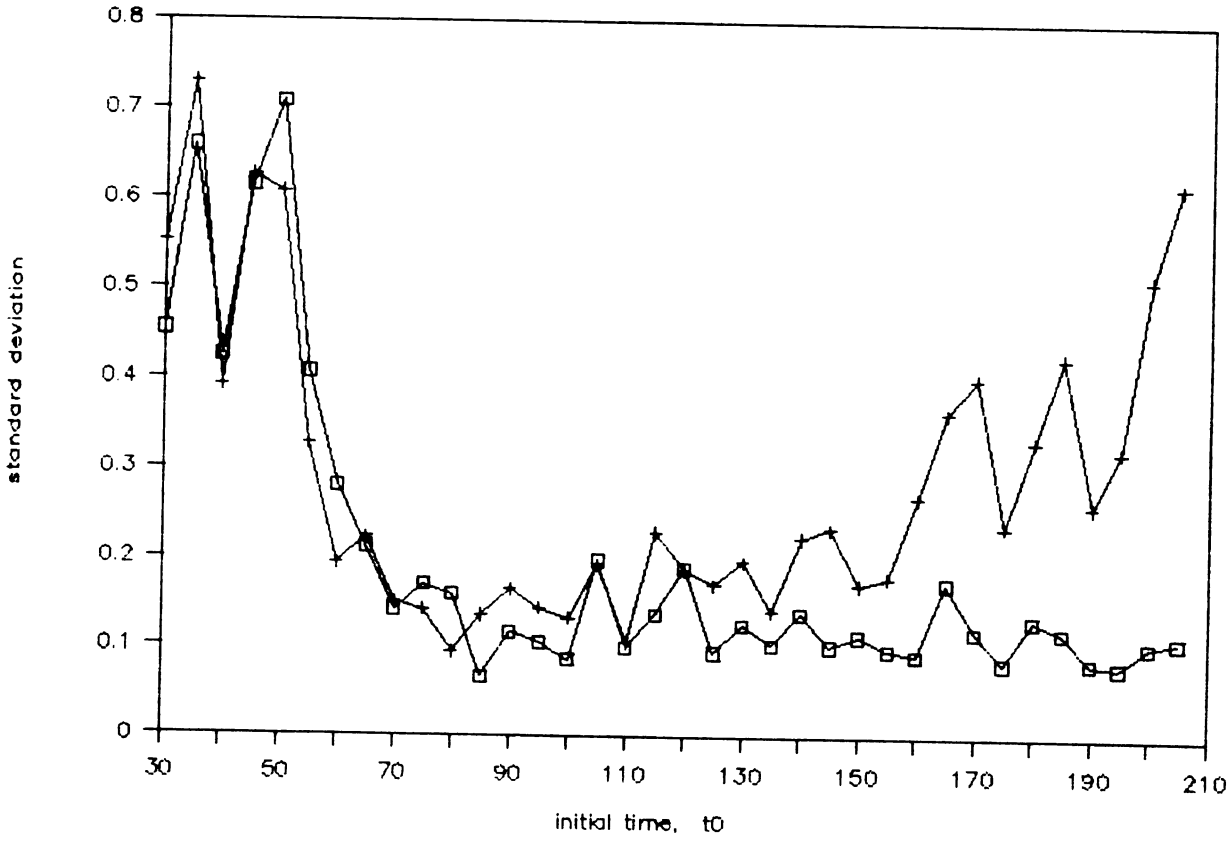


Figure 8.17a Harvest predator/prey variation in population metrics as a function of introduction time with an uncontrolled predator. $T_{10}=25^{o}$, $T_{11}=40^{o}$ F.

▪ prey

\+ predator

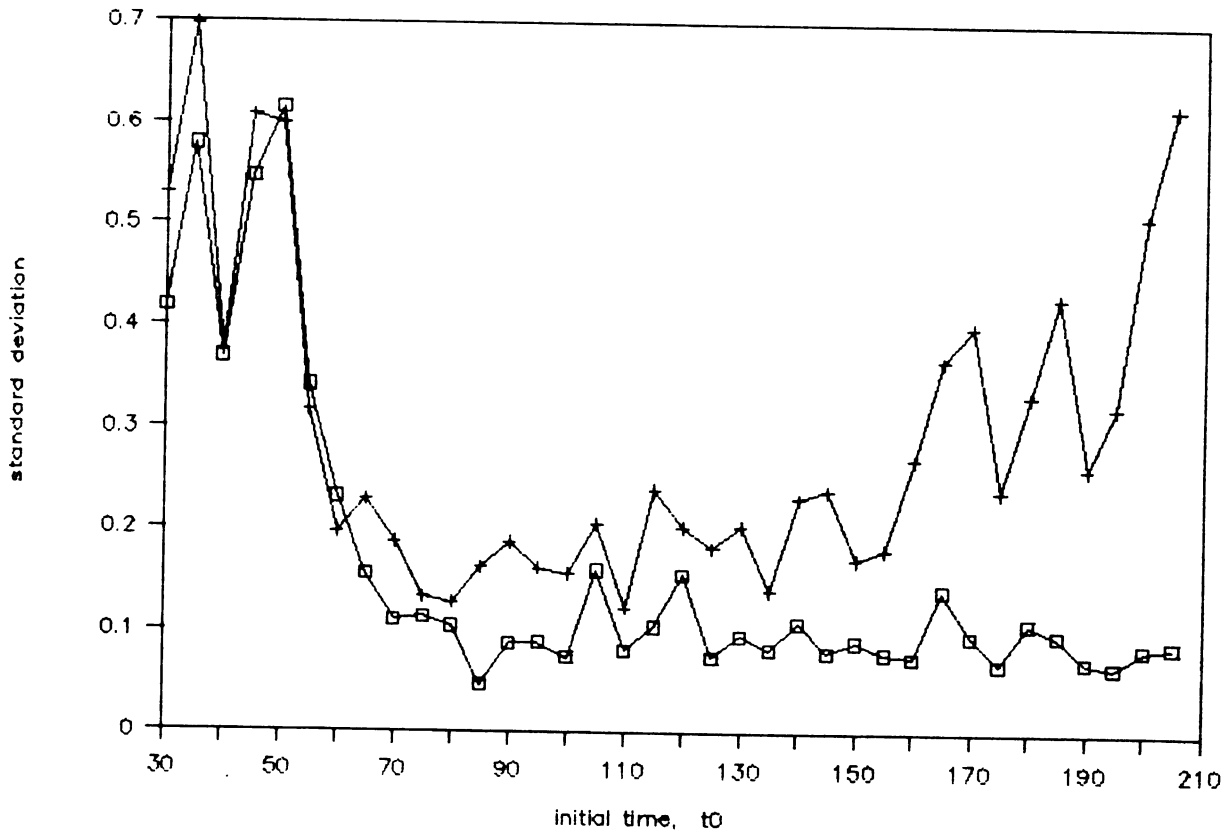


Figure 8.17b Harvest predator/prey variation in population metrics as a function of introduction time with an uncontrolled predator. $T_{10}=25^{o}$, $T_{11}=45^{o}$ F.

▪ prey

\+ predator

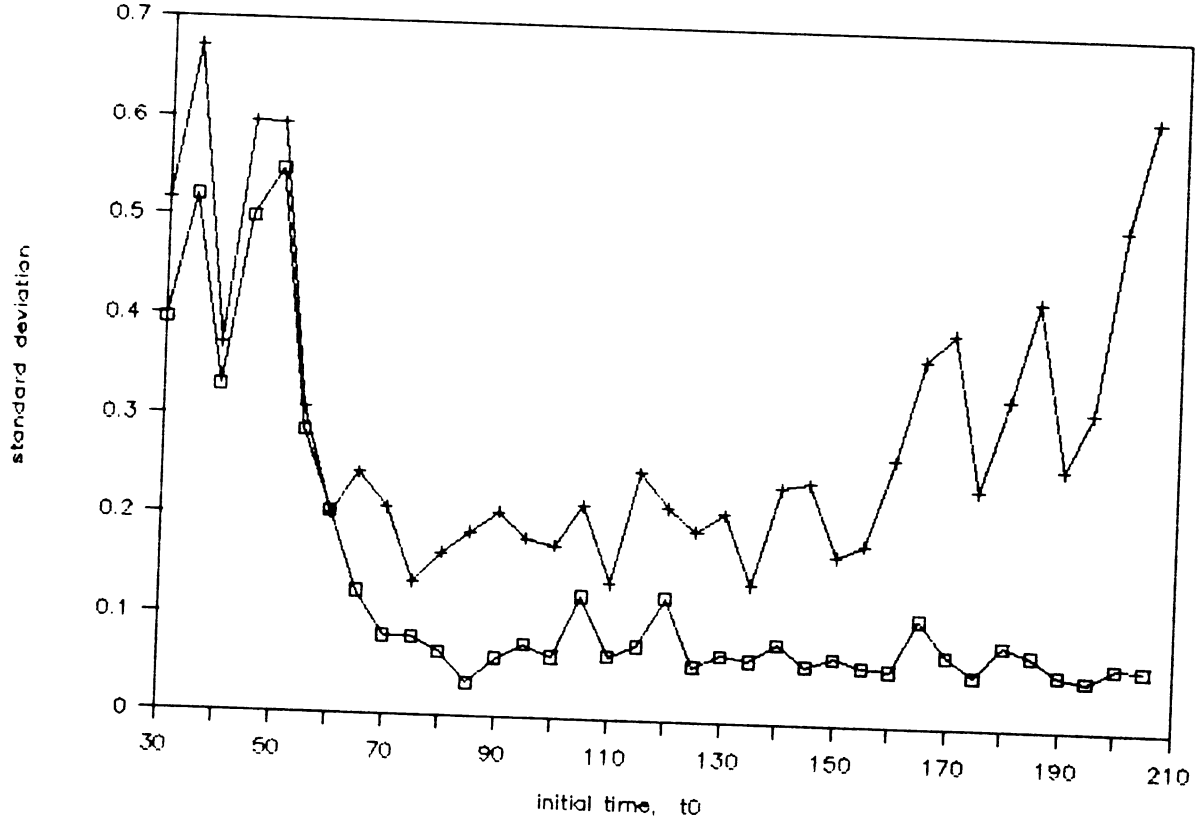


Figure 8.17c Harvest predator/prey variation in population metrics as a function of introduction time with an uncontrolled predator. $T_{10}=25^{o}$, $T_{11}=50^{o}$ F.

▪ prey

\+ predator

IMPLEMENTING CONTROLS

Thus far we have considered results only for systems in which the predator population is uncontrolled. Strictly speaking, as was mentioned earlier, there is a cost associated with control. However, in order to demonstrate further system structure, this will be ignored. The preamble to this chapter outlines a useful approach to this added realism if it is deemed relevant to the problem at hand.

Figures 8.18, 8.19 and 8.20 plot population curves for which feedback fractionating predator controls are exercised. Recall that under this protocol the system is manipulated only when the predator population exceeds a predetermined **control threshold** x_{crit}. Under this circumstance the predator population is **fractionated**, that is to say, it is reduced in proportion to its current size:

[8.23] $$x(t+1) = f \cdot x(t), \qquad 0<f<1$$

where $(1-f)$ is the fraction by which the population is reduced. Figure 8.18 illustrates a response in which $f=0.90$ and $x_{crit}=8$.

Just as predator populations may be fractionated, prey, in the form of crops, might also be fractionated if periodic

harvests are performed. Traditionally this is calendar based rather than feedback based. However, Figure 8.19 shows prey fractionation with f=.75 and x_{crit}=10 with the protocol as outlined above. Figure 8.20 shows fractionating for both predator and prey using the same values.

These population curves are interesting, but in order to compare with their uncontrolled counterparts, harvest results $x_1(t_f)$ are plotted against prey introduction time t_{20} as in the previous section. This has been done in Figures 8.21, 8.22 and 8.23. The obvious effect on control is clear. The predator population is effectually forced to remain constant over the entire active interval. Because of this the prey harvest is considerably improved. A quick comparison of Figures 8.10, 8.11 and 8.12 (uncontrolled predator) against Figures 8.21, 8.22 and 8.23 (controlled predator) reveals an over-all improvement by an approximate 250% factor.

Other differences are more subtle. Figures 8.24 (to be compared against the uncontrolled counterparts: Figures 8.13) show the same optimal interval characteristics and nesting properties as the uncontrolled predator models. However, this time there is little overlap. Apparently by controlling the predator population, a more distinct

difference exists between the dormant thresholds, regardless of the lethal threshold value.

On the other hand, Figures 8.25 (compared against Figures 8.14) illustrate an extremely quick convergence of harvest at low introduction times. No matter what the dormancy threshold value, all trajectories converge quite early on! That is, in a controlled environment, the optimal harvest will be achieved by any prey on a large optimal interval, regardless of the dormancy threshold.

The results in variation are even more striking. While in the uncontrolled environment we see very moderate prey variation, but an increasing and significant predator variation, the controlled results are opposite. As would be expected, a controlled predator will have minimal variation. This is verified in Figures 8.26, 8.27 and 8.28 (Notice the very interesting anomaly in Figure 8.28c)! However, the prey variation has become more erratic and pronounced. Even so, there is a distinct tendency for prey variation to decrease with time. Although this is moderate for low (15^{o} F) death thresholds, as T_{10} increases, the early variation is quite significant.

In summary, a controlled environment results in the variation roles of predator and prey populations being reversed. In the uncontrolled case prey variation is

essentially constant and the predator's variation increases. For the controlled case, predator variation is relatively fixed but the prey variation decreases. There are, indeed, underlying patterns.

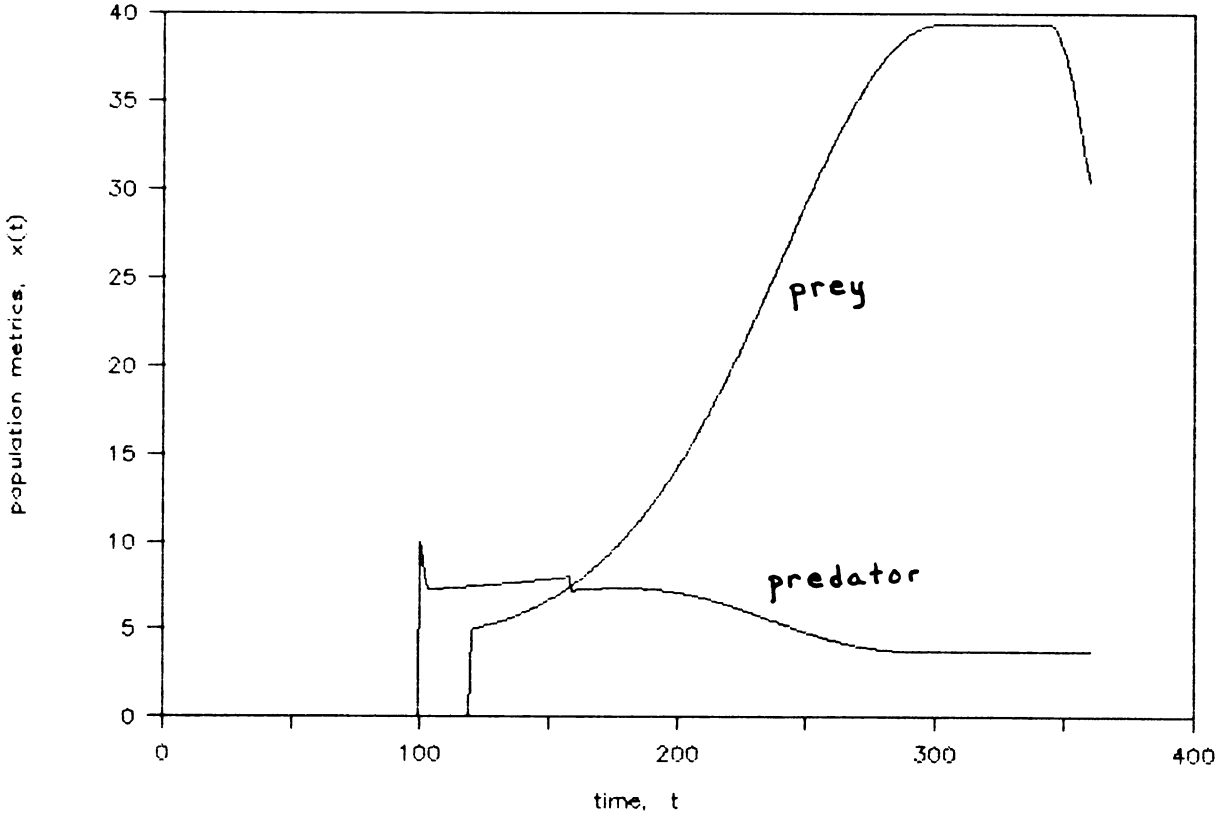


Figure 8.18 Population metric over time with controlled predator, uncontrolled prey.

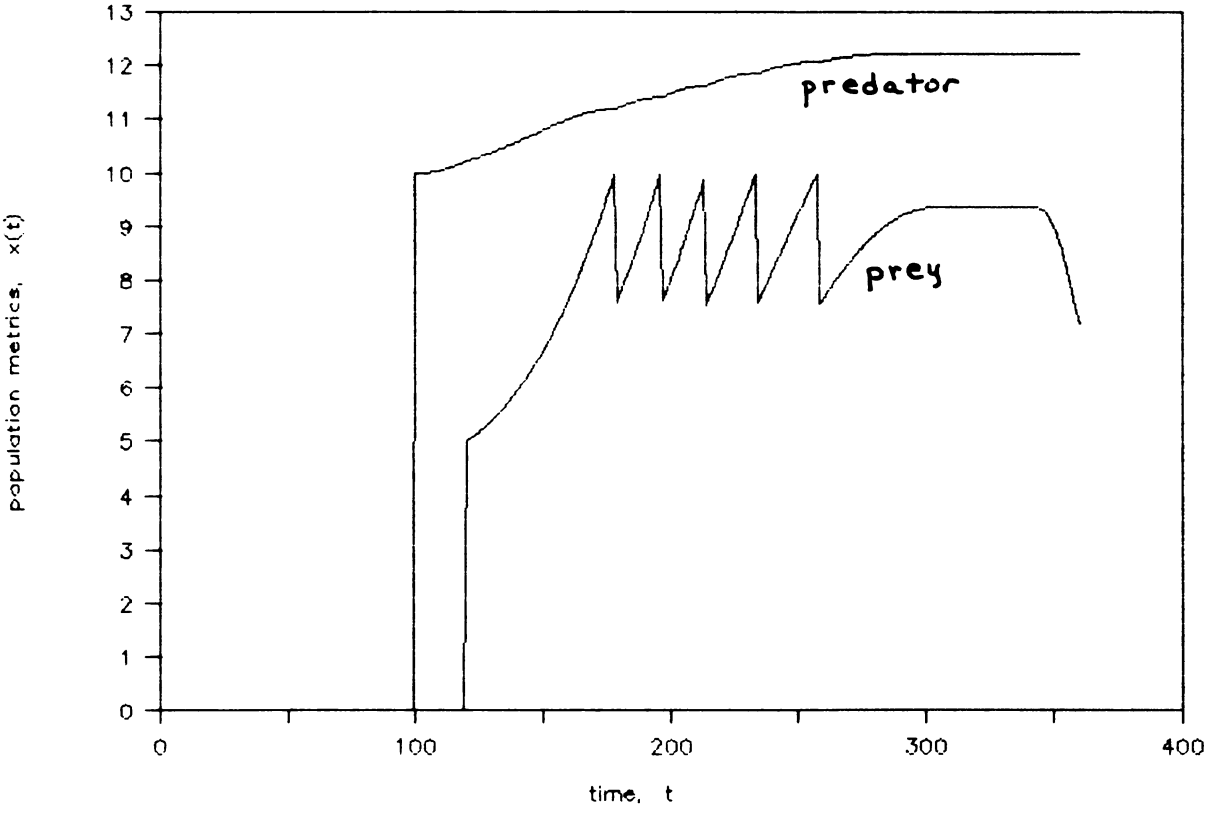


Figure 8.19 Population metric over time with uncontrolled predator, controlled prey.

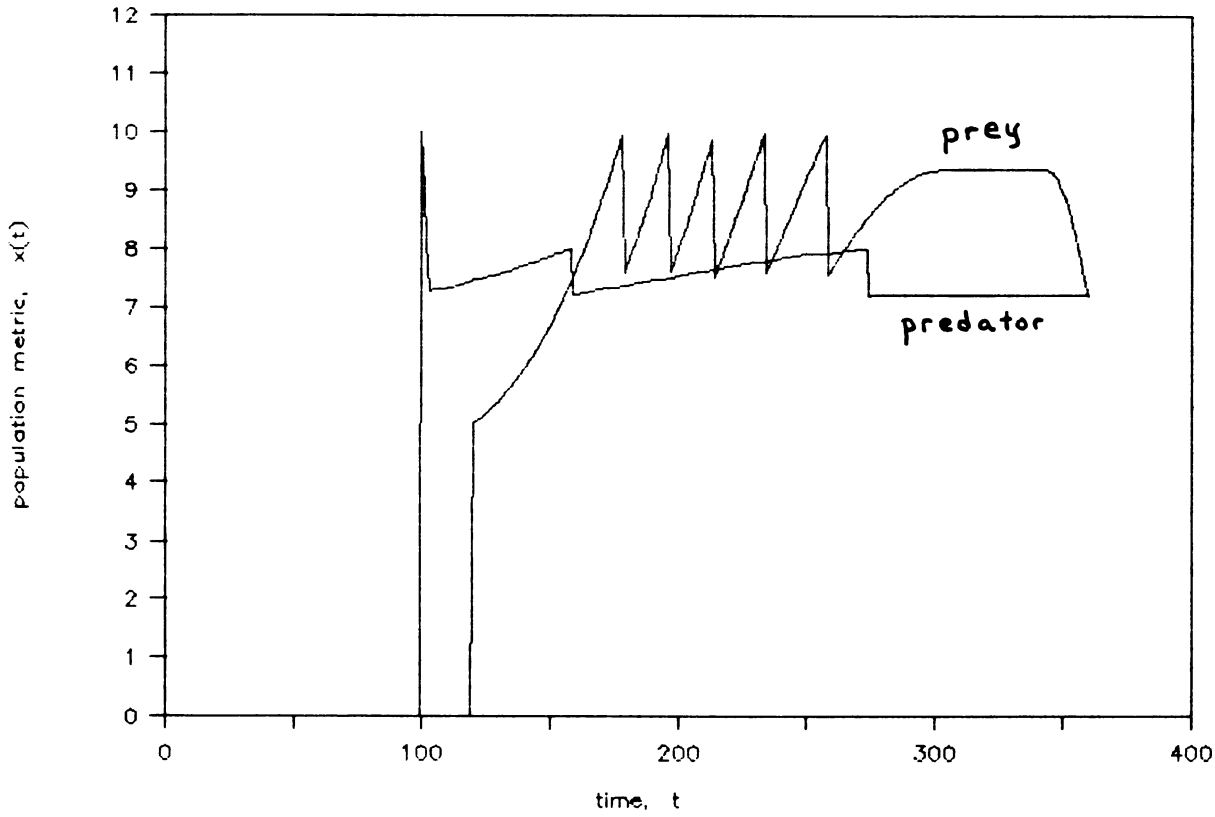


Figure 8.20 Population metric over time with controlled predator, controlled prey.

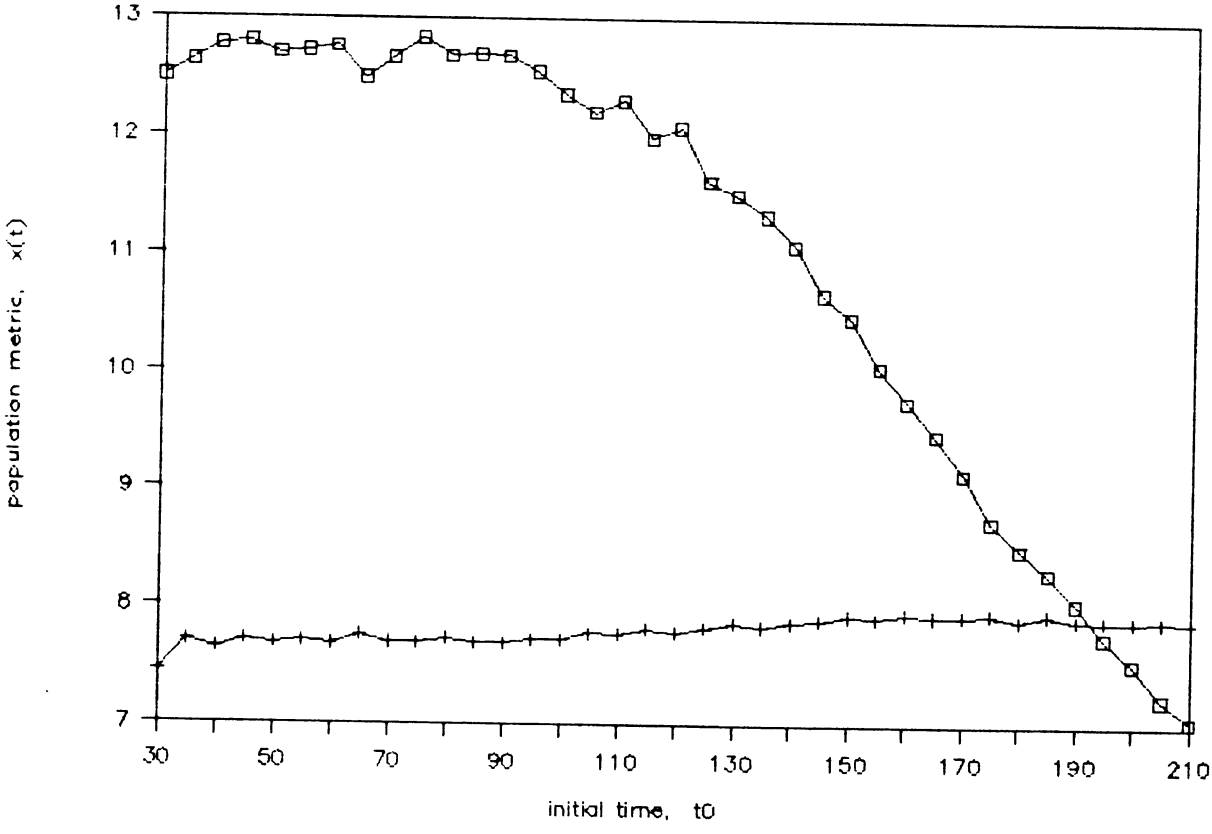


Figure 8.21a Harvest predator/prey mean population metrics as a function of introduction time with a controlled predator. $T_{10}=15^{o}$, $T_{11}=40^{o}$ F.

▪ prey

\+ predator

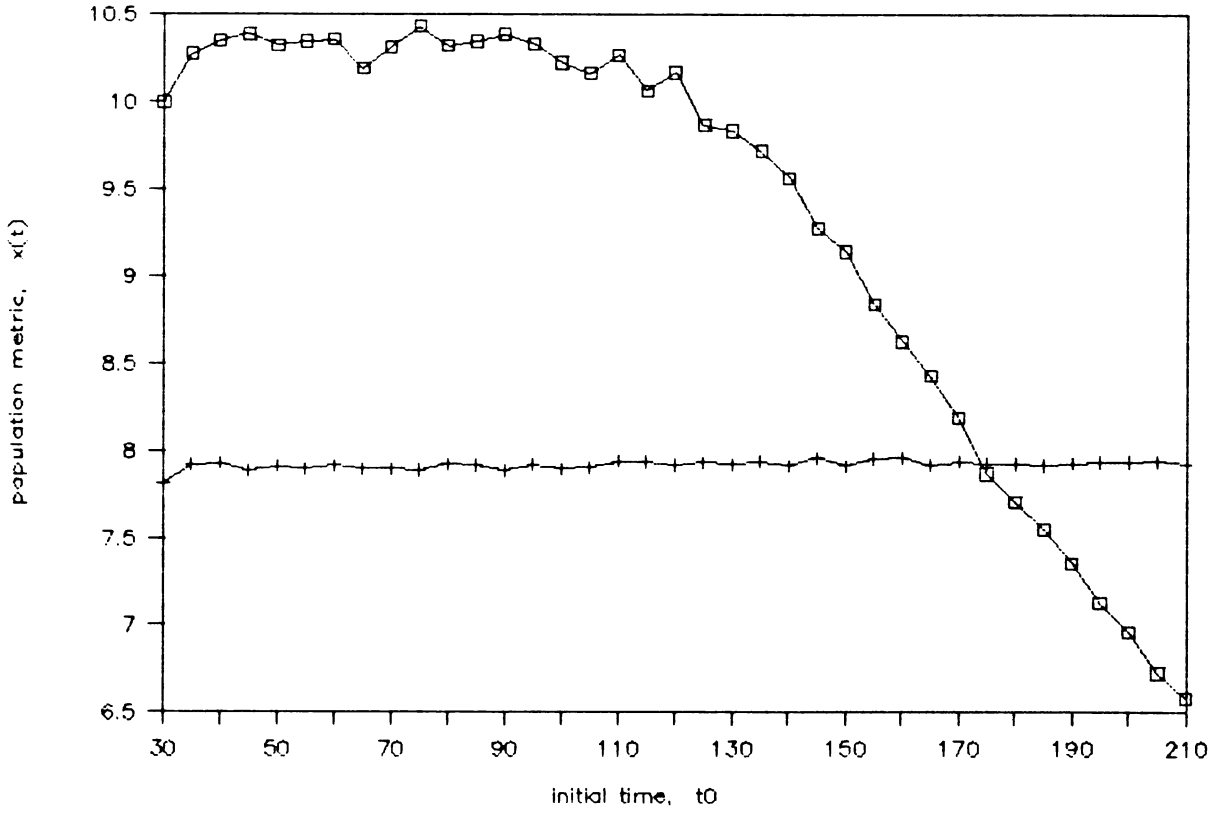


Figure 8.21b Harvest predator/prey mean population metrics as a function of introduction time with a controlled predator. $T_{10}=15^{o}$, $T_{11}=45^{o}$ F.

▪ prey

\+ predator

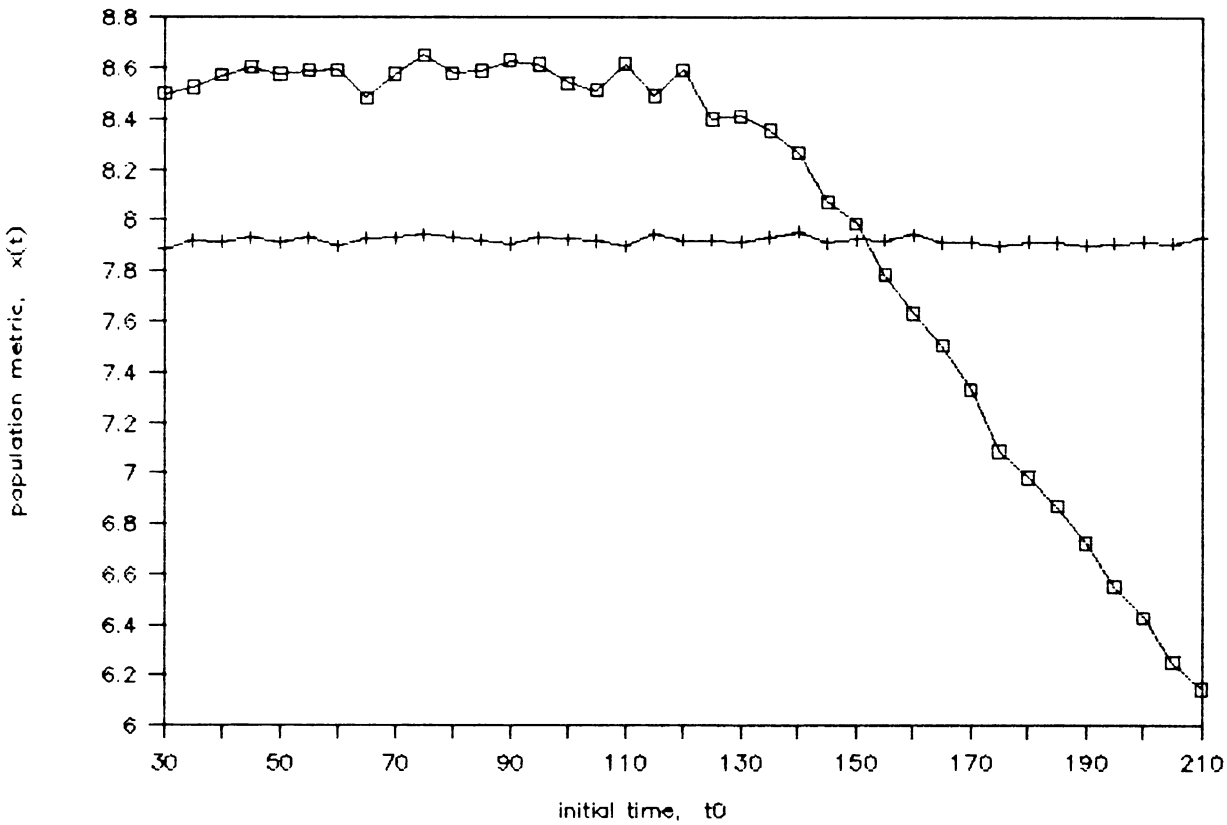


Figure 8.21c Harvest predator/prey mean population metrics as a function of introduction time with a controlled predator. $T_{10}=15^{o}$, $T_{11}=50^{o}$ F.

• prey

+ predator

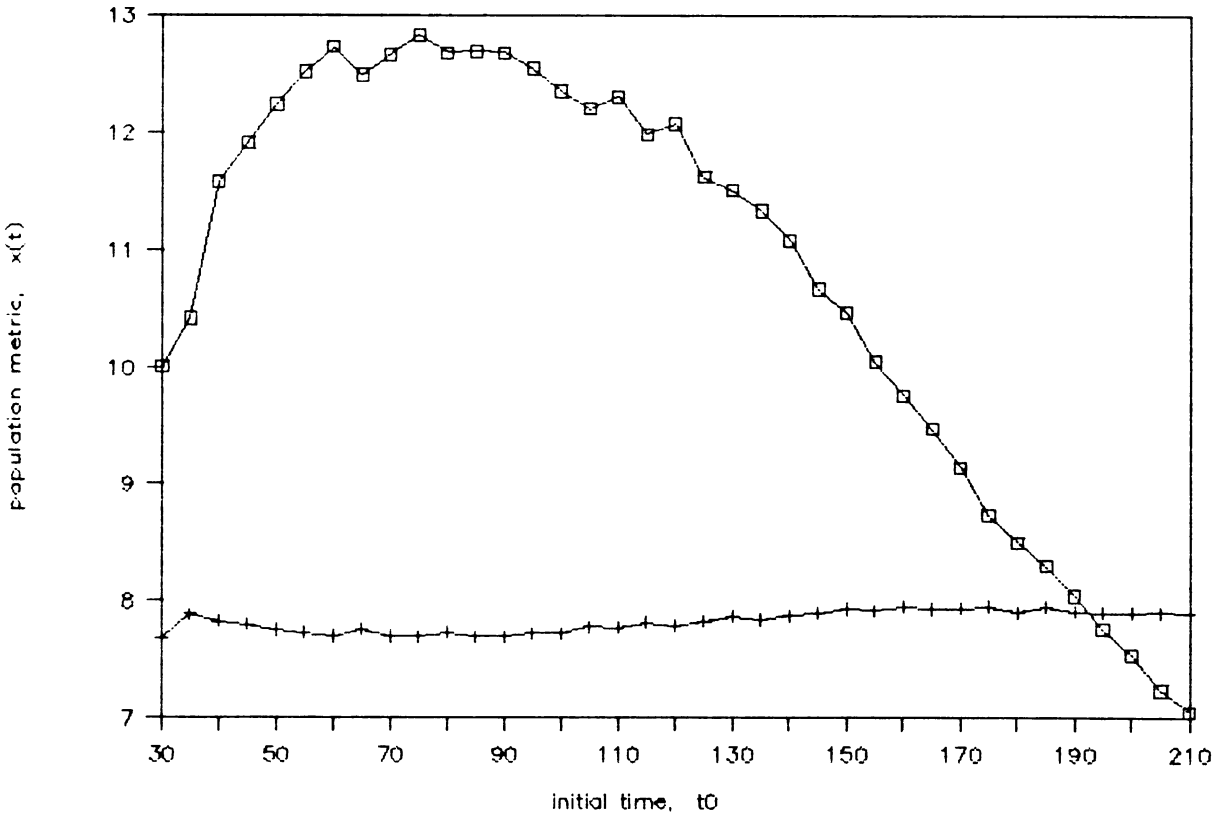


Figure 8.22a Harvest predator/prey mean population metrics as a function of introduction time with a controlled predator. $T_{10}=20^{o}$, $T_{11}=40^{o}$ F.

▪ prey

\+ predator

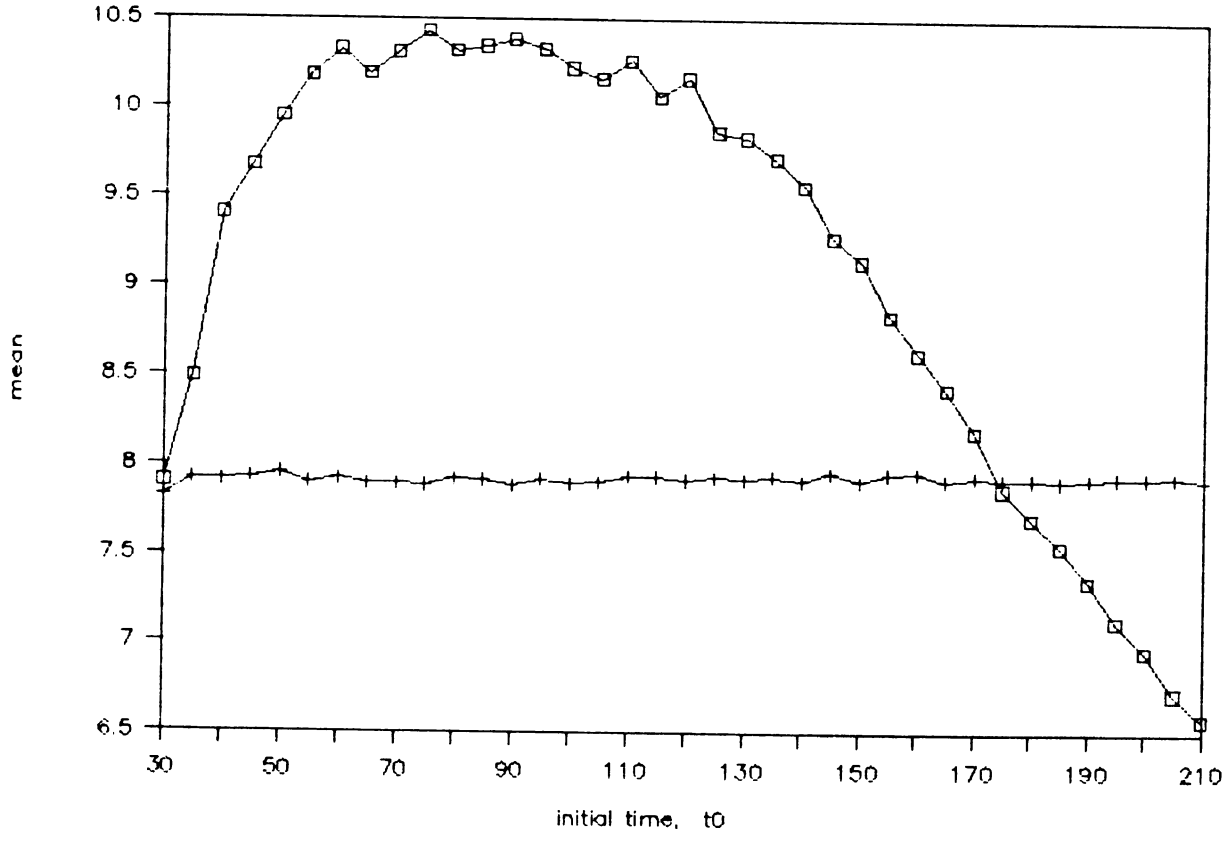


Figure 8.22b Harvest predator/prey mean population metrics as a function of introduction time with a controlled predator. $T_{10}=20^{o}$, $T_{11}=45^{o}$ F.

• prey

\+ predator

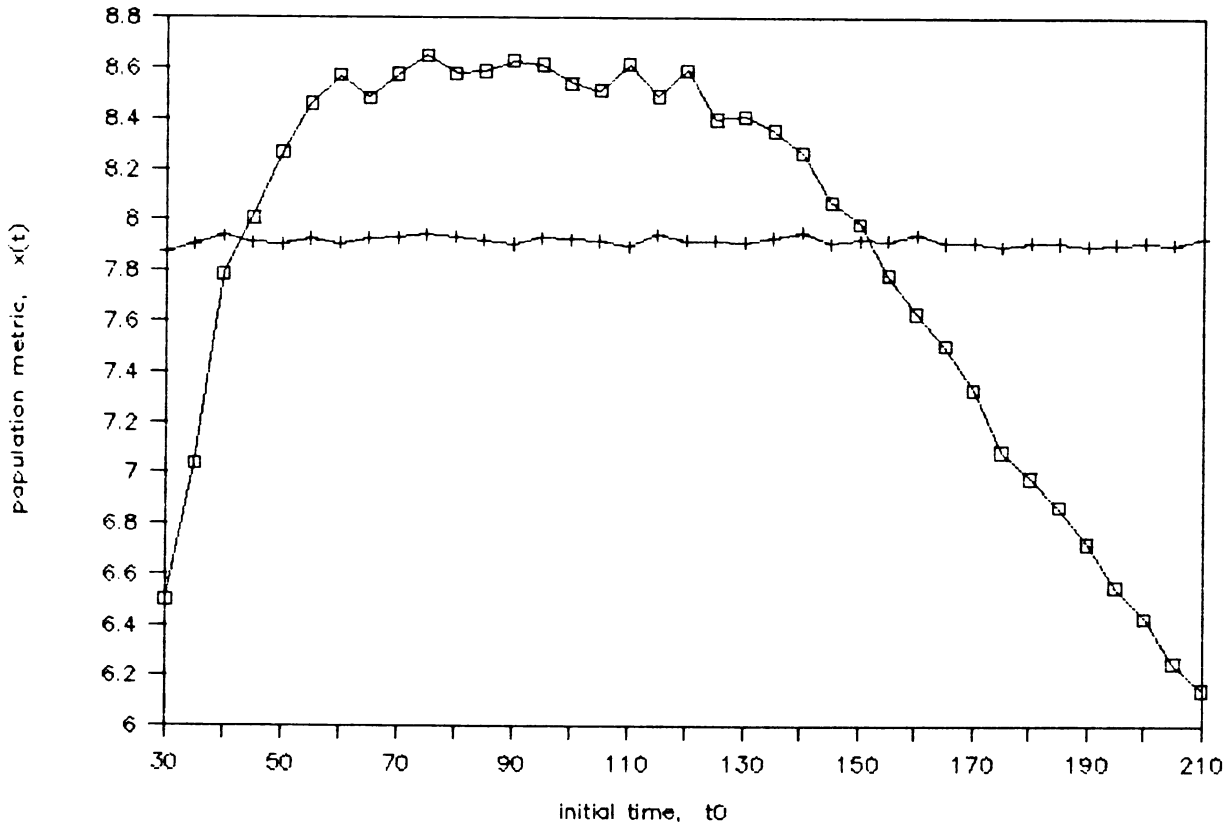


Figure 8.22c Harvest predator/prey mean population metrics as a function of introduction time with a controlled predator. $T_{10}=20^{o}$, $T_{11}=50^{o}$ F.

• prey

+ predator

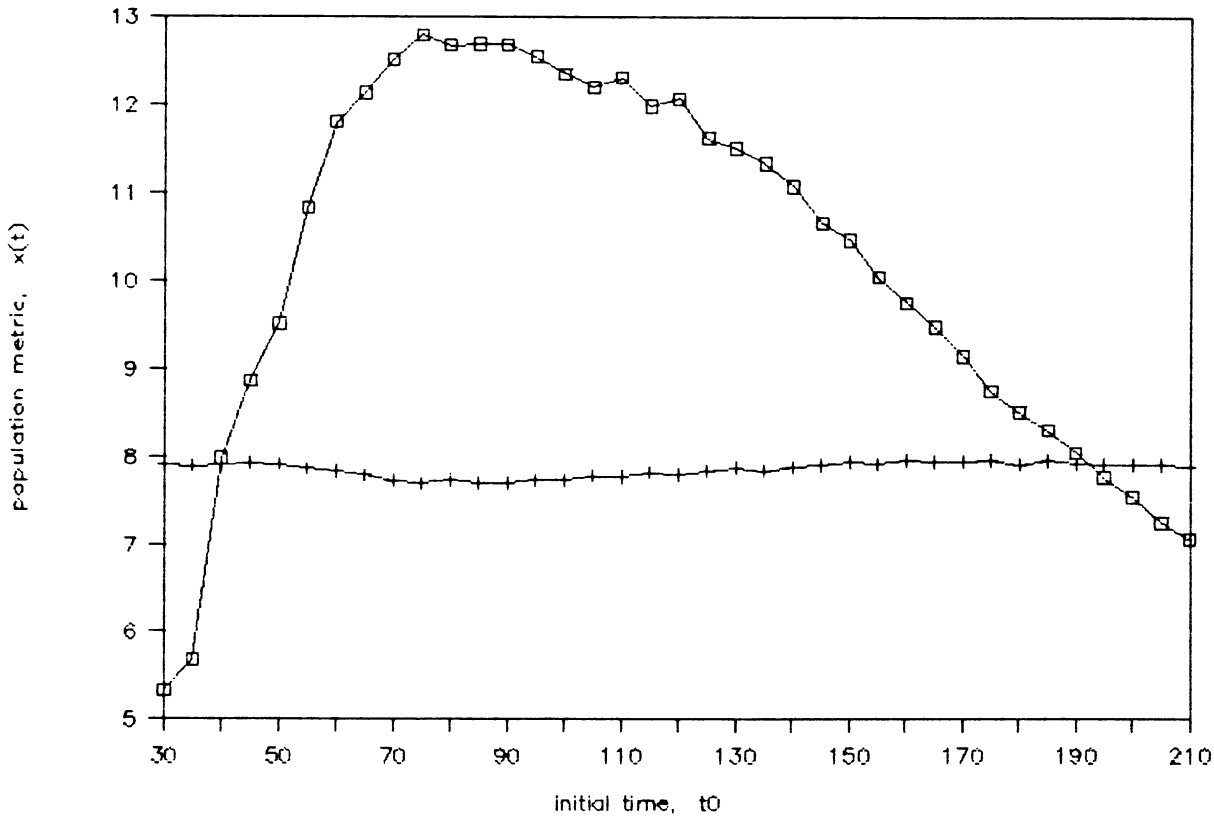


Figure 8.23a Harvest predator/prey mean population metrics as a function of introduction time with a controlled predator. $T_{10}=25^{o}$, $T_{11}=40^{o}$ F.

▪ prey

\+ predator

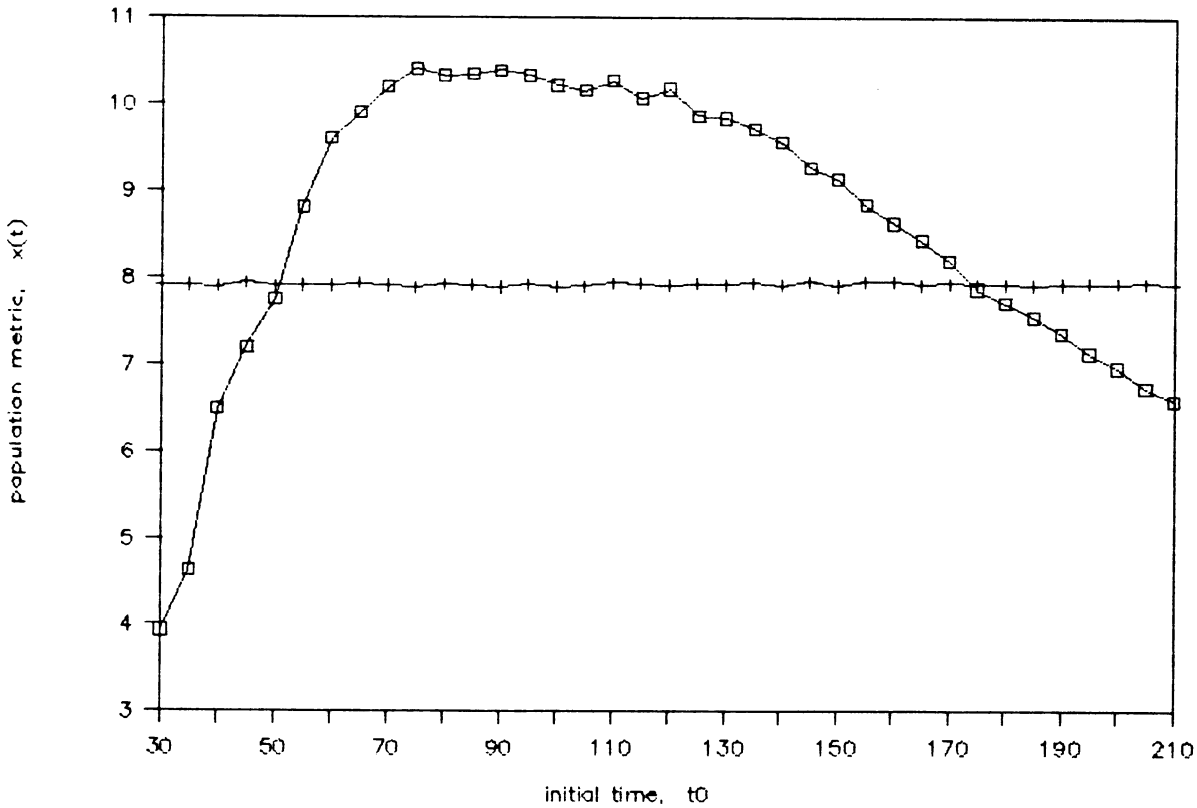


Figure 8.23b Harvest predator/prey mean population metrics as a function of introduction time with a controlled predator. $T_{10}=25^{o}$, $T_{11}=45^{o}$ F.

▪ prey

\+ predator

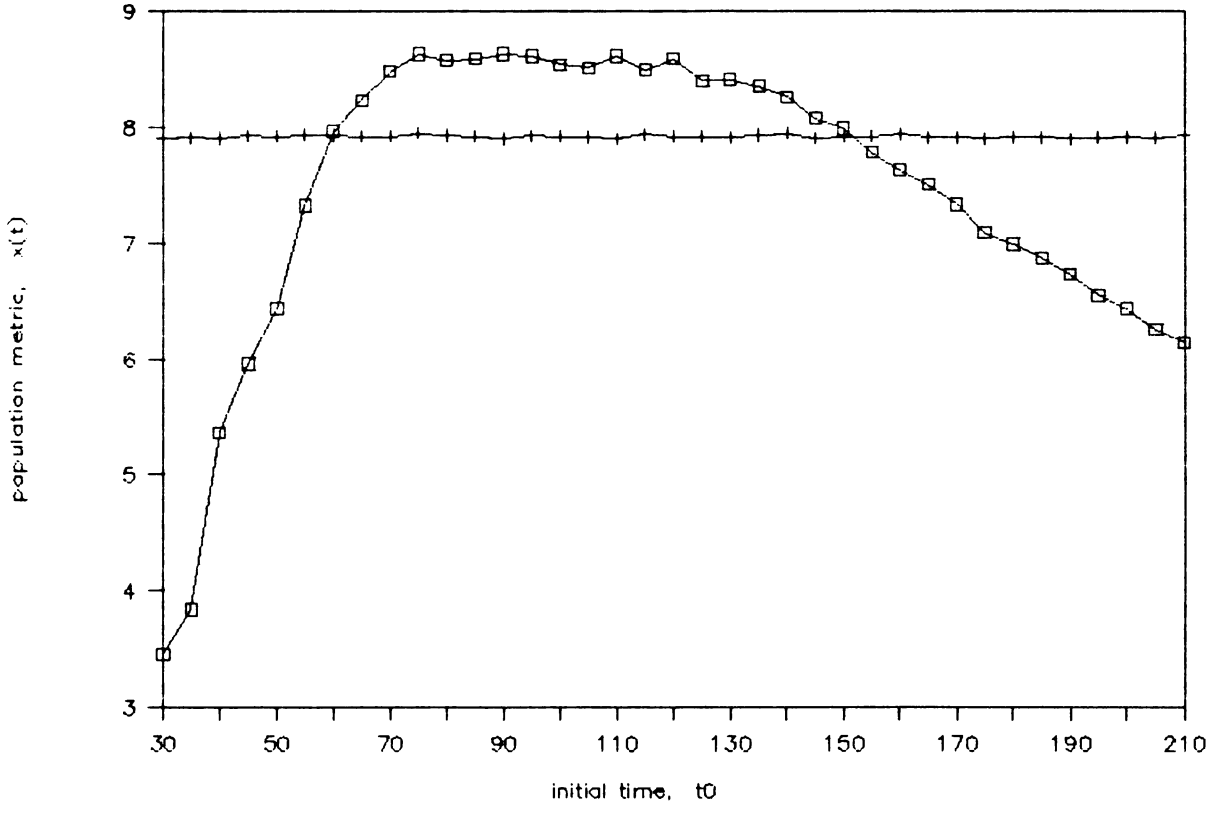


Figure 8.23c Harvest predator/prey mean population metrics as a function of introduction time with a controlled predator. $T_{10}=25^{o}$, $T_{11}=50^{o}$ F.

▪ prey

\+ predator

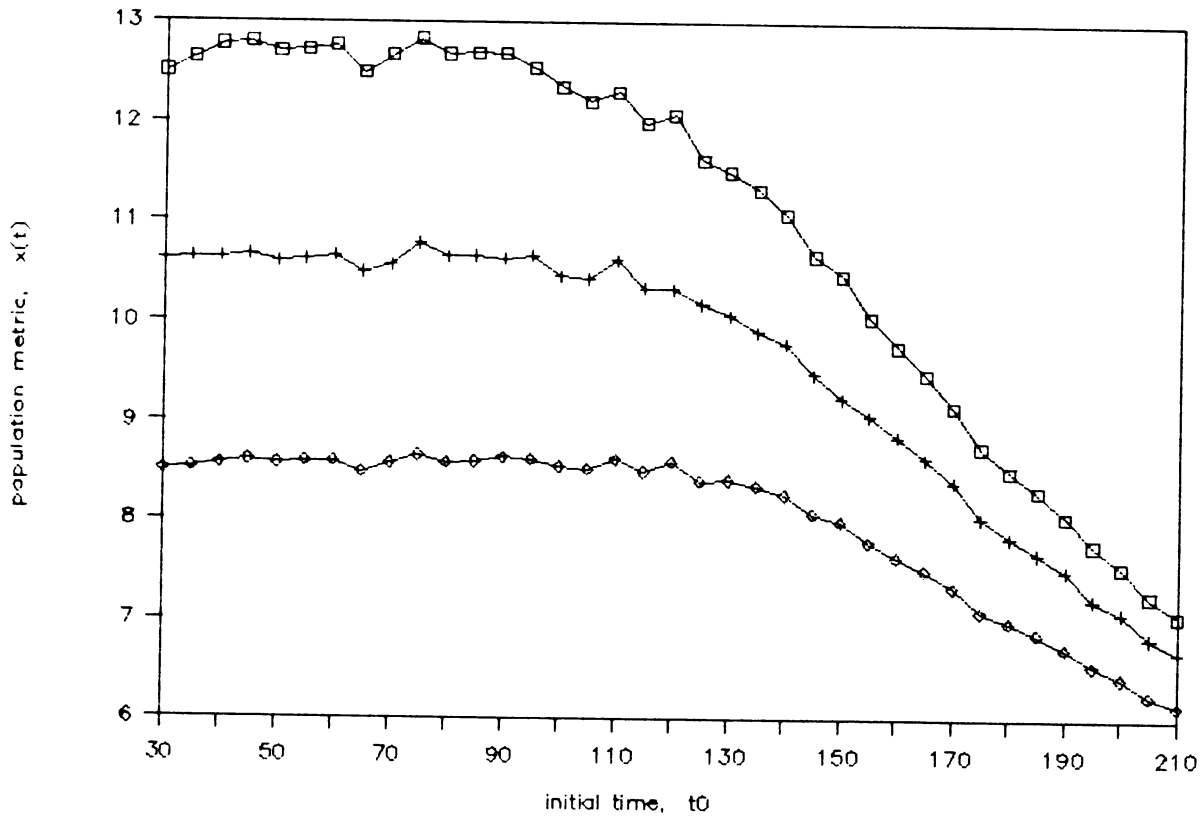


Figure 8.24a Comparison predator/prey mean population metrics summarizing Figures 8.21. $T_{10}=15^{o}$ F.

- ▪ $T_{11}=40^{o}$ F
- \+ $T_{11}=45^{o}$ F
- ♦ $T_{11}=50^{o}$ F

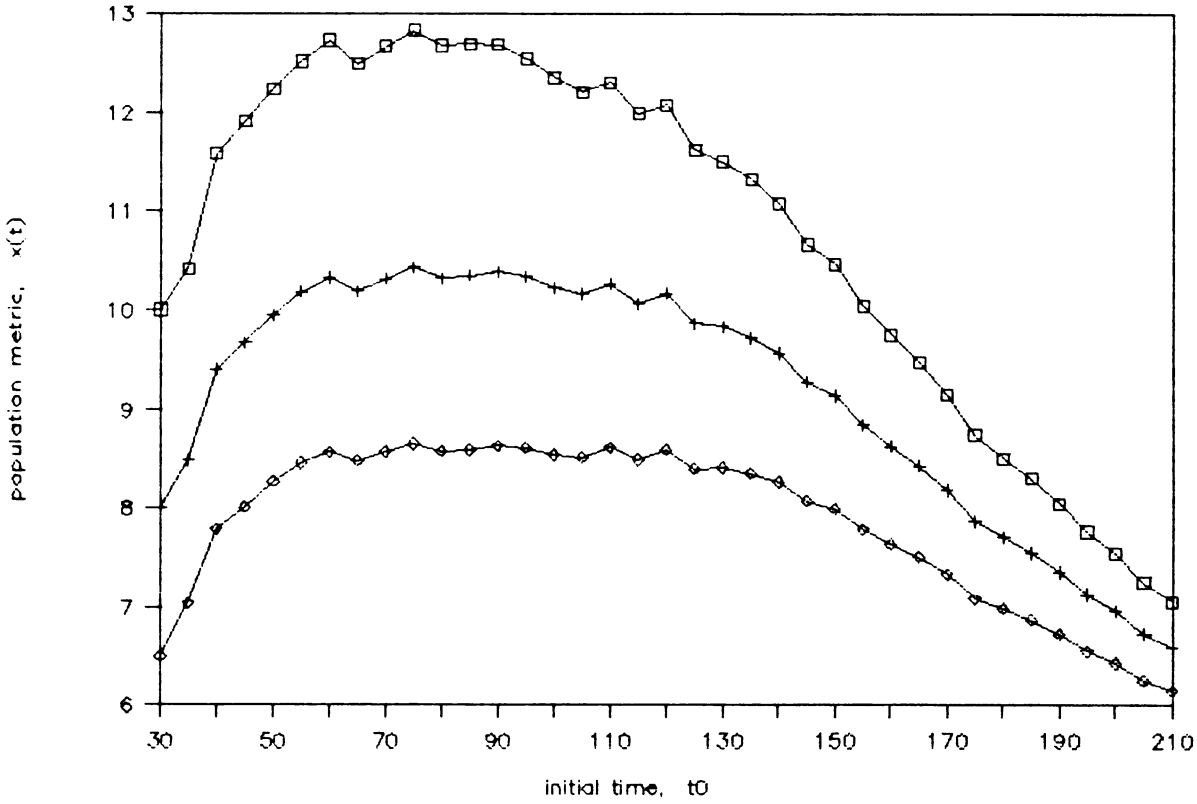


Figure 8.24b Comparison predator/prey mean population metrics summarizing Figures 8.22. $T_{10}=20^{o}$ F.

▪ $T_{11}=40^{o}$ F

\+ $T_{11}=45^{o}$ F

♦ $T_{11}=50^{o}$ F

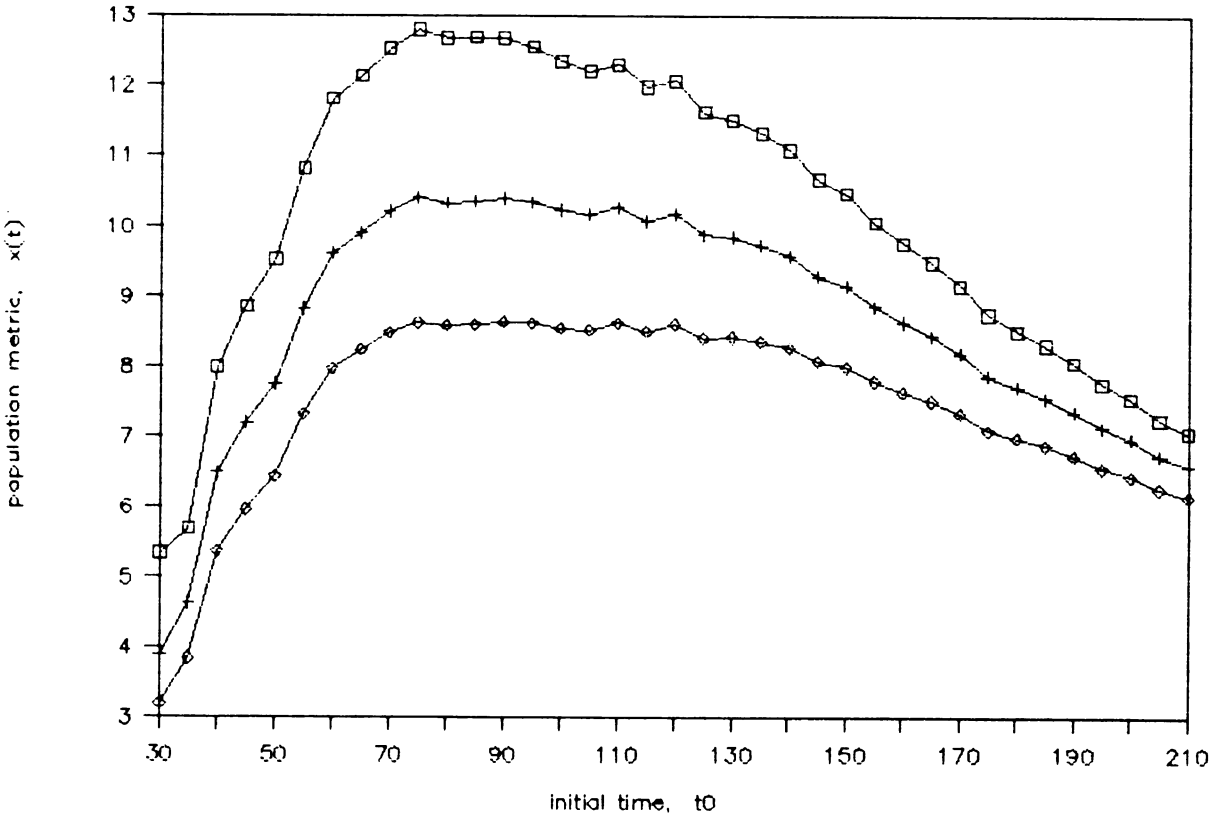


Figure 8.24c Comparison predator/prey mean population metrics summarizing Figures 8.23. $T_{10}=25^{o}$ F.

▪ $T_{11}=40^{o}$ F

\+ $T_{11}=45^{o}$ F

♦ $T_{11}=50^{o}$ F

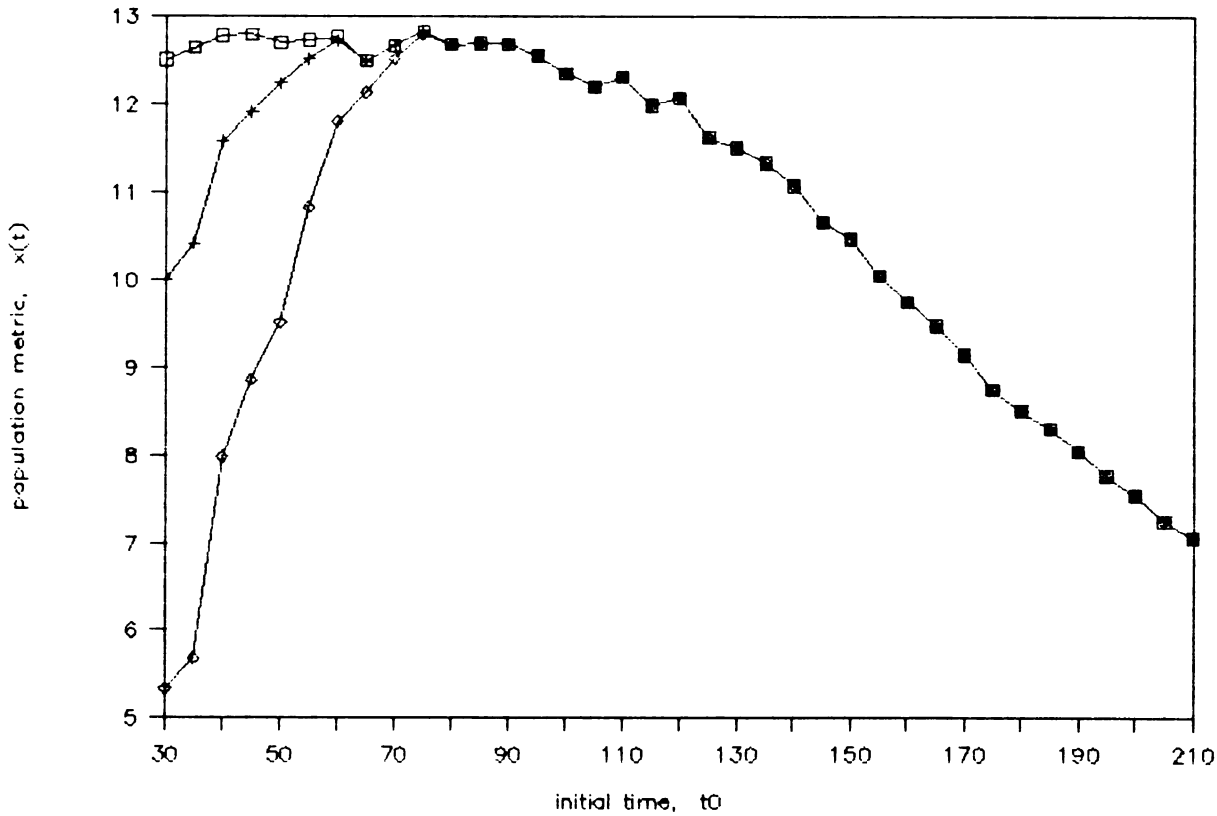


Figure 8.25a Comparison predator/prey mean population metrics summarizing Figures 8.21a, 8.22a, 8.23a. $T_{11}=40^{o}$ F.

▪ $T_{10}=15^{o}$ F

\+ $T_{10}=20^{o}$ F

♦ $T_{10}=25^{o}$ F

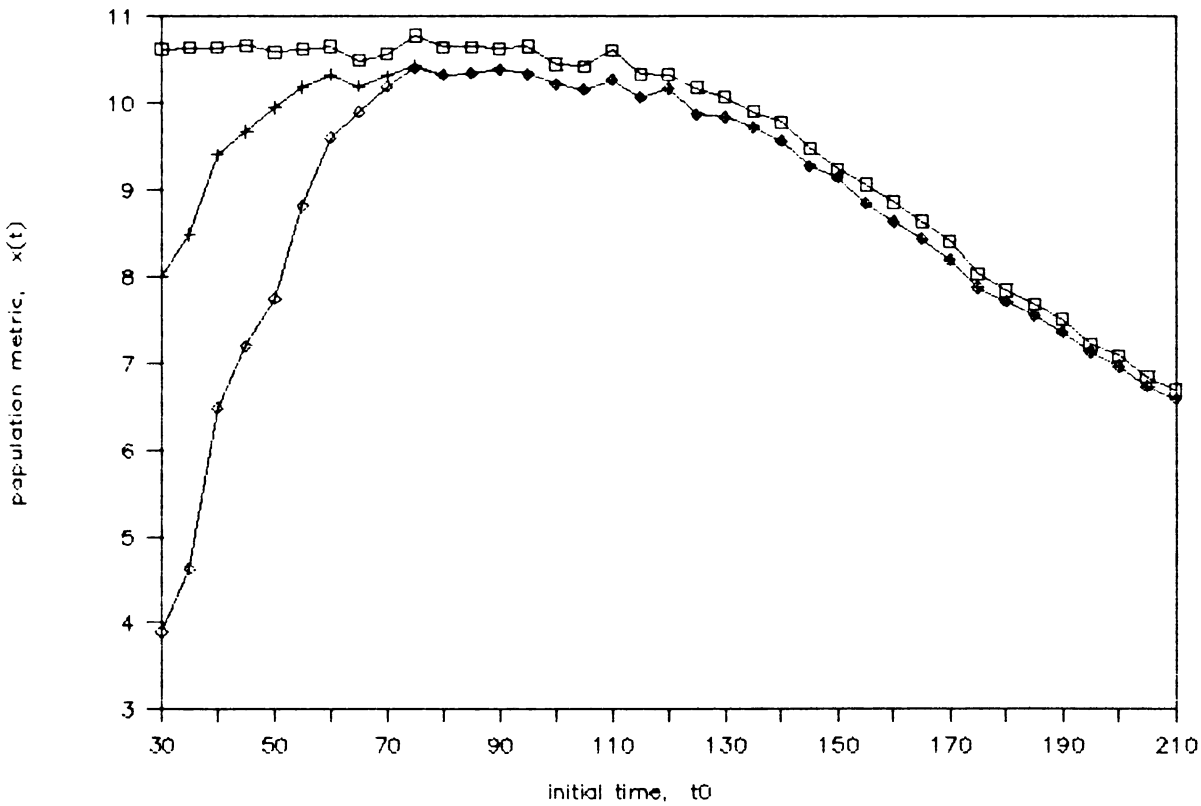


Figure 8.25b Comparison predator/prey mean population metrics summarizing Figures 8.21b, 8.22b, 8.23b. $T_{11}=45^{o}$ F.

▪ $T_{10}=15^{o}$ F

\+ $T_{10}=20^{o}$ F

♦ $T_{10}=25^{o}$ F

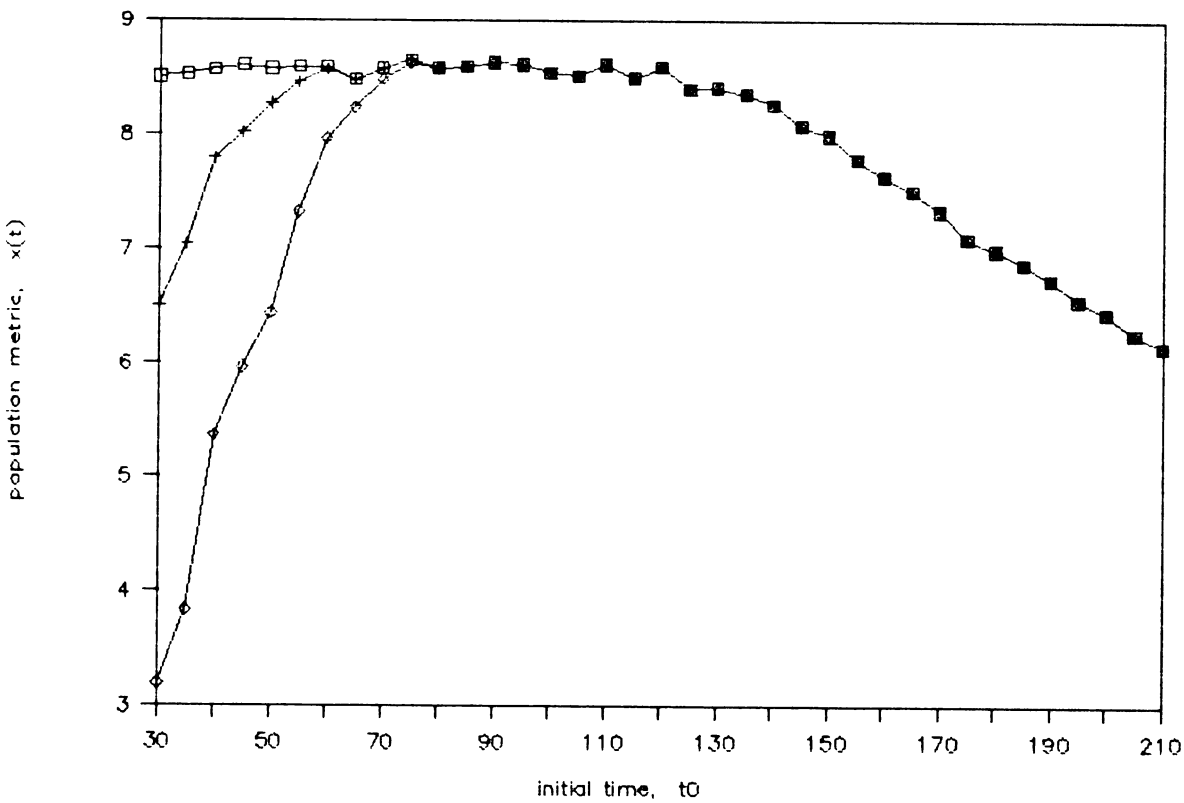


Figure 8.25c Comparison predator/prey mean population metrics summarizing Figures 8.21c, 8.22c, 8.24c. $T_{11}=50^{o}$ F.

- ▪ $T_{10}=15^{o}$ F
- \+ $T_{10}=20^{o}$ F
- ♦ $T_{10}=25^{o}$ F

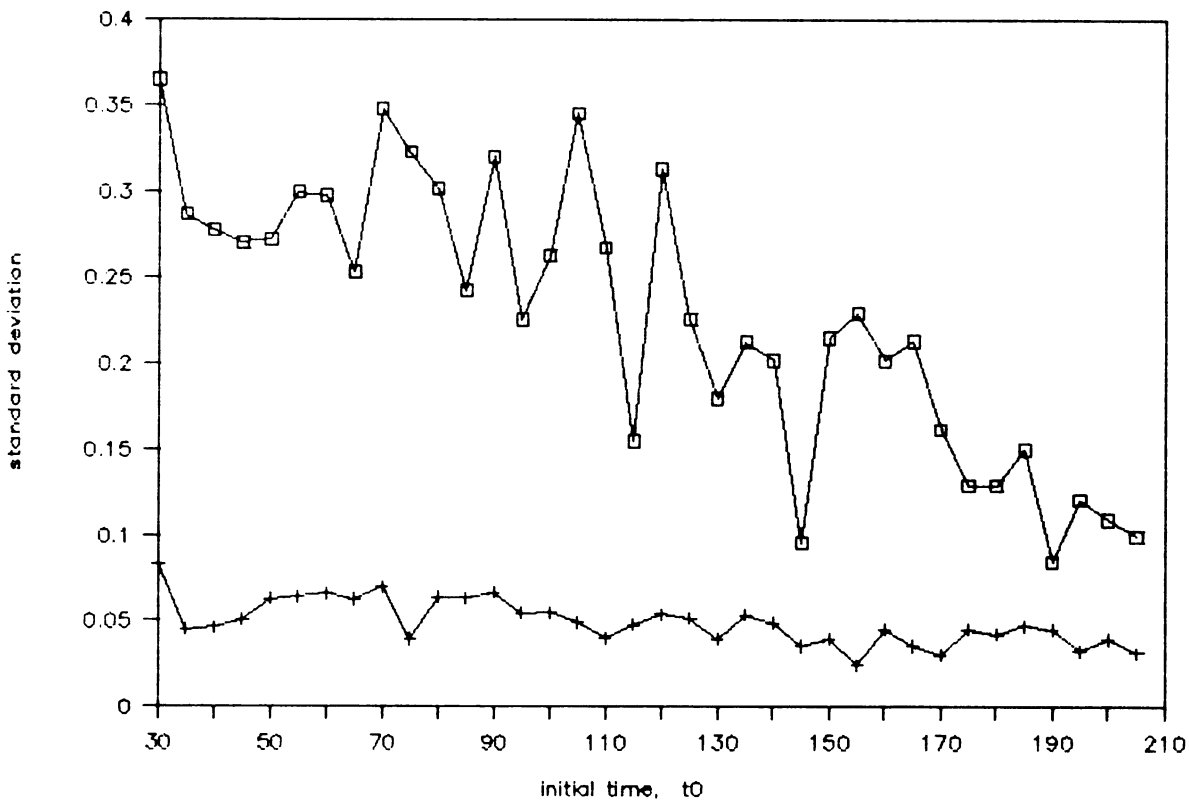


Figure 8.26a Harvest predator/prey variation in population metrics as a function of introduction time with a controlled predator. $T_{10}=15^{o}$, $T_{11}=40^{o}$ F.

▪ prey

\+ predator

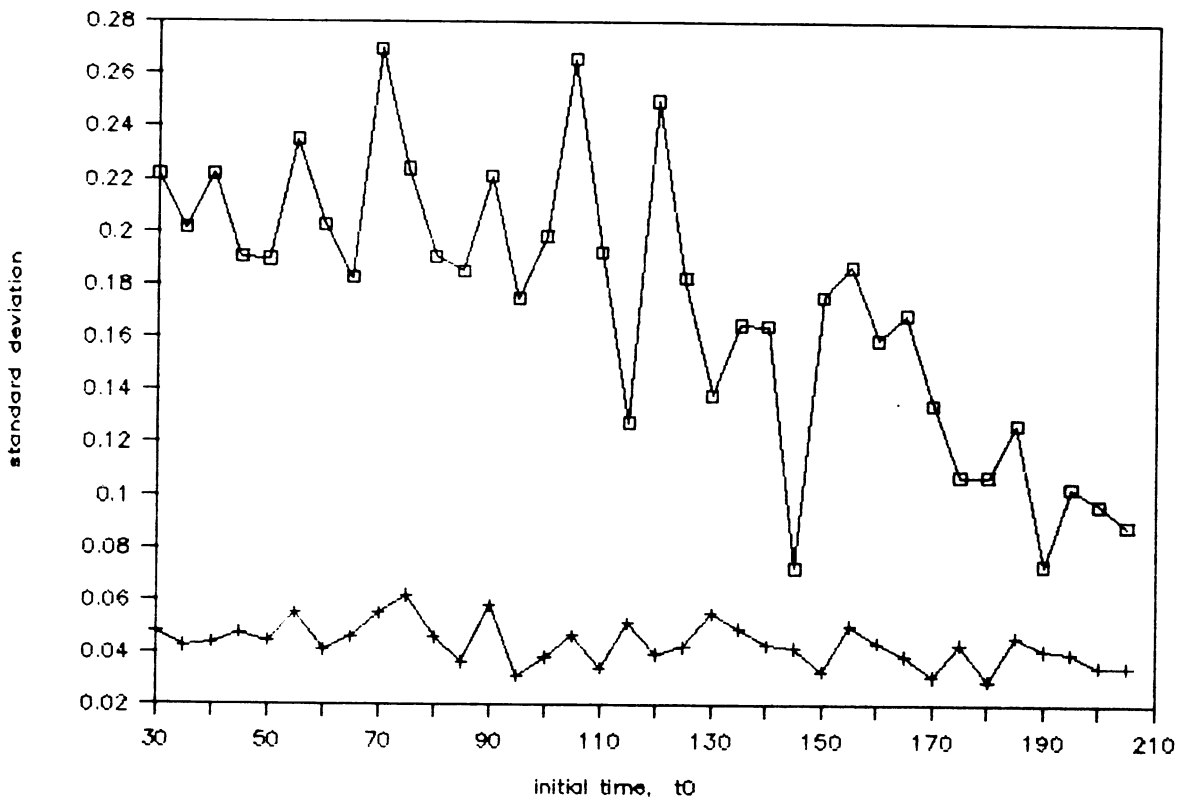


Figure 8.26b Harvest predator/prey variation in population metrics as a function of introduction time with a controlled predator. $T_{10}=15^{o}$, $T_{11}=45^{o}$ F.

* prey

\+ predator

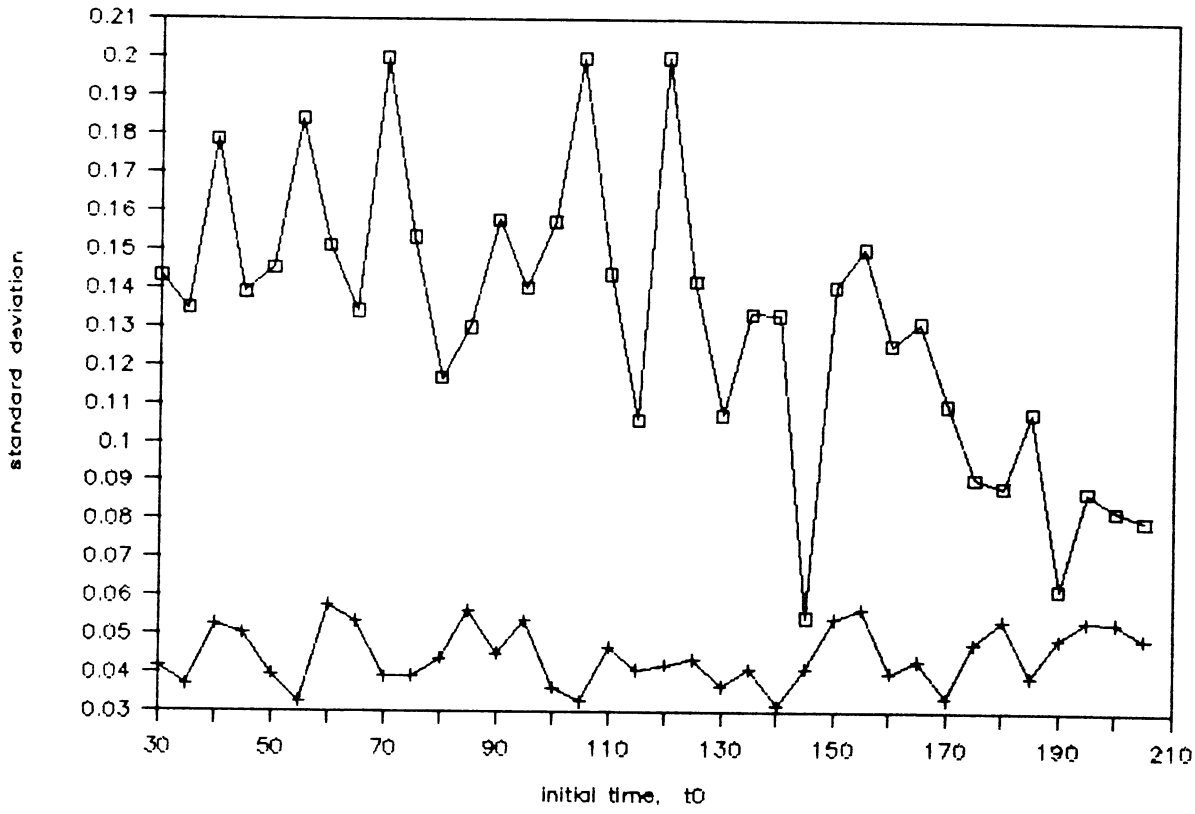


Figure 8.26c Harvest predator/prey variation in population metrics as a function of introduction time with a controlled predator. $T_{10}=15^{o}$, $T_{11}=50^{o}$ F.

▪ prey

\+ predator

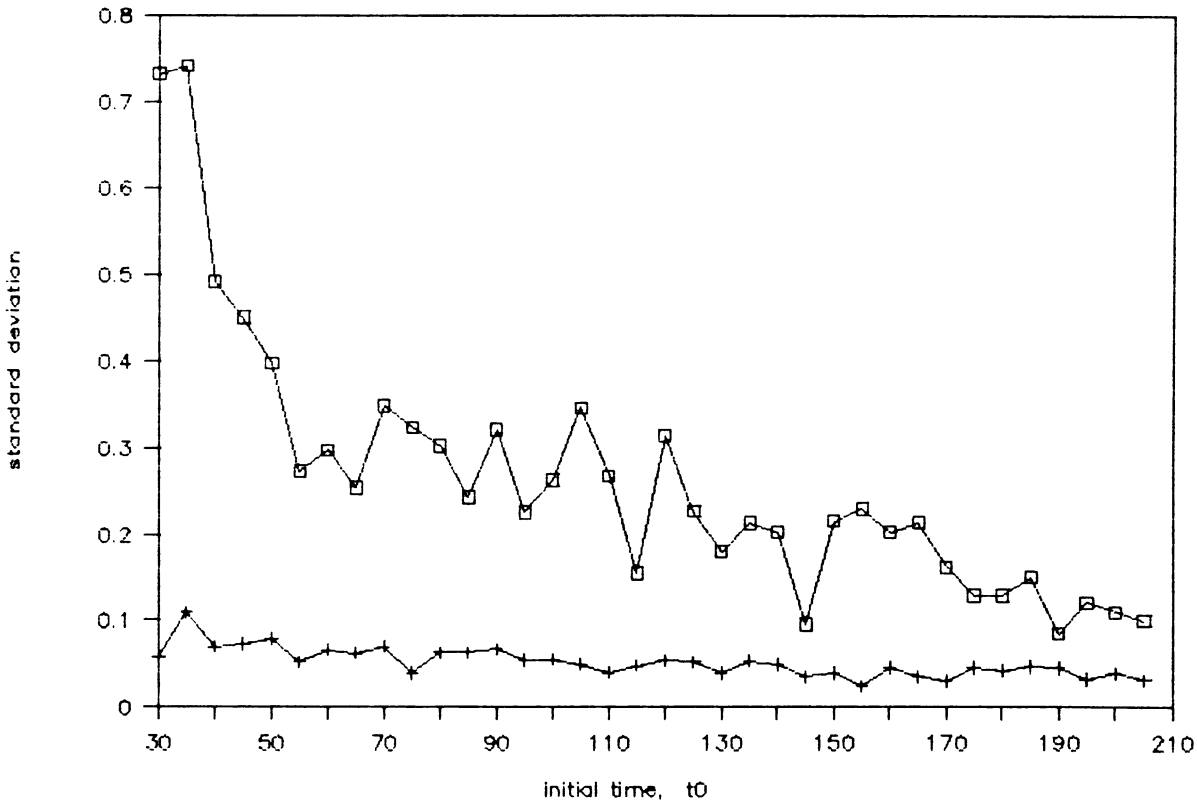


Figure 8.27a Harvest predator/prey variation in population metrics as a function of introduction time with a controlled predator. $T_{10}=20^{o}$, $T_{11}=40^{o}$ F.

▪ prey

\+ predator

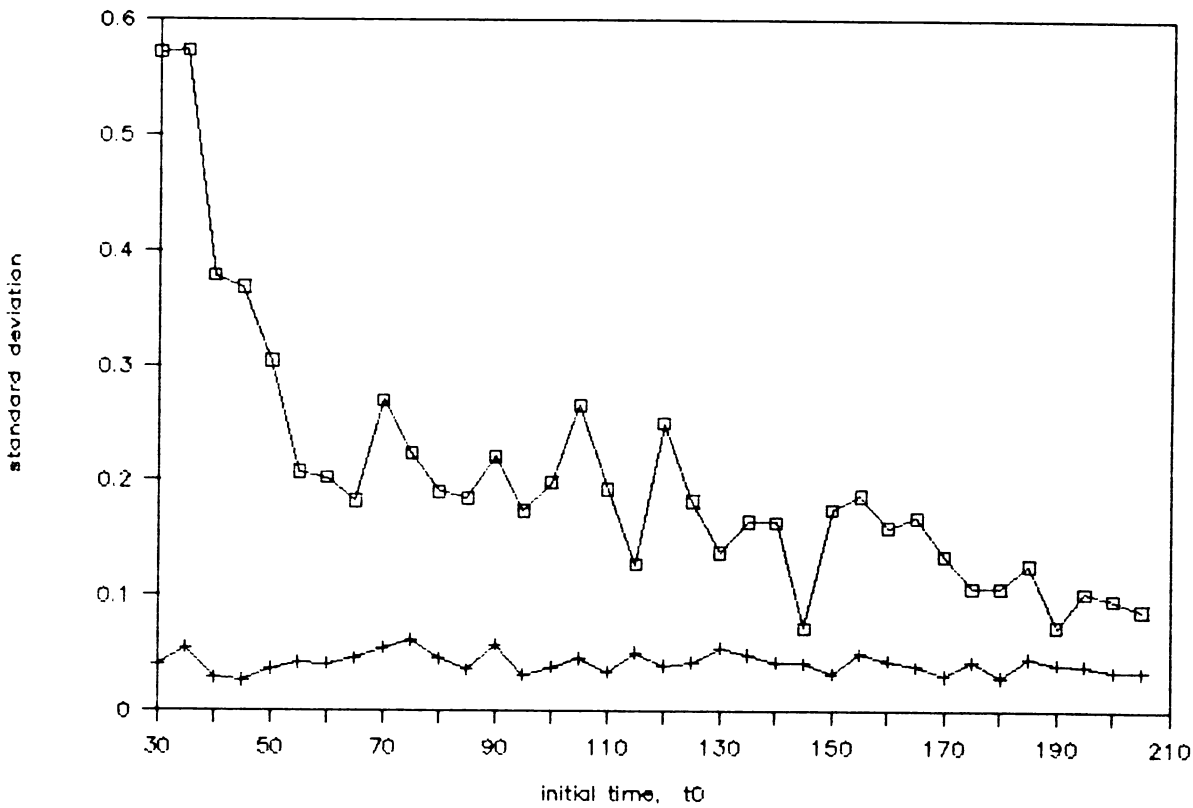


Figure 8.27b Harvest predator/prey variation in population metrics as a function of introduction time with a controlled predator. $T_{10}=20^{o}$, $T_{11}=45^{o}$ F.

* prey

\+ predator

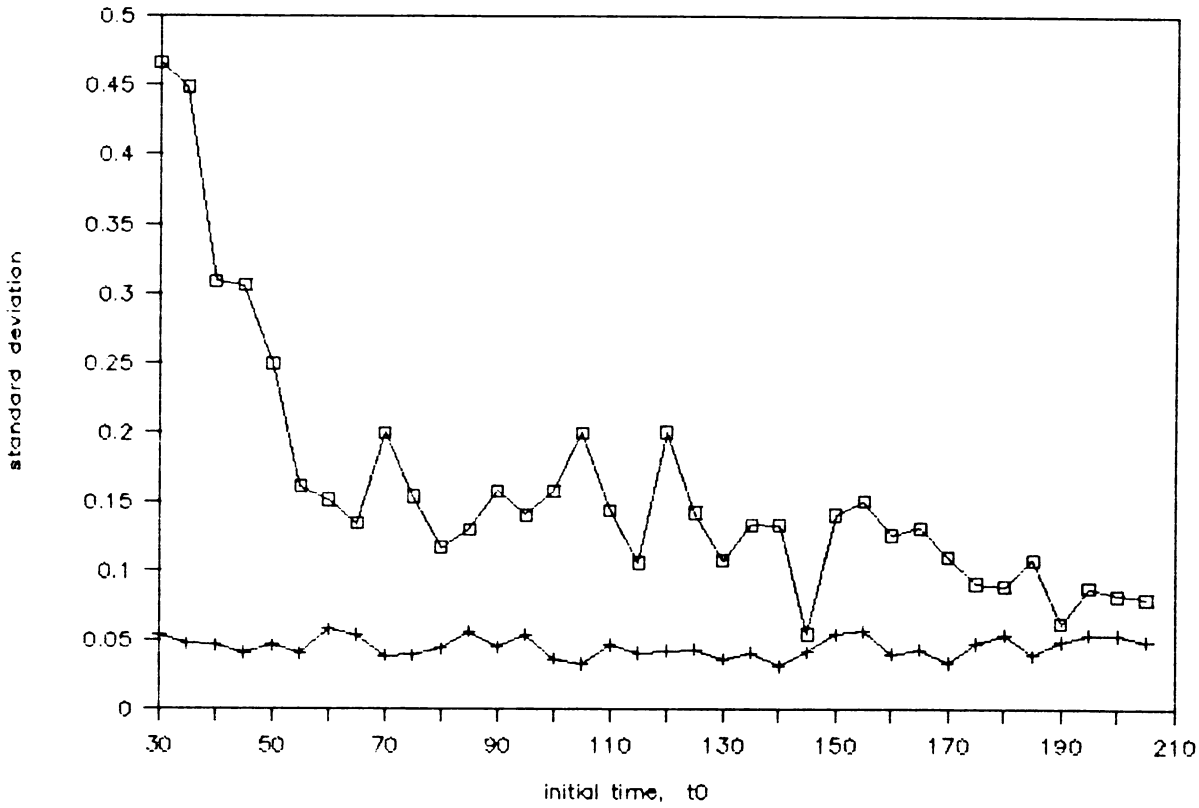


Figure 8.27c Harvest predator/prey variation in population metrics as a function of introduction time with a controlled predator. $T_{10}=20^{o}$, $T_{11}=50^{o}$ F.

▪ prey

\+ predator

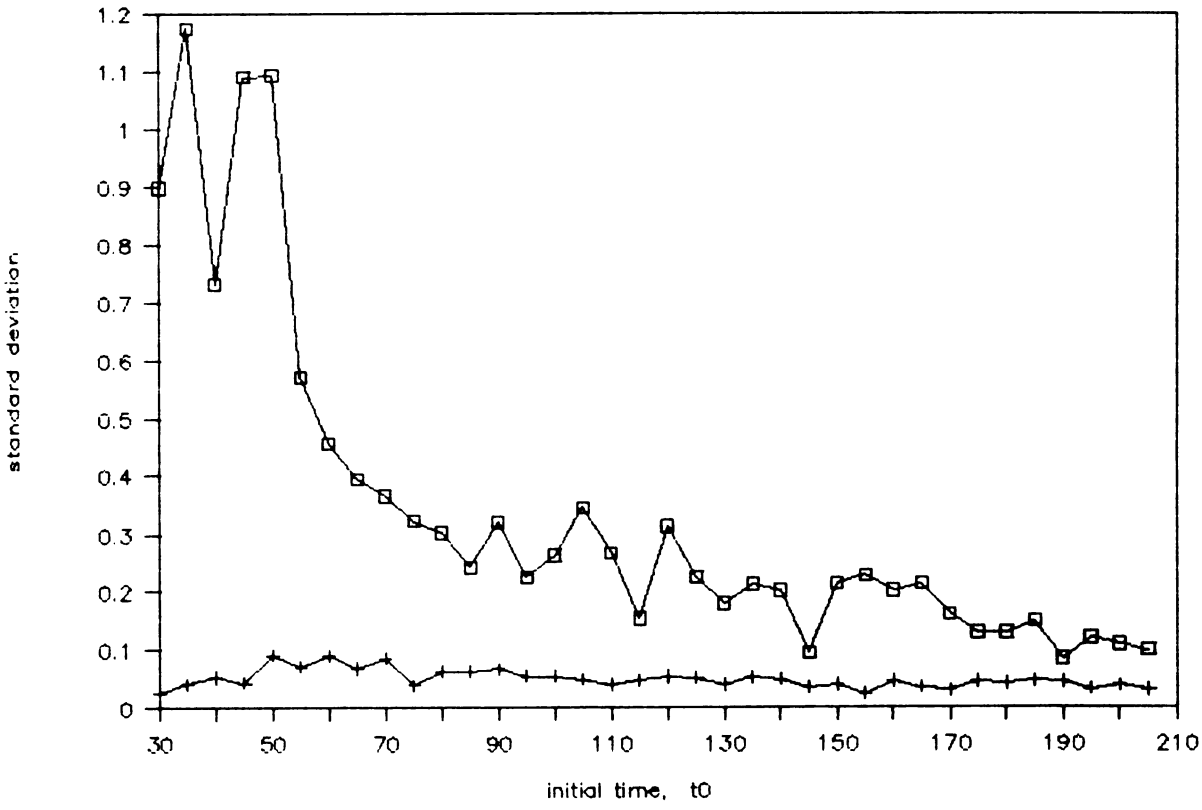


Figure 8.28a Harvest predator/prey variation in population metrics as a function of introduction time with a controlled predator. $T_{10}=25^{o}$, $T_{11}=40^{o}$ F.

* prey

\+ predator

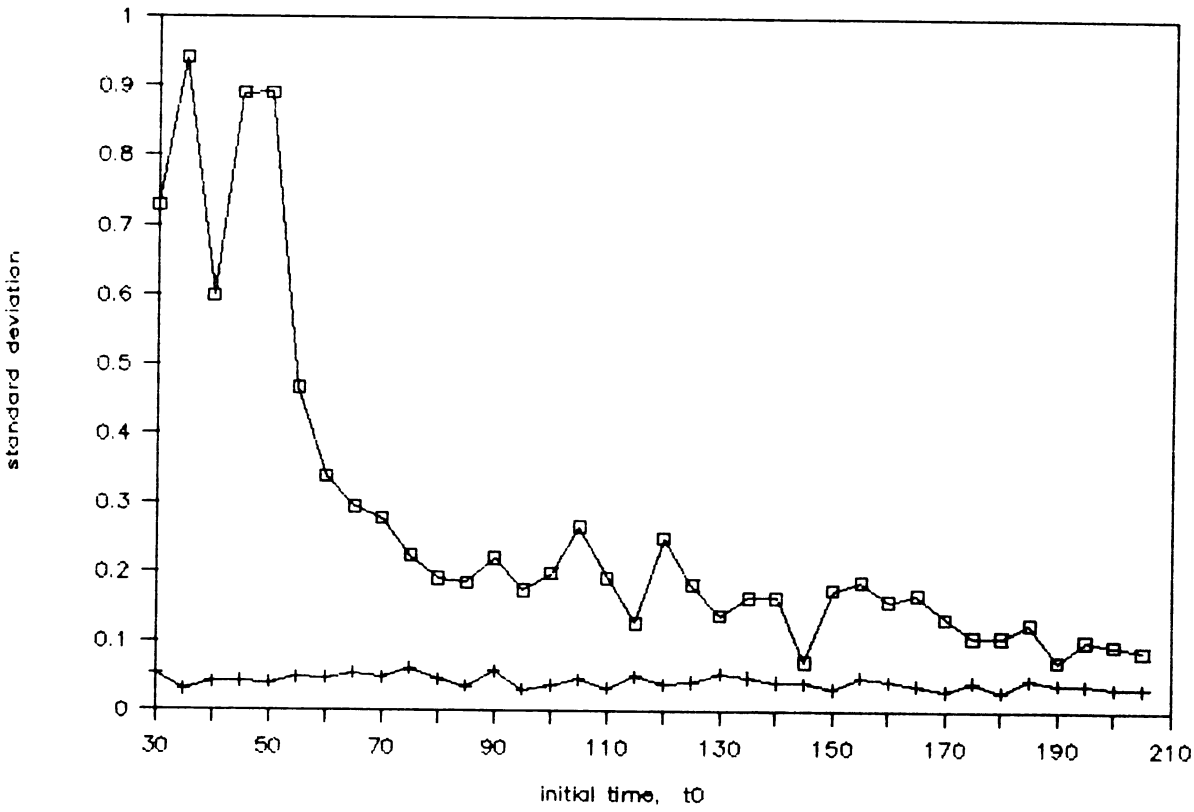


Figure 8.28b Harvest predator/prey variation in population metrics as a function of introduction time with a controlled predator. $T_{10}=25^{o}$, $T_{11}=45^{o}$ F.

• prey

\+ predator

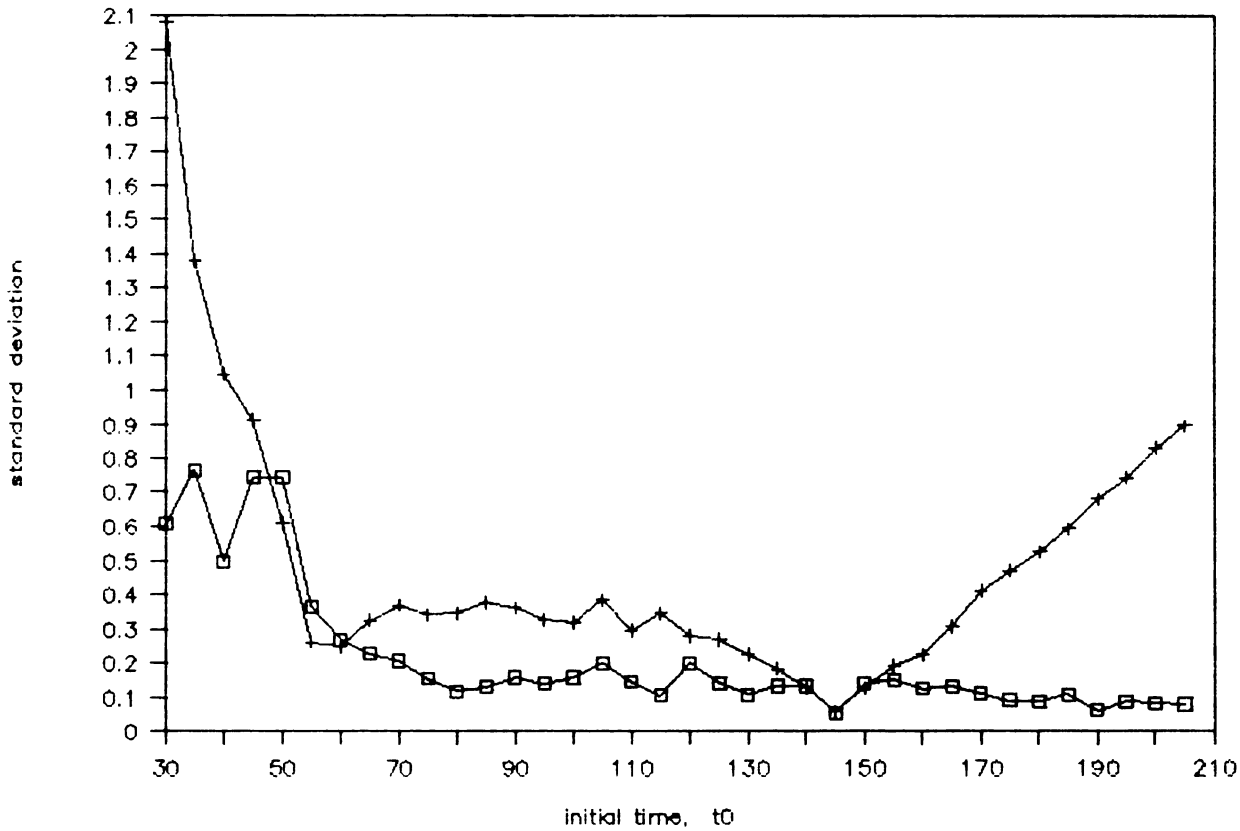


Figure 8.28c Harvest predator/prey variation in population metrics as a function of introduction time with a controlled predator. $T_{10}=25^{o}$, $T_{11}=50^{o}$ F.

▪ prey

\+ predator

CONCLUSION

The purpose of this chapter has been to define a set of control problems useful in population modeling and define practical means for their implementation. These include passive control systems, where the introduction time of a prey species is the variable and active control systems where fractionating and application protocols are variable. In the active case, either open loop or closed loop solutions were considered.

This list is hardly exhaustive! However, rather than produce detailed compartmental models or list "all" possible control systems, it was decided to simply illustrate the passive control system using a simple classical Lotka-Volterra model. This model is admittedly not the one to use for the agricultural case in point. Crops and predators of those crops go through distinct life stages and this is highly significant. Lotka-Volterra ignores this.

Even so, the analysis gives crude results and the methodology of a stochastic driver was illustrated. Applications toward more detailed models should be quite straight forward.

CHAPTER NINE

CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER STUDY

INTRODUCTION

It has been the goal of this research to establish a systematic means for the study of population dynamics systems that invoke the heat unit hypothesis. In doing this care has been taken to establish a rigorous set of definitions and vocabulary by which suitable problems be described. Traditional approaches to applying the heat unit hypothesis have been analyzed and stated in terms of the new vocabulary and definitions. An attempt has been made to demonstrate the validity of these approaches, or failing this, to produce improvements.

It has been a source of surprise to this author that there are a large number of modeling approaches which do not lend themselves to any generalization. Thus, while others might have established shaky but workable models, there was no hope of imbedding them into larger systems or implementing control structures. In other words, not only was the model aesthetically unpleasing but it was impotent as well! Such models have collectively been called ad hoc by this author.

Hopefully this study has provided the basis and direction upon which to establish better models.

Another surprise in reviewing the literature has been the number of seemingly arbitrary thresholds applied. Even relatively recent authors have assumed single threshold (Type I) models as, say, 50^{o} F for vegetables! No attempt was made to validate this guess, to differentiate between species nor to establish other thresholds. In any case, while the concept of multi-threshold models apparently is appreciated, what to do about it is not really understood. This author is unaware of any work that even remotely reflects systematic work establishing model Types or the appropriate endogenous parameter values. This study has produced a framework, a methodology and a set of algorithms so that this process can proceed.

The last surprise in reviewing past work has been that very few researchers have appreciated the fact that these [terrestrial temperature driven] heat-unit models are fundamentally statistical. Temperature profiles were almost invariably deterministic and if not, the temperature generators were primitive and statistically unsound. This author has seen this basic problem as stochastically driven by temperature and, therefore, all results should be couched in statistical terms.

RESULTS OF THIS STUDY

Each of the three "surprises" listed above have been addressed in this work. In doing so, the underlying motivation for this work has been to establish a well defined framework for the heat-unit problem. Having defined the approach, it is hoped that future research will be done with this perspective. In particular, the following tasks have been performed:

i. A suitable vocabulary and a [mathematically] rigorous set of definitions have been established for the descriptions of heat-unit problems.

ii. Basic relationships between the classical models, how to imbed them into heat-unit systems and the new vocabulary has been established.

iii. Data reduction formulas have been derived for both the deterministic and stochastic cases using HIGH-LOW data in order that an effective temperature profile be established.

iv. A review of numeric methods as applied to the heat-unit models has been conducted. While even more precision is available, within the

constraint of the problem at hand, Eulerian methods are sufficient but curve-fitting methods generally are not.

v. An exhaustive study was made of parameter estimation methods. These include a data independent method for endogenous parameters (primarily of only academic interest), algorithmic methods for exogenous parameters (primarily an application of existing techniques) and an algorithmic method for endogenous parameters (most important). A methodology was put forward for the establishment of endogenous parameter values.

vi. A series of stochastic temperature generators was defined with a sound statistical basis. The statistical properties of these generators and the corresponding results were analyzed.

vii. The principles outlined for heat-unit models were applied to an elementary optimal control problem for illustrative purposes. In particular, this was to establish the definition of a biological activity function and the statistical nature of the results.

There have been other, more personal results, of this study as well. Many researchers have paid lip service to the difficulty of data collection, analysis and retrieval. However, it was surprising to note the number of people who were content to recollect the data rather than simply use those already obtained. Modern database technology and data reduction approaches, many of which were applied in this dissertation, could (and should) be applied with immediate success. Even so, the difficulty of bringing the right data to bear at the right time has led this author to a profound resect for this problem in particular and data based research in general.

RECOMMENDATIONS FOR FURTHER STUDY

One needn't go far into this field until the forks in the road become more than evident. There are so many areas where more research is needed. A number of these have been alluded to throughout this work. However, in conclusion, it is probably also a good idea that several of the more salient be given special note. Although this study has been mathematically oriented, this list will include both applied as well as theoretical suggestions.

i. Theoretical Generalizations

At the outset of this work it was stated that a direct analogy exists between mechanical force and biological activity (recall equations [1.1] and [1.2]). This should be developed and exploited further. Other authors imply this to be the case, but none use the definitions in this work, nor have any developed this analogy to any significant degree. It is the feeling of this author that further development of this will yield insights and progress producing most satisfying results.

Another cue might be taken from the relativity example in chapter one. Since multiple time scale

problems exist in both biological and physical sciences, it should be possible to generalize this concept. The design engineer should rightfully ask, not if, but how can multiple time scale concepts be applied to his world too. This author's experience in computer system design simply reaffirms this. There are a tremendous number of applications where a multiple time basis exists. Indeed, this is almost the fundamental principle of parallel computing systems. A more general theory of multiple time scale based systems just might prove not only useful, but well within an insightful grasp.

ii. Theoretical Extensions.

It was stated at the outset that time varying thresholds would not be studied in that they were somehow "impure". There is apparently no theoretical justification for their existence. Even so, they are used because of their convenience in implementation as compared with highly compartmentalized models. Further work examining the relationship between these two approaches for a general class of systems is needed. This is so because the very large class

of agricultural production models are comprised almost inevitably of homogeneous populations, each individual going through life at the same rate.

Let us define this further. In order to model multistage models "properly", it is necessary to use a compartmental model over n stages, each using a logistic equation, with a suitably imbedded logistic model for each stage. Each stage would also have constant temperature thresholds and growth rates as usual. The conjecture is that an n-stage compartmental model could be made equivalent to a single stage logistic model with time varying thresholds.

In general, we should attempt to show that any multistage model might be reducible to a single stage model with time varying thresholds. To prove this would be a nice piece of mathematics and a genuine utility in practice. This would allow the practitioner to find, from the individual threshold and rate constants, the equivalent model with a single stage and thresholds. To say the least, this later model would be significantly easier to implement, especially as n becomes very large. Perhaps the theory could also be extended to the Lotka-

Volterra cases and a network of generalizations would exist.

iii. Applying the results

This work has been largely theoretical. In order to produce tangible results it will be necessary to validate the methodology for determining endogenous constants and once validated, tabulate results of the biological activity function for each stage of various species. Conceptually this should not be difficult. However, in practice there is a significant amount of data reduction to be performed and discovered. Even so, in this author's view, this task is most worthy since only after a basic set of endogenous models is agreed upon can simulation systems be compared.

In order to validate these models it will probably be required that thresholds for species data from a variety of sources in a variety of environments be compared and hopefully found to be consistent. Indeed, the validation and methodological problems alone are most appropriate research topics. This author is enthusiastic that such studies could build on this dissertation.

iv. Applications

The purpose of model building is usually that of either system design or performance optimization. There is no end to the significant problems to be considered here. It is important to generate insight to the mechanisms fundamentally at work rather than simply striving for realism. There is a great deal of insight yet to be found, both individually and collectively, by looking at problems of the type exemplified in chapter eight. A great deal more experimentation needs to be done on classes of problems with different classes of controls.

Consider first optimal control problems. The basic control types of species introduction time, population fractionating, control protocols and feedback controls need to be explored in a systematic manor. This will require a great many simulations of the type done in chapter eight. With good fortune and perseverance, generalizations and insight will result. In fact, this author's limited experience has led him to believe that there is more insight to be gained from this experimental method using simulation.

System design is the other major application area alluded to in this project. This problem really has two parts: the modelling verification portion and a stability study(In doing this, recall the generalized notion of stability for biological systems). Once the existing system is satisfactorily understood, the results of adding a foreign species need to be studied. In most circumstances the system designer will most likely wish to impose stability on the new system as well.

It may be that the system is fundamentally unstable. In this case even more interesting questions arise. For instance, what kinds of controls (as defined above) would lead to a forced stability? These controls might even be studied in the context of optimizing an index of performance. The list is unlimited. Never the less, studying heat-unit based elementary stability and control problems will yield useful results.

CONCLUSION

The experience of preparing this dissertation has been truly unique. As is no doubt common, this author has spent many hours exploring unreported paths, learning those things he already knew and trying to focus on an attainable goal that was also significant.

From the outset it was decided that any results should be easily accessible to the biological system designer. Since these people are the ultimate audience for this work, each result was written explicitly in algorithmic form so as to facilitate ease in implementation. At the same time, no grandiose black-box type software was produced. This was done purposely as it is felt that the principles and their applicability must be understood rather than simply used. Since the majority of this audience will probably not have advanced mathematics or systems background, the text has been kept as rudimentary as possible.

It is worth noting that every program, graph and for that matter the text itself was produced on only a personal computer (IBM PC, 512 K, dual floppy drives). Any mainframe implementation was considered off limits and the software was only of the common off-the-shelf variety (GWBASIC, Turbo Pascal, Lotus 123, dBase III and Wordperfect).

Certainly there have been 24 hour simulation runs and much time has been spent on the most menial of clerical tasks. However, it epitomizes a project in self sufficiency! Even so, every algorithm is easily implemented and every procedure well within the grasp of any competent programmer. The tools used are only those most familiar.

Although this has been made to sound so very easy, the experience has proved the old maxim: Research is 20% inspiration, 30% perspiration and 50% organization. I encourage others to take this path of academic self sufficiency, so be it considerably more demanding. The lessons learned in tool usage and research organization are well worth the additional expenditure in time and effort.

LIST OF REFERENCES

1. Arnold, C. Y. 1959. The determination and significance of the base temperature in a linear heat unit system. Proceedings of the American Society of Horticultural Scientists. 74:430-445.

2. Baskerville, G. L. and P. Emin. 1969. Rapid estimation of heat accumulation from maximum and minimum temperature. Ecology 50:514-517.

3. Bazaraa, M. S. and C. M. Shetty. 1979. Nonlinear Programming. John Wiley & Sons.

4. Boyne, J. V., G. C. Rock and R. E. Stinner. 1985. Temperature dependent models for simulating nondiapause development in Platynota-Idaeusalis Lepidoptera Tortricidae in North Carolina USA. Environment Entomology. 14:785-789.

5. Bulgren, W. G. 1982. Discrete System Simulation. Prentice Hall.

6. Castonguay, Y., J. Boisvert and P. A. Dube. 1984. Comparison of two statistical methods applied to development of phenoclimatic models. Agriculture Forest Meteorology. 34:273-288.

7. Coakley, S. M., W. S. Boyd and R. F. Line. 1982. Statistical models for predicting stripe rust on winter wheat Triticum-Aestivum in the pacific northwest USA. Phytopathology. 72:1539-1542.

8. Daubenmire, R. F. 1947. Plants and Environment - A Textbook of Plant Autoecology. John Wiley & Sons.

9. Davis H. T. 1962. Introduction to Nonlinear Differential and Integral Equations. Dover.

10. Dennis, B., W. P. Kemp and R. C. Beckwith. 1986. Stochastic Model of insect phenology estimation and testing. Environmental Entomology. 15:540-546.

11. Flinn, P. W., A. J. Taylor and A. A. Hower. 1986. Predictive model for the population dynamics of Potato Leafhopper Empoasca-Fabae Homoptera Cicadellidae on alfalfa. Environmental Entomology. 15:898-904.

12. Fritsch, K. 1961. Thermische Constanten fur die Bluthe und Fruchtreife nvon 889 Pflanzenarten, abgeleitet aus zehnjahrigen Beobachtungen im K. K. Botanischen Garten zu Xein. Wien Denkschriften. 25:711-718.

13. Gutierrez, A. P., J. U. Baumgaertner and C. G. Summers. 1984. Multitrophic models of predator prey energetics. Canadian Entomology. 116:923-964.

14. Gutierrez, A. P., D. W. Williams and H. Kido. 1985. A Model of grape Vitis-Vinifera growth and development: the mathematical structure and biological considerations. Crop Science. 25:721-728.

15. Hesketh, J. D. 1980. [effects of temperature on growth]. Crop Physiology. p.85, 123.

16. Hogg, D. B. 1985. Potato Leafhopper Empoasca-Fabae Homoptera Cicadellidae immature development life tables and population dynamics under fluctuation temperature regimes. Environmental Entomology. 14:349-355.

17. Holmes, R. M. and G. W. Roberston. 1959. Heat units and crop growth. Canada Department of Agriculture Publication number 1042. Ottawa, Ontario.

18. Hoppensteadt, F. C. 1975. Mathematical Theories of Populations Demographics, Genetics and Epidemics. Society for Industrial and Applied Mathematics.

19. Kemp, W. P., B. Dennis and R. C. Beckwith. 1986. Stochastic phenology model for the Western Spruce Budworm Choristoneura-Occidentalis Lepidoptera Tortricidae. Environmental Entomology. 15:547-554.

20. Kemp, W. P. and J. A. Onsager. 1986. Rangeland Grasshoppers Orthoptera Acrididae modeling phenology of natural populations of six species. Environmental Entomology. 15: 924-930.

21. Kuester, J. L. and J. H. Mize. 1973. Optimization Techniques with Fortran. McGraw-Hill Book Company.

22. Lasack, P. M., W. C. Bailey and L. P. Pedigo. 1987. Assessment of Stalk Borer Lepidoptera Nictuidae population dynamics by using logistic development curves and partial life tables. Environmental Entomology, 16:296-303.

23. Lehenhauser, P. A. 1914. Growth of maise seedlings in relation to temperature. Physiol. Res. 1:247-288.

24. Leslie, P. H. 1945. On the use of matrices in certain population mathematics. Biometrika. 33:183-212.

25. Lindsey, A. A. and J. E. Newman. 1956. Use of official weather data in spring time - temperature analysis of an Indiana phenological record. Ecology. 37:812-823.

26. Linvill, E. E., R. F. Dale and H. F. Hodges. 1978. Solar radiation weighting for weather and corn growth models. Agronomy Journal. 1978:257-263.

27. Lotka, A. J. 1956. Elements of Mathematical Biology. Dover.

28. Lund, C. T. and D. C. Herne. 1980. Relationships between degree-days and computer simulated events within a population of the European Red Mite Panonychus-Ulmi Acarina Tetranychidae. Proceedings of the Entomological Society of Ontario. 111:33-38.

29. Manetsch, T. J. 1976. Time-varying distributed delays and their use in aggregative models of large systems. IEEE Transactions on Systems, Man and Cybernetics. 6:547-553.

30. Manetsch, T. J. 1980. Bilateral distributed delays and the use in modeling classes of distributed parameter processes. IEEE Transactions on Systems, Man and Cybernetics. 10:61-67.

31. Manetsch, T. J. 1966. Transfer function representation of the aggregate behavior of a class of economic processes. IEEE Transactions on Automatic Control. 11:693-698.

32. McKendrick, A. G. 1926. The application of mathematics to medical problems. Proceedings of the Edinburgh Mathematical Society. 44:98-130.

33. Nisbet, R. M. and W. S. C. Gurney. 1985. Modelling Fluctuating Populations. Wiley-Interscience.

34. Nuttonson, M. Y. 1948. Some preliminary observations of phenological data as a tool in the study of photoperiodic and thermal requirements of various plant material. Chronica Botanica, Waltham, Massachusetts.

35. Osawa, A., C. A. Shoemaker and J. R. Stedinger. 1983. A stochastic model of balsam fir abies-balsamea bud phenology utilizing maximum likelihood parameter estimation. Forestry Science. 29:478-490.

36. Plant, R. E. and L. T. Wilson. 1986. Models for age structured populations with distributed maturation rates. Journal of Mathematical Biology. 23:247-262.

37. Press, W. H., B. P. Flannery, S. A. Teukolsky and W. T. Vetterling. 1986. Numerical Recipes - The Art of Scientific Computing. Cambridge University Press.

38. Raworth, D. A. 1984. Population dynamics of the Cabbage Aphid Brevicoryne-Brassicae Homoptera Aphididae at Vancouver British-Columbia Canada 2. Development Fecundity and Longevity. Canadian Entomology. 116:871-878.

39. Reaumur, R. A. F. 1735. Observation du thermometre, faites a Paris pendant Pannee 1735, comparees avee celles qui ont ete faites sous la ligne, a l'Isle de France, a Alger et en quelques-unes de nos isles de l'Amerique. M'm. Academy des Science. Paris. p. 545.

40. Ring, D. R., V. R. Calcote, J. N. Cooper, R. Olszak, J. E. Begnaud, T. W. Fuchs, C. W. Neevb, R. D. Parker, J. L. Henson, et-al. 1983. Generalization and application of a degree-day model prediction pecan nut Casebearer Acrobasis-Nuxvorella Lepidoptera Pyralidae activity. Journal of Economic Entomology. 76:831-835.

41. Schreiber, M. M., G. E. Miles, D. A. Holt and R. J. Bula. 1978. Sensitivity analysis of SIMED: a crop growth model. Agronomy Journal. 70:105-108.

42. Shugart, H. H., W. R. Emanuel, D. C. West and D. L. Deangelis. 1980. Environmental gradients in a simulation model of a Beech Fagus-Grandifolia Yellow-Poplar Liriodendron-Tulipifera stand. Mathematical Bioscience. 50:163-170.

43. Slobodkin, L. B. 1953. An algebra of population growth. Ecology 34:513-519.

44. Thornthwaite, C. W. 1952. Temperature relation to time of maturity of vegetable crops. Proceedings of the 78th New Jersey State Horticultural Society. Johns Hopkins University.

45. Van Sickle, J. 1977. Attrition in distributed delay models. IEEE Transactions on Systems, Man and Cybernetics. 7:635-638.

46. Waggoner, P. E. 1984. The hatching of Gypsy Moth Lymantria-Dispar eggs: A Phenological Model. Agricultural Forest Meteorology. 33:53-66.

47. Wang, J. Y. 1958a. An Evaluation of Some Techniques in Agrometeorology. Ph.D. dissertation, University of Wisconsin.

48. Wang, J. Y. 1958b. More accurate predictions of corn maturity data. Food Packer. 39:36-37.

49. Wang, J. Y. 1960. A critique of the heat unit approach to plant response studies. Ecology. 41:785-790.

50. Werner, P. A. and H. Caswell. 1977. Population growth rates and age versus stage-distribution models for teasel (Dipsacus sylvestris Huds.). Ecology. 58:1103-1111.

51. Wilson, L. T. and W. W. Barnett. 1983. Degree-days: an aid in crop and pest management. California Agriculture. 37:4-7.

52. Wolf, S., J. Rudich, A. Marani and Y. Rekah. 1986. Predicting harvesting date of processing tomatoes by a simulation model. Journal of American Society of Horticultural Scientists. 111:11-16.

GENERAL REFERENCES

1. Aguirre, L. A. and M. K. Harris. 1986. Predicting biological events of the pecan nut casebearer using a degree-day model in Coahuila Mexico. Southwest Entomology. 11:263-268.

2. Albrechyt, F., H. Gatzke, A. Haddad and N. Wax. 1974. The dynamics of two interacting populations. Journal of Mathematical Analysis Applications. 46:658-670.

3. Ayala, F. J., M. E. Gilpin and J. G. Ehrenfeld. 1973. Competition between species: theoretical models and experimental tests. Theoretical Population Biology. 4:331-356.

4. Beckwith, R. C. and W. P. Kemp. 1984. Shoot growth models for Douglas-Fir Pseudotsuga-Menziesii-Var-Glauca and Grank-Fir Abies-Grandis. Forestry Science. 30:743-746.

5. Blackburn, W. J. 1985. Use of models in apple pest management. Canadian Journal of Plant Science. 65:818.

6. Blythe, S. P. 1984. The dynamics of population models with distributed maturation periods. Theoretical Population Biology. 25:289-311.

7. Burden, R. L., J. D. Faires and A. C. Reynolds. 1978. Numerical Analysis (second edition). Prindle, Weber & Schmidt.

8. Butler, G. S. 1980. Modeling and Differential Equations in Biology, T. Burton, ed. Marcel Dekker, New York.

9. Caswell, E. P. and I. J. Thomason. 1986. Physiological time based model for egg production by Heterodera-Schachtii. Journal of Nematology. 18:603.

10. Christiansnen, F. B. and T. M. Fenchel. 1977. Theories of Populations in Biological Communities. Springer-Verlag, Heidelberg.

11. D'Azzo, J. J. and C. H. Houpis. 1981. Linear Control System Analysis and Design (second edition). McGraw Hill Book Company.

12. Dintenfass, L. P., D. P. Bartell and M. A. Scott. 1987. Predicting resurgence of western flower Thrips Thysanoptera Thripidae on onions after insecticide application in the texas high plains. Journal of Economic Entomology. 80:502-506.

13. Dym, C. L. and E. S. Ivey. 1980. Principles of Mathematical Modeling. Academic Press.

14. Eskridge, K. M. and E. J. Stevens. 1987. Growth curve analysis of temperature-dependent phenology models. Agronomy Journal. 79:291-297.

15. Even, S. 1979. Graph Algorithms. Computer Science Press.

16. Francl, L. J., L. V. Madden, R. C. Rowe and R. M. Riedel. 1986. Modeling yield loss due to potato early dying disease. Journal of Nematology. 18:609.

17. Goodell, P. B. and H. Ferris. 1986. Validation of an overwintering simulation for Meloidogyne-Incognita in cotton fields in the San-Joaquin Valley of California USA. Journal of Nematology. 18:610.

18. Goodenough, J. L., A. W. Hartstack and E. G. King. 1983. Developmental models for Trichogramma-Pretiosum Hymenoptera Trichogrammatidae reared on four hosts. Journal of Economic Entomology. 76:1095-1102.

19. Gutierrez, A. P., J. E. Devay, G. S. Pullman and G. E. Friebertshauser. 1983. A model of Verticillium Wilt in relation to Cotton Gossypium-Hirsutum Cultivar Acala-s-2 growth and development. Phytopathology. 73:89-95.

20. Hamming, R. W. 1973. Numerical Methods for Scientists and Engineers (second edition). Dover.

21. Hardman, J. M., W. A. Charnetski and M. K. Muderji. 1985. A model simulating grasshopper damage to wheat planted in infested stubble. Journal of Applied Ecology. 22:373-394.

22. Hochberg, M. E., J. Pickering and W. M. Getz. 1986. Evaluation of phenology models using field data case study for the Pea Aphic Acyrthosiphon-Pisum and the Blue Alfalfa Aphid Acyrthosiphon-Kondoi Homoptera Aphididae. Environmental Entomology. 15:227-231.

23. Hoppensteadt, F. C. 1982. Mathematical Methods of Population Biology. Cambridge University Press, Cambridge.

24. Isaacson, E. and H. B. Keller. 1966. Analysis of Numerical Methods. John Wiley & Sons.

25. Jacquez, J. A. 1972. Compartmental Analysis in Biology and Medicine. Elsevier.

26. May, R. M. 1973. Stability and Complexity in Model Ecosystems. Princeton University Press, Princeton.

27. McNichols, C. W. 1984. Microcomputer-Based Data Analysis for the IBM PC. Reston Publishing.

28. Mikkelsen, S. A. 1981. Predicting the date of harvest of vining peas by means of growing degree-day models. ACTA Horticulture. 122:211-221.

29. Morse, J. G. and K. R. Laking. 1987. A degree-day model for Fuller Rose beetle. Citrograph. 72.

30. Noling, J. W. and H. Ferris. 1987. Nematode degree-days: a density-time model for relating epidemiology and crop losses in perennials. Journal of Nematology. 19:108-118.

31. Pinkerton, J. N., G. S. Santo and H. Mojtahedi. 1986. Population dynamics of Meloidogyne-chitwoodi in relation to Russet-burbank potato tuber penetration. Journal of Nematology. 18:627.

32. Raworth, D. A. 1984. Dynamics of the Cabbage Aphid Brevicoryne-Brassicae Homoptera Aphididae at Vancouver British-Columbia Canada 4. Predation by Aphidoletes-Aphidimyza Diptera Cecidomyiidae. Canadian Entomology. 116:889-894.

33. Reader, R. J. 1983. Modeling geographic variation in the timing of shoot extension by Ericaceous Shrubs. Canadian Journal of Botany. 61:2032-2037.

34. Rice, J. R. 1983. Numerical Methods, Software and Analysis (IMSL reference version). Mcgraw-Hill Book Company.

35. Richardson, J. C., C. D. Jorgensen and B. A. Croft. 1982. Embryogenesis of the Codling Moth Laspeyresia-Pomonella use in validating phenology models. Annual Entomological Society of America. 75:201-209.

36. Silver, S. B. and C. W. Basham. 1980. PSIM: a crop scheduling simulation. Hortoscience. 15:411.

37. Simon, W. 1972. Mathematical Techniques for Biology and Medicine. Dover.

38. Tenenbaum, M. and H. Pollard. 1963. Ordinary Differential Equations. Dover.

39. Tricomi, F. G. 1897. Integral Equations. Dover.

40. Valentine, H. T. 1983. Bud break and leaf growth functions for modeling Herbivory in some Gypsy Moth Lymantria-Dispar Hosts. Forestry Science. 29:607-617.

41. Vidyasagar, M. 1978. Nonlinear Systems Analysis. Prentice Hall.

42. Waltman, P. 1983. Competition Models in Population Biology. Society for Industrial and Applied Mathematics.