Reliable Modeling and Optimization for Chemical Engineering Applications: Interval Analysis Approach

Youdong Lin, C. Ryan Gwaltney and Mark A. Stadtherr

Department of Chemical and Biomolecular Engineering, University of Notre Dame

Notre Dame, IN, USA

NSF Workshop on Reliable Engineering Computing, Savannah, GA, September 15–17, 2004

Outline

- Motivation – Reliability in Computing
- Problem Solving Methodology
- Applications in Chemical Engineering
  - Overview
  - Parameter estimation in modeling of vapor-liquid equilibrium (VLE)
  - Nonlinear dynamics – ecological modeling
  - Molecular modeling – transition state analysis
- Concluding Remarks

Motivation

- Many applications in chemical engineering deal with nonlinear models of complex physical phenomena, on scales from macroscopic to molecular
- A common problem is the need to solve a nonlinear equation systems in which the variables are constrained physically within upper and lower bounds; that is, to solve:

$$\mathbf{f}(\mathbf{x}) = \mathbf{0}$$

$$\mathbf{x}^L \leq \mathbf{x} \leq \mathbf{x}^U$$

- These problems may:
  - Have multiple solutions – Have all been found?
  - Have no solution – Can this be verified?
  - Be difficult to converge to any solution using standard methods

Motivation (Cont'd)

- Another common problem is the need to globally minimize a nonlinear function, subject to nonlinear equality and/or inequality constraints:

$$\min_{\mathbf{x}} \phi(\mathbf{x})$$

subject to

$$\mathbf{h}(\mathbf{x}) = \mathbf{0}$$

$$\mathbf{g}(\mathbf{x}) \geq \mathbf{0}$$

$$\mathbf{x}^L \leq \mathbf{x} \leq \mathbf{x}^U$$

- These problems may:
  - Have multiple local minima – Has the global minimum been found?
  - Require finding all local minima or stationary points – Have all been found?
  - Have no solution (infeasible NLP) – Can this be verified?
  - Be difficult to converge to any local minima using standard methods

Interval Analysis

- One approach for dealing with these issues is interval analysis
- Interval analysis can
  - Provide the tools needed to solve modeling and optimization problems with complete certainty
  - Provide problem-solving reliability not available when using standard local methods
  - Deal automatically with rounding error, thus providing both mathematical and computational guarantees

Interval Methodology

- Core methodology is Interval Newton/Generalized Bisection (IN/GB)
  - Given a system of equations to solve, an initial interval (bounds on all variables), and a solution tolerance:
  - IN/GB can find (enclose) with mathematical and computational certainty either all solutions or determine that no solutions exist
  - IN/GB can also be extended and employed as a deterministic approach for global optimization problems
- A general purpose approach; in general requires no simplifying assumptions or problem reformulations
- No strong assumptions about functions need to be made

Interval Methodology (Cont'd)

Problem: Solve $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ for all roots in interval $\mathbf{X}^{(0)}$

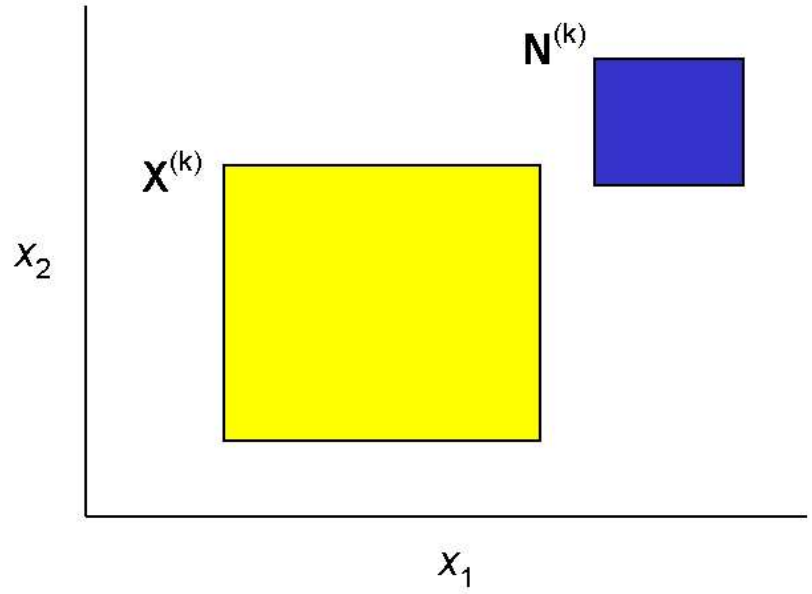
Basic iteration scheme: For a particular subinterval (box), $\mathbf{X}^{(k)}$, perform root inclusion test:

- (Range Test) Compute the interval extension $\mathbf{F}(\mathbf{X}^{(k)})$ of $\mathbf{f}(\mathbf{x})$ (this provides bounds on the range of $\mathbf{f}(\mathbf{x})$ for $\mathbf{x} \in \mathbf{X}^{(k)}$)
  - If $0 \notin \mathbf{F}(\mathbf{X}^{(k)})$, delete the box. Otherwise,
- (Interval Newton Test) Compute the *image*, $\mathbf{N}^{(k)}$, of the box by solving the linear interval equation system

$$\mathbf{F}'(\mathbf{X}^{(k)})(\mathbf{N}^{(k)} - \tilde{\mathbf{x}}^{(k)}) = -\mathbf{f}(\tilde{\mathbf{x}}^{(k)})$$

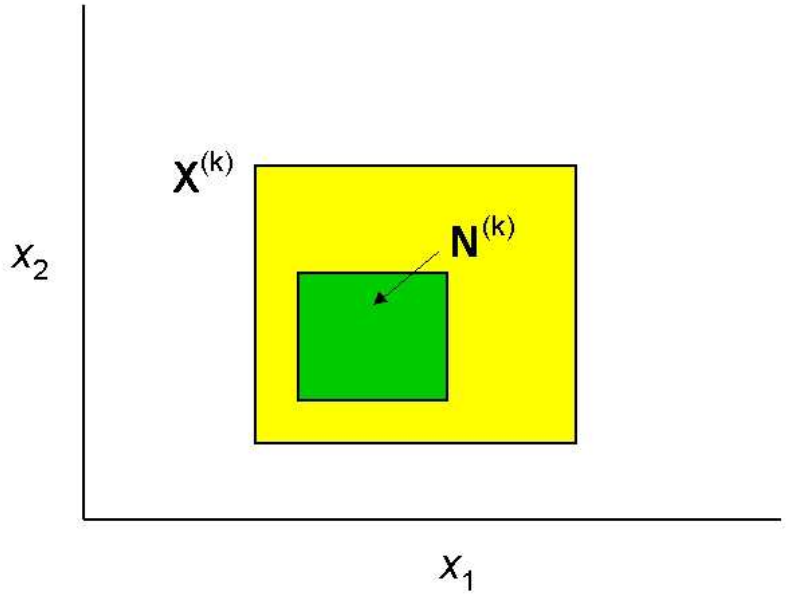
  - $\tilde{\mathbf{x}}^{(k)}$ is some point in $\mathbf{X}^{(k)}$
  - $\mathbf{F}'(\mathbf{X}^{(k)})$ is an interval extension of the Jacobian of $\mathbf{f}(\mathbf{x})$ over the box $\mathbf{X}^{(k)}$

Interval Methodology (Cont'd)



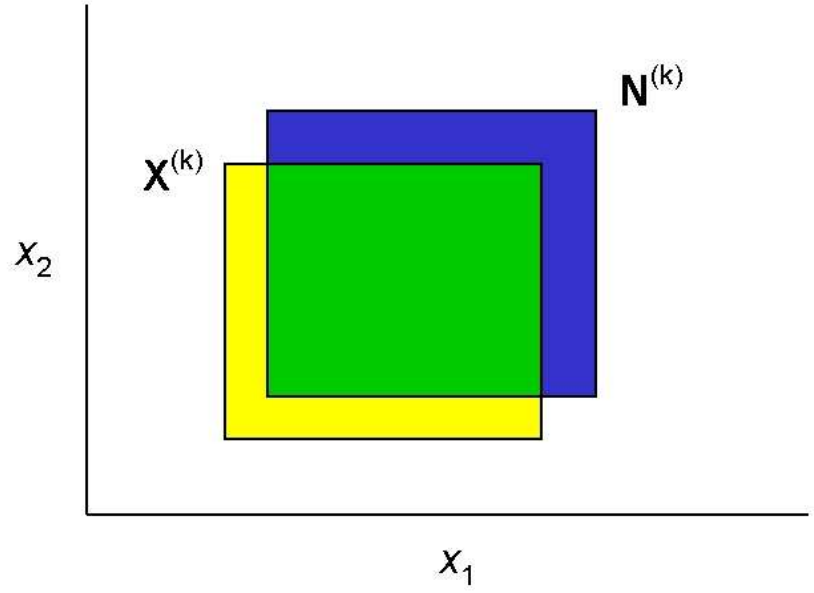
- There is no solution in $\mathbf{X}^{(k)}$

Interval Methodology (Cont'd)



- There is a *unique* solution in $\mathbf{X}^{(k)}$
- This solution is in $\mathbf{N}^{(k)}$
- Additional interval-Newton steps will tightly enclose the solution with quadratic convergence

Interval Methodology (Cont'd)



- Any solutions in $\mathbf{X}^{(k)}$ are in intersection of $\mathbf{X}^{(k)}$ and $\mathbf{N}^{(k)}$
- If intersection is sufficiently small, repeat root inclusion test
- Otherwise, bisect the intersection and apply root inclusion test to each resulting subinterval
- This is a branch-and-prune scheme on a binary tree

Interval Methodology (Cont'd)

- Can be extended to global optimization problems
- For unconstrained problems, solve for stationary points ($\boldsymbol{\nabla}\phi = \mathbf{0}$)
- For constrained problems, solve for KKT or Fritz-John points
- Add an additional pruning condition (objective range test):
  - Compute interval extension of objective function
  - If its lower bound is greater than a known upper bound on the global minimum, prune this subinterval
- This combines IN/GB with a branch-and-bound scheme on a binary tree

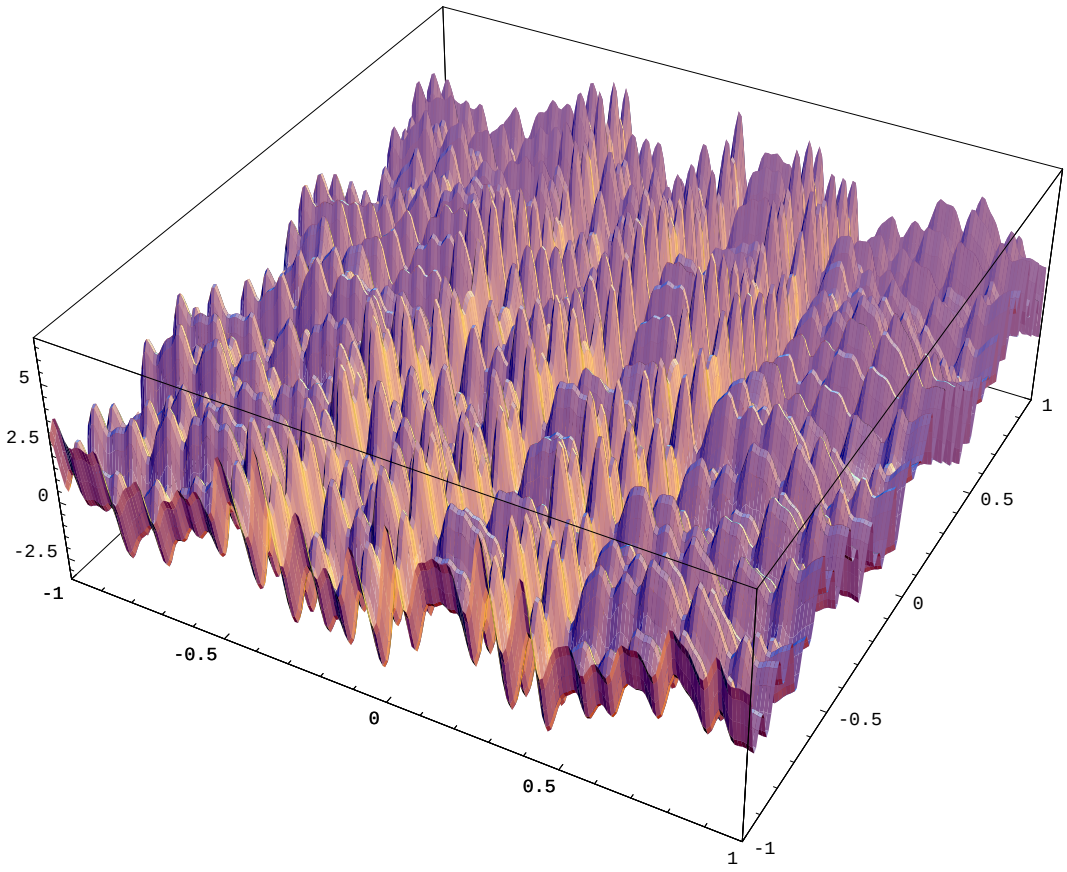
Interval Methodology (Cont'd)

Enhancements to basic methodology:

- Hybrid preconditioning strategy (HP) for solving interval-Newton equation (Gau and Stadtherr, 2002)
- Strategy (RP) for selection of the real point $\tilde{\mathbf{x}}^{(k)}$ in the interval-Newton equation (Gau and Stadtherr, 2002)
- Use of linear programming techniques to solve interval-Newton equation — LISS/LP (Lin and Stadtherr, 2003, 2004)
  - Exact bounds on $\mathbf{N}^{(k)}$ (within roundout)
- Constraint propagation (problem specific)
- Tighten interval extensions using known function properties (problem specific)

Example

- Trefethen (2002) Challenge Problem #4 — Find the Global Minimum



$$f(x,y) = \exp(\sin(50x)) + \sin(60\exp(y)) + \sin(70\sin(x)) + \sin(\sin(80y)) - \sin(10(x+y)) + (x^2+y^2)/4; \qquad x \in [-1,1]; \qquad y \in [-1,1]$$

Example (Cont'd)

- Global minimum is easily found using interval approach

$$x \in [-0.0244030796943751 7, -0.0244030796943751 6]$$

$$y \in [0.21061242715535 57, 0.21061242715535 58]$$

$$f \in [-3.30686864747524 5, -3.30686864747523 2]$$

- CPU time (LISS/LP): 0.16 seconds on SUN Blade 1000 model 1600 workstation

Another Example

- Find the global minimum of the function (Siirola et al., 2002):

$$f(\mathbf{x}) = 100 \prod_{i=1}^{N} \sum_{j=1}^{5} \left(\frac{j^5}{4425} \cos(j + j x_i) \right) + \frac{1}{N} \sum_{i=1}^{N} (x_i - x_{0,i})^2$$

  where $x_{0,i} = 3, \quad x_i \in [x_{0,i} - 20, x_{0,i} + 20], \quad i = 1, ..., N.$

- For $N = 6$, there are $\approx 10^{10}$ local optima.

- Results:

	Global Minimizer Points			
N	x_i^*	$x_{j \neq i}^*$	Global Minimum	CPU time (s)
2	4.6198510288	5.2820519601	-88.1046253312	0.07
3	4.6201099154	5.2824296177	-87.6730486951	2.12
4	4.6202393815	5.2826184940	-87.4572049443	33.95
5	4.6203170683	5.2827318347	-87.3276809494	413.61
6	4.6203688625	5.2828074014	-87.2413242244	4566.42

CPU times on Dell workstation – 1.7 GHz Xeon running Linux

Some Applications in Chemical Engineering

- Fluid phase stability and equilibrium
  - Activity coefficient models (Stadtherr *et al.*, 1995; Tessier *et al.*, 2000)
  - Cubic EOS (Hua *et al.*, 1996, 1998, 1999)
  - SAFT EOS (Xu *et al.*, 2002)
- Combined reaction and phase equilibrium (Burgos *et al.*, 2004)
- Location of azeotropes: Homogeneous, Heterogeneous, Reactive (Maier *et al.*, 1998, 1999, 2000)
- Location of mixture critical points (Stradi *et al.*, 2001)
- Solid-fluid equilibrium
  - Single solvent (Xu *et al.*, 2000, 2001)
  - Solvent and cosolvents (Scurto *et al.*, 2003)

Applications (cont'd)

- General process modeling problems (Schnepper and Stadtherr, 1996)
- Parameter estimation

$\Longrightarrow$ Relative least squares (Gau and Stadtherr, 1999, 2000)

  - Error-in-variables approach (Gau and Stadtherr, 2000, 2002)

- Nonlinear dynamics

$\Longrightarrow$ Equilibrium states and bifurcations in ecological models (Gwaltney *et al.*, 2004)

- Molecular Modeling
  - Density-functional-theory model of phase transitions in nanoporous materials (Maier *et al.*, 2001)

$\Longrightarrow$ Transition state analysis (Lin and Stadtherr, 2004)

  - Molecular conformations (Lin and Stadtherr, 2004)

Example – Parameter Estimation in VLE Modeling

- Goal: Determine parameter values $\boldsymbol{\theta}$ in activity coefficient models (e.g., Wilson, van Laar, NRTL, UNIQUAC):

$$\gamma_{\mu i,\mathrm{calc}} = f_i(\mathbf{x}_\mu, \boldsymbol{\theta})$$

- Use a relative least squares objective; thus, seek the minimum of:

$$\phi(\boldsymbol{\theta}) = \sum_{i=1}^{n} \sum_{\mu=1}^{p} \left[\frac{\gamma_{\mu i,\mathrm{calc}}(\boldsymbol{\theta}) - \gamma_{\mu i,\mathrm{exp}}}{\gamma_{\mu i,\mathrm{exp}}} \right]^2$$

- Experimental values $\gamma_{\mu i,\mathrm{exp}}$ of the activity coefficients are obtained from VLE measurements at compositions $\mathbf{x}_\mu, \mu = 1, \ldots, p$
- This problem has been solved for many models, systems, and data sets in the DECHEMA VLE Data Collection (Gmehling *et al.*, 1977-1990)

Parameter Estimation in VLE Modeling

- One binary system studied was benzene (1) and hexafluorobenzene (2)
- Ten problems, each a different data set from the DECHEMA VLE Data Collection were considered
- The model used was the Wilson equation

$$\ln \gamma_1 = -\ln(x_1 + \Lambda_{12} x_2) + x_2 \left[\frac{\Lambda_{12}}{x_1 + \Lambda_{12} x_2} - \frac{\Lambda_{21}}{\Lambda_{21} x_1 + x_2} \right]$$

$$\ln \gamma_2 = -\ln(x_2 + \Lambda_{21} x_1) - x_1 \left[\frac{\Lambda_{12}}{x_1 + \Lambda_{12} x_2} - \frac{\Lambda_{21}}{\Lambda_{21} x_1 + x_2} \right]$$

- This has binary interaction parameters

$$\Lambda_{12} = (v_2/v_1) \exp(-\theta_1/RT)$$
$$\Lambda_{21} = (v_1/v_2) \exp(-\theta_2/RT)$$

  where v_1 and v_2 are pure component molar volumes

- The energy parameters θ_1 and θ_2 must be estimated

Results

- Each problem was solved using the IN/GB approach to determine the globally optimal values of the θ_1 and θ_2 parameters
- For each problem, the number of local minima in $\phi(\boldsymbol{\theta})$ was also determined (branch and bound steps were turned off)
- Table 1 compares parameter estimation results for θ_1 and θ_2 with those given in the DECHEMA Collection – New globally optimal parameter values are found in five cases
- CPU times on Sun Ultra 2/1300

Table 1: IN/GB results vs. DECHEMA values

Data Set	Data points	T (^oC)	DECHEMA			IN/GB			No. of Minima	CPU time(s)
			θ_1	θ_2	$\phi(\boldsymbol{\theta})$	θ_1	θ_2	$\phi(\boldsymbol{\theta})$		
1*	10	30	437	-437	0.0382	**-468**	**1314**	**0.0118**	2	15.1
2*	10	40	405	-405	0.0327	**-459**	**1227**	**0.0079**	2	13.7
3*	10	50	374	-374	0.0289	**-449**	**1157**	**0.0058**	2	12.3
4*	11	50	342	-342	0.0428	**-424**	**984**	**0.0089**	2	10.9
5	10	60	-439	1096	0.0047	-439	1094	0.0047	2	9.7
6	9	70	-424	1035	0.0032	-425	1036	0.0032	2	7.9

Data Set	Data points	P (mmHg)	DECHEMA			IN/GB			No. of Minima	CPU time(s)
			θ_1	θ_2	$\phi(\boldsymbol{\theta})$	θ_1	θ_2	$\phi(\boldsymbol{\theta})$		
7*	17	300	344	-347	0.0566	**-432**	**993**	**0.0149**	2	17.4
8	16	500	-405	906	0.0083	-407	912	0.0083	2	14.3
9	17	760	-407	923	0.0057	-399	908	0.0053	1	13.9
10	17	760	-333	702	0.0146	-335	705	0.0146	2	20.5

***New globally optimal parameters found**

Discussion

- Does the use of the globally optimal parameters make a significant difference when the Wilson model is used to predict vapor-liquid equilibrium (VLE)?
- A common test of the predictive power of a model for VLE is its ability to predict azeotropes
- Experimentally this system has two homogeneous azeotropes
- Table 2 shows comparison of homogeneous azeotrope prediction when the locally optimal DECHEMA parameters are used, and when the global optimal parameters are used

Table 2: Homogeneous azeotrope prediction

Data Set	$T(^oC)or$ P (mmHg)	DECHEMA			IN/GB		
		x_1	x_2	P or T	x_1	x_2	P or T
1	T=30	0.0660	0.9340	P=107	0.0541	0.9459	P=107
					0.9342	0.0658	121
2	40	0.0315	0.9685	168	0.0761	0.9239	168
					0.9244	0.0756	185
3	50	NONE			0.0988	0.9012	255
					0.9114	0.0886	275
4	50	NONE			0.0588	0.9412	256
					0.9113	0.0887	274
7	P=300	NONE			0.1612	0.8388	T=54.13
					0.9315	0.0685	52.49

- Based on DECHEMA results, one would conclude Wilson is a poor model for this system. But actually Wilson is a reasonable model if the parameter estimation problem is solved correctly

Example – Nonlinear Dynamics

- Nonlinear dynamic systems are of frequent interest in engineering and science

$$\dot{\mathbf{x}} = \frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, \mathbf{p}); \qquad \mathbf{x} = \text{state variables}; \qquad \mathbf{p} = \text{parameters}$$

- Common problems include computing
  - Equilibrium states ($\dot{\mathbf{x}} = \mathbf{0}$)
  - Bifurcations of equilibria
  - Limit cycles
  - Bifurcations of cycles
- Of specific interest are food chain/web models
  - Use to predict impact on ecosystems of introducing new materials (ionic liquids) into the environment

Ionic Liquids

- Ionic liquids (ILs) are salts that are liquid at or near room temperature
- Many attractive properties
  - No measurable vapor pressure – ILs do not evaporate
  - Many potential applications, including replacement of volatile organic compounds (VOCs) currently used as industrial solvents
  - Eliminates a major source of air pollution
- Could enter the environment via aqueous waste streams
  - Very little environmental toxicity information available
  - Single species toxicity information is not sufficient to predict ecosystem impacts
- Need for ecological risk assessment – Modeling can play an important role

Finding Equilibrium States and Bifurcations

- Equilibrium states: Solve equilibrium conditions for $\mathbf{x}$

$$\dot{\mathbf{x}} = \frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}) = \mathbf{0}$$

- Bifurcations of equilibria: Solve augmented equilibrium conditions for $\mathbf{x}$ and parameter(s) of interest

- Augmenting conditions (in terms of Jacobian matrix $J = d\mathbf{f}/d\mathbf{x}$)
  - Fold and transcritical bifurcations: $\det(J(\mathbf{x}, \alpha)) = 0$
  - Hopf bifurcation: $\det(2J(\mathbf{x}, \alpha) \otimes I) = 0$
  - Fold-fold or fold-Hopf bifurcations: $\det(J(\mathbf{x}, \alpha, \beta)) = 0$ and $\det(2J(\mathbf{x}, \alpha, \beta) \otimes I) = 0$

Finding Equilibrium States and Bifurcations (cont'd)

- These equation systems commonly have multiple solutions
- Typically these systems are solved using a continuation-based strategy (e.g., Kuznetsov, 1991; AUTO software)
  - Initialization dependent
  - No guarantee of locating all solution branches
- Interval mathematics provides a method that is
  - Initialization independent
  - Capable of locating all solution branches with certainty
- As a relatively simple test problem (Gwaltney *et al.*, 2004), consider a tritrophic food chain with logistic prey and hyperbolic predator and superpredator response functions (Rosenzweig-MacArthur model)

Rosenzweig-MacArthur model

In terms of prey(1), predator(2) and superpredator(3) biomasses x_1, x_2 and x_3, the model is given by

$$\frac{dx_1}{dt} = x_1 \left[r\left(1 - \frac{x_1}{K}\right) - \frac{a_2 x_2}{b_2 + x_1} \right]$$

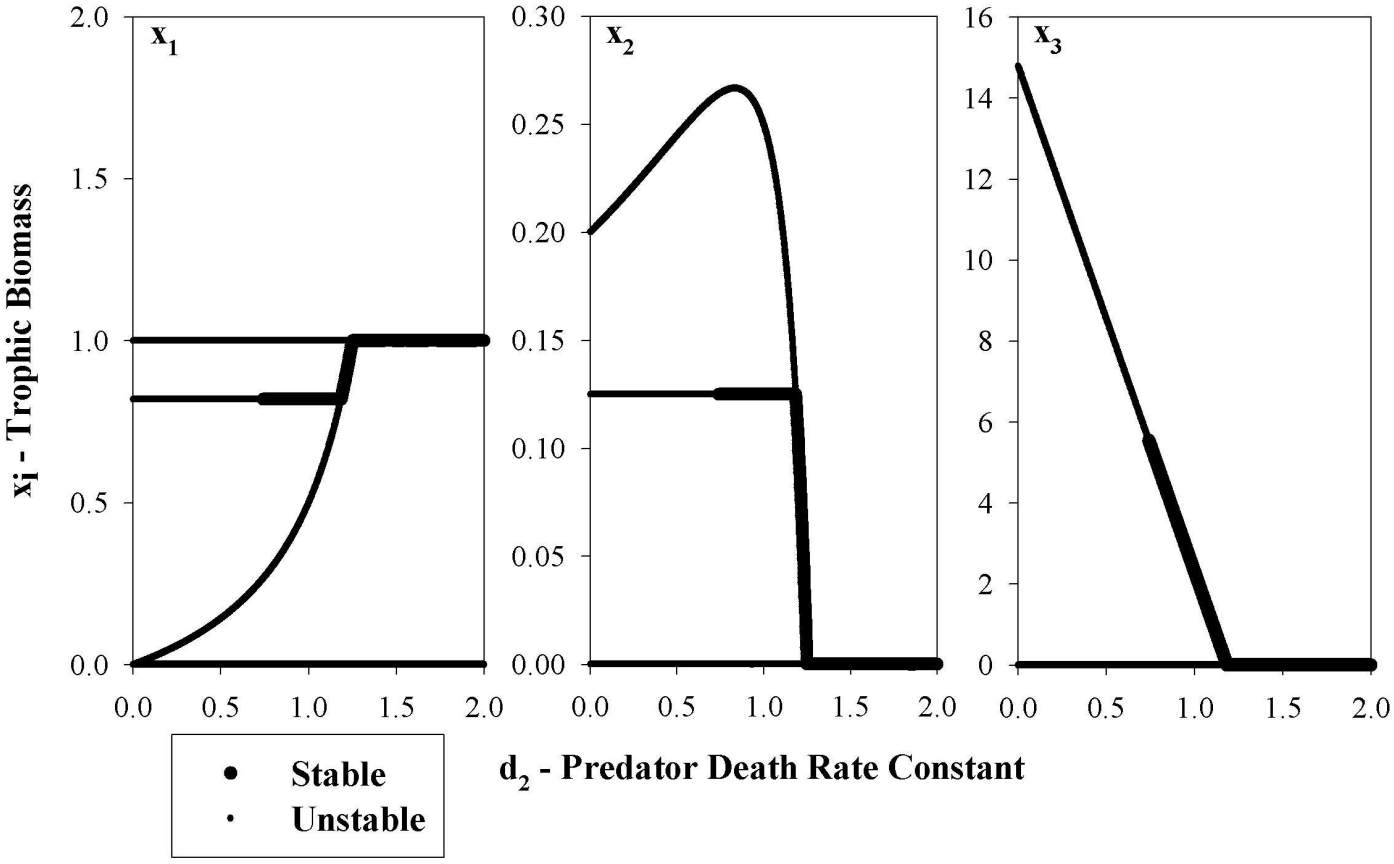
$$\frac{dx_2}{dt} = x_2 \left[e_2 \frac{a_2 x_1}{b_2 + x_1} - \frac{a_3 x_3}{b_3 + x_2} - d_2 \right]$$

$$\frac{dx_3}{dt} = x_3 \left[e_3 \frac{a_3 x_2}{b_3 + x_2} - d_3 \right]$$

Here r is the prey growth rate constant, K is the prey carrying capacity of the ecosystem, the d_i are death rate constants, the a_i represent maximum predation rates, the b_i are half-saturation constants, and the e_i are predation efficiencies

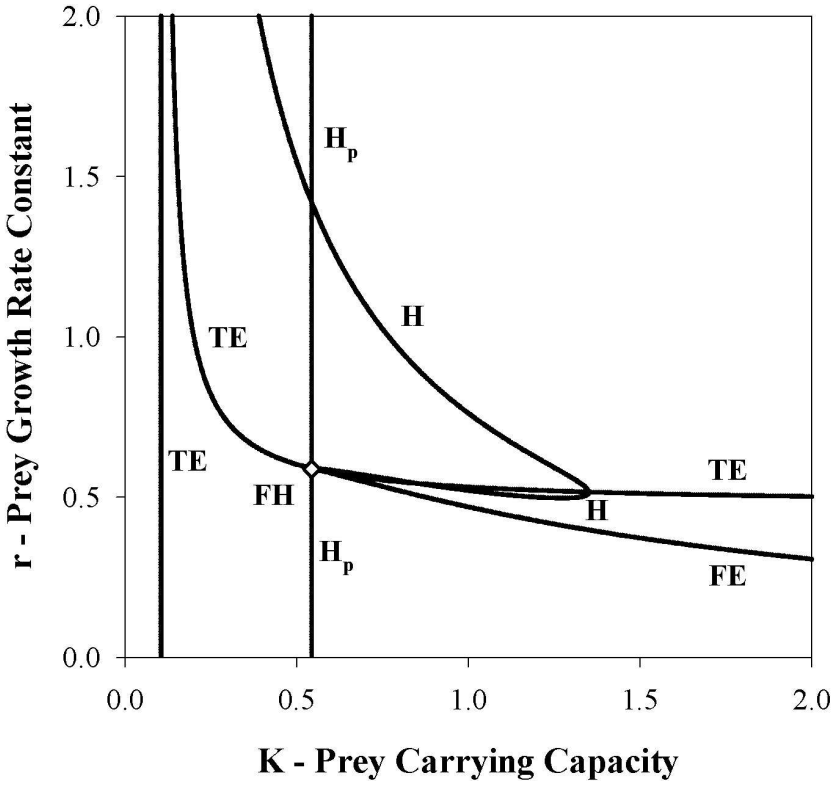
Results – Rosenzweig-MacArthur Model

Example of a solution-branch diagram (equilibrium states vs. one parameter with other parameters fixed) – here x_1, x_2 and x_3 vs. d_2



Results – Rosenzweig-MacArthur Model

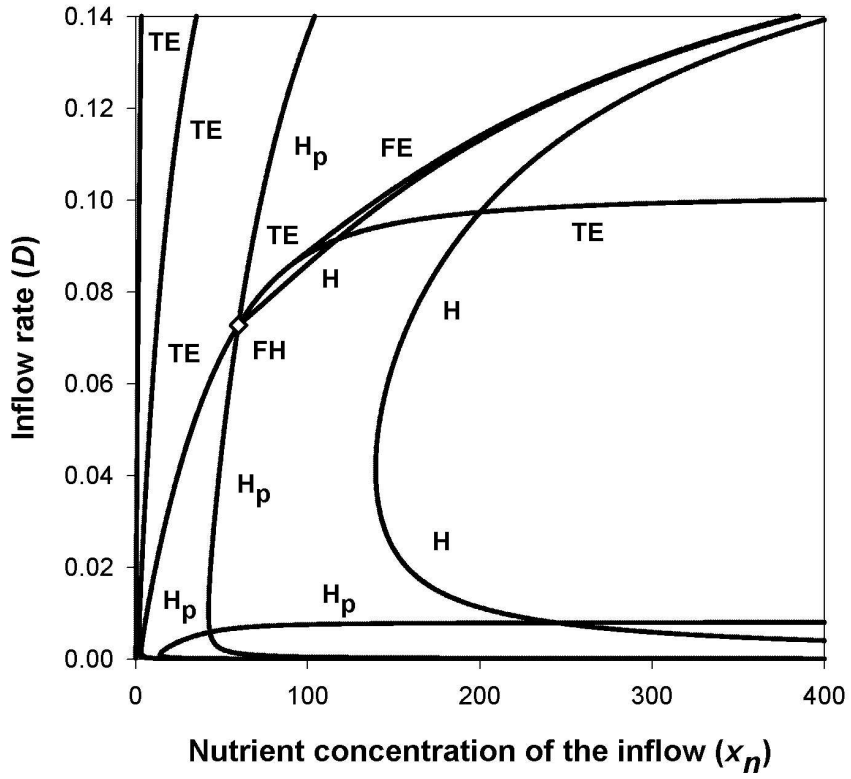
Example of a bifurcation diagram (parameter value at which bifurcation occurs vs. another parameter – here r vs. K with other parameters fixed)



TE = Transcritical of equilibrium; FE = Fold of equilibrium; H = Hopf; H_p = Planar Hopf; FH = Fold-Hopf

Results – Canale's Chemostat Model

This is a more complex model (4 state variables). This is the computed D vs. x_n bifurcation diagram



TE = Transcritical of equilibrium; FE = Fold of equilibrium; H = Hopf; H_p = Planar Hopf; FH = Fold-Hopf

Example – Transition State Analysis

- Transition state analysis is widely used in engineering and science to study the kinetics of various phenomena, e.g.,
  - Chemical reactions
  - Adsorption/desorption to/from surfaces
  - Diffusion through a porous media (e.g., zeolites)
- The key step is identifying stationary points on the potential energy surface $\mathcal{V}$ that characterizes the intermolecular and intramolecular interactions governing the system
- Motion in the system then is assumed to proceed as a series of hops from one local minimum to another, passing through a saddle point (transition state)
- Need a method that is guaranteed to find all stationary points of $\mathcal{V}$

Transition State Analysis (cont'd)

- One example problem – diffusion of xenon in silicalite (June *et al.*, 1991; Lin and Stadtherr, 2004)
- Use truncated Lennard-Jones 6-12 potential

$$\mathcal{V} = \sum_{i=1}^{N} \mathcal{V}_i$$

$$\mathcal{V}_i = \begin{cases} \dfrac{a}{r_i^{12}} - \dfrac{b}{r_i^{6}} & r_i < r_{\text{cut}} \\ 0 & r_i \geq r_{\text{cut}} \end{cases}$$

$$r_i^2 = (x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2$$

where (x, y, z) are the Cartesian coordinates of the xenon, and $(x_i, y_i, z_i), i = 1, \ldots, N$ are the Cartesian coordinates of the $N = 192$ oxygen atoms in a unit cell of the silicalite lattice

- Problem is to solve $\nabla \mathcal{V}(x, y, z) = 0$ for all stationary points (x, y, z)

Results using interval-Newton methodology (LISS_LP)

No.	Type	Energy(kcal/mol)	x(Å)	y(Å)	z(Å)	Connects
1	minimum	-5.9560	3.9956	4.9800	12.1340	
2	minimum	-5.8763	0.3613	0.9260	6.1112	
3	minimum	-5.8422	5.8529	4.9800	10.8790	
4	minimum	-5.7455	1.4356	4.9800	11.5540	
5	minimum	-5.1109	0.4642	4.9800	6.0635	
6	1st order	-5.7738	5.0486	4.9800	11.3210	(1, 3)
7	1st order	-5.6955	0.0000	0.0000	6.7100	(2′, 2)
8	1st order	-5.6060	2.3433	4.9800	11.4980	(1, 4)
9	1st order	-4.7494	0.1454	3.7957	6.4452	(2, 5)
10	1st order	-4.3057	9.2165	4.9800	11.0110	(3, 4)
11	1st order	-4.2380	0.0477	3.9147	8.3865	(2, 4)
12	1st order	-4.2261	8.6361	4.9800	12.8560	(3, 5′)
13	1st order	-4.1405	0.5925	4.9800	8.0122	(4, 5)
14*	2nd order	-4.1404	0.5883	4.8777	8.0138	(4,5),(4,4′)
15	2nd order	-4.1027	9.1881	4.1629	11.8720	(2,3),(4,5)

* Not found by June *et al.* (1991)

Concluding Remarks

- Interval analysis provides a powerful general purpose and model independent approach for solving a wide variety of modeling and optimization problems, giving a mathematical and computational guarantee of reliability.
- Guaranteed reliability of interval methods comes at the expense of CPU time. Thus, there is a choice between fast local methods that are not completely reliable, or a slower method that is guaranteed to give the correct answer.
- The modeler must make a decision concerning how important it is to get the correct answer.
- Continuing advances in computing hardware and software will make this approach even more attractive.
  - Compiler support for interval arithmetic (Sun Microsystems)
  - Parallel computing

Concluding Remarks (cont'd)

- With effective load management strategies, interval methods can be implemented very efficiently using MPI on a networked cluster of workstations (Gau and Stadtherr, 2002).
  - Good scalability
  - Exploit potential for superlinear speedup in optimization
- Parallel computing technology can be used not only to solve problems faster, but to solve problems more reliably.
- Reliability issues are often overlooked:

Are we just getting the wrong answers faster?