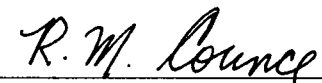
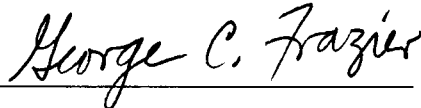


To the Graduate Council:

I am submitting herewith a thesis written by Charles F. Weber entitled "Kinetic Modeling and Parameter Optimization of Iodine Hydrolysis Reactions." I have examined the final copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Chemical Engineering.


J. S. Watson, Major Professor

We have read this thesis
and recommend its acceptance:

Accepted for the Council:


Vice Provost
and Dean of the Graduate School

**KINETIC MODELING AND PARAMETER OPTIMIZATION
OF IODINE HYDROLYSIS REACTIONS**

A Thesis
Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Charles F. Weber

August 1990

ACKNOWLEDGEMENTS

The author is greatly indebted to several staff members of the Oak Ridge National Laboratory for assistance in this effort. As project leader, E. C. Beahm provided overall direction. Both he and W. E. Shockley provided considerable service in the experimental effort, with additional help from S. R. Daish, guest researcher from the United Kingdom. T. S. Kress also assisted in both programmatic and technical planning and support. Wanda Merriweather and Lindy Norris spent considerable time and effort typing the many drafts of this thesis.

Appreciation is also directed to J. S. Watson for many helpful suggestions and to committee members G. C. Frazier and R. M. Counce for their service. Finally, the author is greatly indebted to Martin Marietta Energy Systems and the U.S. Department of Energy for their support of higher education generally, and for the specific assistance received in accomplishing this work.

ABSTRACT

Modeling of iodine hydrolysis is discussed, and rate constants for two particular formulations are obtained from a mathematical optimization procedure. The two models include six and ten principal reaction steps, and are derived to be relevant in situations where radiolysis is also occurring. The resulting kinetic rate equations form a nonlinear system of ordinary differential equations whose solution is straightforward, given values for the rate coefficients and initial species inventories. Since the former are not known *a priori*, they are treated as parameters whose optimal values are determined by fitting the calculated system solution to data. To accomplish this, a methodology is developed which utilizes a quasi-Newton method to minimize a least squares residual. In addition, an experimental effort has resulted in over 130 data points taken in a range of pH between 2.8 and 9.0, and at temperatures 25, 50, and 90°C. The optimization procedure is applied to yield rate constants at 25°C for both models and activation energies for the ten-reaction model. Using the optimized values, the rate equations are solved to obtain time-dependent iodine species inventories. These calculated results compare favorably with actual data, indicating that the models are useful for simulating both forward hydrolysis and the reverse (Dushman) reaction.

Table Of Contents

Section	Page
1. INTRODUCTION	1
2. IODINE HYDROLYSIS REACTIONS	3
2.1. Overall Reaction	3
2.2. Six-Step Reaction	4
2.3. Ten-Step Reaction	5
2.4. Variations and Additional Reactions	7
2.5. Model vs Mechanism	8
3. MATHEMATICAL MODELS OF REACTION SYSTEMS	10
3.1. Basic Rate Equations	10
3.2. Dimensional Analysis and Scaling	11
3.3. Application	14
4. HYDROLYSIS DATA	15
4.1. Experimental Procedure	15
4.2. Results and Analysis	16
5. OPTIMIZATION THEORY FOR PARAMETER ESTIMATION	22
5.1. Solution of Basic Rate Equations	22
5.2. Formulation of the Optimization Problem	23

5.3.	Mathematical Solution of the Optimization Problem	25
5.4.	Parameter Scaling	29
6.	APPLICATION AND RESULTS OF OPTIMIZATION	
	PROCEDURES	31
6.1.	Obtaining Initial Parameter Estimates	31
6.2.	Optimization at 25°C	32
6.3.	Calculation of Species Concentrations	37
6.4.	Comparison of Different Models	39
6.5.	Temperature Dependence of Rate Coefficients	41
7.	SUMMARY AND CONCLUSION	60
	REFERENCES	63
	APPENDIXES	
	APPENDIX A. Rate Equations for the Ten-Reaction System	66
	APPENDIX B. Extensions to the Basic Systems	68
	APPENDIX C. Numerical Solution Procedures	71
	C.1. BFGS Algorithm for Minimizing a Single Nonlinear Equation	71
	C.2. Step-Length Selection	73
	APPENDIX D. Efficient Procedures for Determining the Objective	
	Function Gradient	76

APPENDIX E. Optimization of Initial Inventories	79
APPENDIX F. Description of Optimization Procedures for	
Six-Reaction Model	81
APPENDIX G. Optimization Procedures for Ten-Reaction	
Model at 25°C	87
APPENDIX H. Optimization Theory Extension to Arrhenius Constants	90
APPENDIX I. Computer Coding for Optimization Calculations	92
VITA	97

List Of Tables

Table	Page
1. Suggested Parameters for Six-Reaction Hydrolysis Model	5
2. Dushman Reaction Data	18
3. Hydrolysis Data	19
4. Mean ($\bar{X}$) and Standard Deviation (σ) of Initial Species Inventory (g-atom/dm ³) x 10 ⁴	20
5. Optimal Rate Constants at 25°	33
6. Optimal Initial Inventories at 25° Using Ten-Constant (g-atom/dm ³) x 10 ⁴ Model	34
7. Error Residual Comparison	34
8. Choices of Pre-exponential Factor and Corresponding Activation Energy	43
9. Optimized Arrhenius Parameters for Ten-Constant Model	44
F.1 Initial Scoping Calculations for Six Constant Model	82
F.2 Optimization Results for Initial Scoping Calculations	83
F.3 Advanced Scoping Calculations for Six Constant Model	84
F.4 Optimization of Advanced Scoping Calculations	85
F.5 Parameter Changes and Uncertainty in Six-Constant Model	85
G.1 Primary Optimization for 10-Constants Model at 25°	88

G.2 Optimal Rate Constants for 10-Constant Model at 25° Using	89
---	----

List Of Figures

Figure	Page
1. Species concentration for three representative transients	38
2. Transient 1 data and calculations	45
3. Transient 2 data and calculations	46
4. Transient 3 data and calculations	47
5. Transient 4 data and calculations	48
6. Transient 5 data and calculations	49
7. Transient 6 data and calculations	50
8. Transient 7 data and calculations	51
9. Transient 8 data and calculations	52
10. Transient 9 data and calculations	53
11. Transient 10 data and calculations	54
12. Transient 11 data and calculations	55
13. Transient 12 data and calculations	56
14. Transient 13 data and calculations	57
15. Transient 14 data and calculations	58

16. Transient 15 data and calculations	59
C.1 Residual dependence on step length	74
F.1 Interpolating quadratic polynomial for determining parameter uncertainties	86
I.1 Main routine block diagram	94
I.2 Block diagram for step-length calculation	95
I.3 Block diagram for parameter update calculation	96

1. INTRODUCTION

A major application of aqueous iodine chemistry is in the safety analyses of nuclear power plants. Radioactive isotopes of iodine have received great attention since they are present in large amounts and pose special risks when ingested by people. Iodine may react in a variety of ways within the reactor containment, some of which tend to stabilize and isolate it. However, it is also possible that certain processes will promote release to areas outside the plant; hence, a detailed treatment is necessary in order to assess the dangers that might arise in the event of an accident.

Conditions expected in reactor containments include fairly low concentrations of iodine and impurities (usually $< 10^{-4}\text{M}$). Temperatures exceed 40°C and are often much higher (80°C - 100°C); pH may vary between 3 and 9. During an actual accident, considerable radiation would be absorbed by water pools where iodine could deposit, creating water radiolysis products that react readily with many iodine compounds. It is important to understand what iodine species might result, since some may be volatile enough to be released in large quantities.

The quantitative description of aqueous iodine behavior has been described in hundreds of articles, yet remains a somewhat elusive subject. Most studies have been made at 25°C , using fairly high concentrations ($\geq 10^{-2}\text{M}$), and often contain additional salts or impurities. Very little work has been done in a radiation environment, which is known to be quite important. Because previous work has limited usefulness in nuclear reactor safety, further efforts are needed in both experimentation and modeling.

As a prerequisite to understanding iodine radiolysis, the major focus of this work is to understand iodine hydrolysis more fully. This includes the formation of

various species at different temperature and pH, but without radiation. If a credible model of hydrolysis can be formulated, then it can serve as a foundation upon which a reliable radiolysis model could be built.

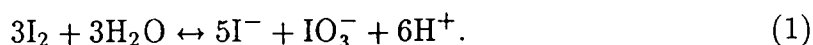
In developing a model of iodine hydrolysis, several different sets of reactions are considered; each is described in detail in Section 2. Each set of reactions results in a set of rate equations, which must be solved to calculate the transient distribution of species. These equations are derived in Section 3 and put in dimensionless form. In order to alleviate the lack of data, an experimental program has been undertaken. A brief description of the experimental procedures and a complete listing of results are given in Section 4.

A major deficiency of all models for iodine hydrolysis is the lack of accurate values for reaction rate coefficients. It was thus necessary to obtain reliable values by extensive fitting to data. This required considerable development of a mathematical optimization procedure in Section 5, which was applied to the two hydrolysis models in Section 6, where reliable rate coefficients are obtained for both models at 25°C. Additional optimizations were made with higher temperature data to yield values for activation energies. Results are summarized, and conclusions presented in Section 7.

2. IODINE HYDROLYSIS REACTIONS

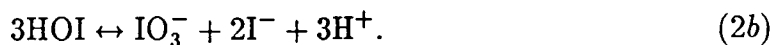
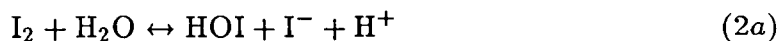
2.1. Overall Reaction

In pure water, I_2 hydrolyzes to form I^- and IO_3^- through processes that can be described by the overall reaction



At equilibrium, only I^- , I_2 , and IO_3^- exist as stable end products (occupying oxidation states -1, 0, and 5, respectively). However, during the transient reaction phase, various intermediate species have been suggested to occupy oxidation states 1, 2, 3, and 4, the most popular being HOI, OI, HIO₂, and IO₂, respectively. In addition, several reaction progressions involving these have been suggested. Although some investigators claim to have measured certain of these intermediates,¹ their involvement in hydrolysis reactions and even their existence remains open to speculation. In the face of such uncertainty, the primary goal of the present modeling effort is to include all essential components, and to provide a simple but realistic description of the overall process which accurately matches experimental data.

The most popular model of iodine hydrolysis is the two-stage reaction²⁻⁴

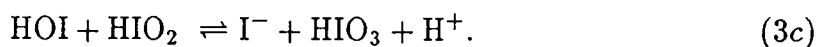
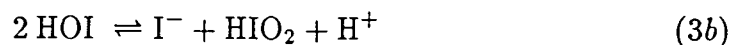
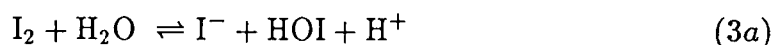


The first reaction is usually quite rapid, while the second step is much slower under many pH and temperature conditions.⁴ It is unlikely that this actually occurs in a single reaction step; rather, Eq. (2b) is viewed as the result of several reaction subprocesses. The major deficiency of this approach for reactor safety calculations

is that several intermediate forms are neglected, since Eq. (2b) converts HOI (oxidation state 1) directly into IO_3^- (oxidation state 5). In radiation environments, water radiolysis products may interact with neglected intermediate species, further altering the progression of reaction (2b).

2.2. Six-Step Reaction

Several researchers have developed complicated models to describe iodine radiolysis. These include reactions between many different iodine species and the free radicals that result from irradiating water. Basic hydrolysis processes (the aqueous iodine reactions that would occur in the absence of radiation) are also included, since they would occur simultaneously. These can be extracted and used by themselves; two such formulations^{5,6} include the following reversible steps to model the overall reaction:



Both authors give values for the rate constants at 25°C; however, these vary considerably, as shown in Table 1. Furthermore, transient calculations using either parameter set are not capable of matching the data presented in Section 4.

This system includes species with oxidation states -1, 0, 1, 3, and 5, which is a considerable improvement over the previous system. The last two reactions involve transitions of ± 2 in oxidation state, corresponding to the transfer of two electrons from one reactant to the other. In addition, the intermediate reaction step allows for greater flexibility, and the additional specie HIO_2 is available for interactions

with water radiolysis products. Nevertheless, the situation is still not complete until all oxidation states are represented and reaction steps are simplified still further.

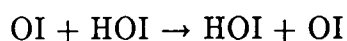
Table 1. Suggested Parameters for Six-Reaction Hydrolysis Model

Parameter	Reaction	Parameter Value ($\text{M}^{-1} \text{s}^{-1}$)	
		Ref. 5	Ref. 6
k_1	Eq.(3a) forward	0.054	31
k_2	reverse	5.87×10^{12}	3.5×10^{15}
k_3	Eq.(3b) forward	10^5	10^{10}
k_4	reverse	5×10^6	10^{10}
k_5	Eq.(3c) forward	3	2.3×10^5
k_6	reverse	1800	3.7×10^{-5}

2.3. Ten-Step Reaction

Because of the deficiencies in previous formulations, it is now useful to develop a model that includes species at each oxidation state from -1 to +5. Since actual reaction mechanisms are unknown, reaction substeps are based on the following simple premises: 1. only oxidation state changes of ± 1 occur, and 2. no specie can be oxidized by another specie with a lower oxidation state.

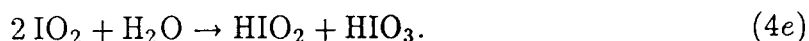
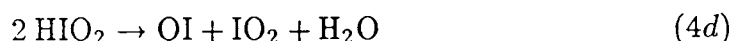
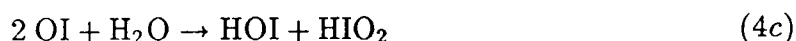
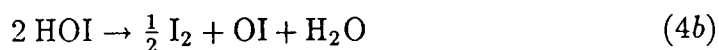
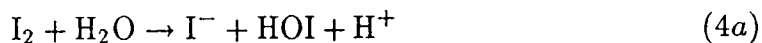
If both conditions hold, reactions involving species at adjacent oxidation states are superfluous and can be omitted. As an example, the reaction



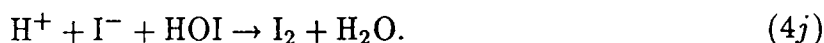
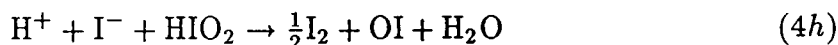
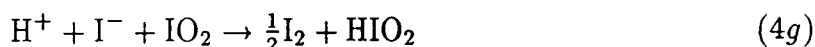
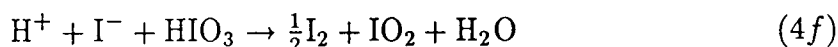
produces no net change in the concentration of either reacting species.

The overall forward reaction of iodine hydrolysis [cf. Eq. (1)] results in a "spreading out" of the occupied oxidation states, beginning at zero and moving to the extremes -1 and 5. Thus, the reaction substeps should also reflect this pattern.

The resulting reactions for the forward process are those involving like species, as modeled by the following irreversible reaction steps:



The substeps for the overall reverse reaction (also known as the Dushman reaction) also utilize the two assumptions stated above. This basic process involves oxidation of I^- and reduction of HIO_3 and other intermediates to form I_2 . Since I^- is the strongest reducing agent, and because it is generally present in greater abundance than other iodine species, the individual reaction steps use only I^- in the stepwise reduction of species in each oxidation state from HIO_3 to I_2 :



The reactions (4a) - (4j) involve species from each oxidation state in both forward and reverse directions. In addition, it is possible for I_2 to be created directly from

I^- and HIO_3 [cf. Eq. (4f)], rather than indirectly through a series of intermediate reaction steps.

2.4. Variations and Additional Reactions

As written, both systems (3) and (4) depend on pH through the H^+ term appearing as a reactant in the back reactions. In fact, data suggest a strong pH dependence, the forward (hydrolysis) reaction being significant at 25°C for $pH > 7$ and the reverse reaction for $pH < 3.2$. In between, very little reaction is noticeable except at higher temperatures. Because neither system involves pH dependence for the forward reaction, it may be difficult to fit data correctly.

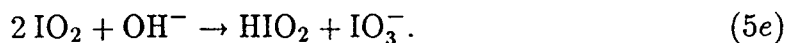
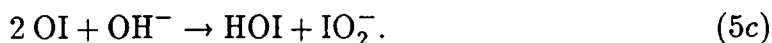
This situation can be remedied by subtracting H^+ from each side of Eq. (3a) (forward only) and Eq. (4a) to obtain



The hydroxide concentration OH^- varies inversely with pH according to the water dissociation equilibrium

$$[OH^-][H^+] = K_w,$$

which is assumed to hold continuously throughout this work. Other variations of the above systems should also be considered. For example, reactions analogous to Eq. (5a) can be written as replacements for Eqs. (4c) and (4e) as well:

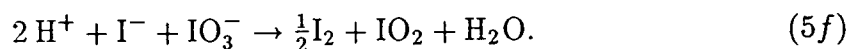


Both of these would give more pH dependence to the overall forward reaction, but add the complication of dealing with both ionized and unionized forms. In the case

of HIO_2 , little is known about its dissociation behavior; hence, Eq. (5c) is no more reliable than Eq. (4c). However, Eq. (5e) is likely to be a better description than Eq. (4e), since iodate should be completely ionized under all conditions encountered in this work. The term IO_3^- in Eq. (5e) can be related to HIO_3 in Eq. (4f) by the acid dissociation equilibrium:

$$K_a = \frac{[\text{H}^+][\text{IO}_3^-]}{[\text{HIO}_3]}$$

or Eq. (4f) could be replaced by the 4th order reaction



One additional reaction that should be considered is the formation of tri-iodide



which will occupy some I^- and I_2 , preventing their participation in other reactions. Although the amount of I_3^- formed should generally be quite small compared to I^- and I_2 , it is mentioned here for the sake of completeness.

2.5. Model vs Mechanism

Even though the reactions in systems (3) and (4) were selected by reasonable criteria, it is important to note that none of the reactions except for Eq. (3a) [equivalent to Eqs. (4a) and (4j)] have been studied very extensively or can be verified. As mentioned previously, only I^- , I_2 , and HIO_3 (or the ionized form IO_3^-) have been documented as stable species; the others are somewhat speculative. The reaction steps do not reflect known chemical mechanisms, but are formulated from simple and reasonable premises as a logical way to transfer iodine between various oxidation states. As a result, these systems are not so much mechanistic

as empirical, with parameters (i.e., the rate coefficients) which are determined by fitting the models to data of overall system behavior.

3. MATHEMATICAL MODELS OF REACTION SYSTEMS

In order to be used in practical calculations, each system of reactions mentioned in the previous section needs to be converted into mathematical equations describing the transient variation of species concentrations. Such equations are put in dimensionless form in Section 3.2 and solved simultaneously to simulate the overall system behavior. This computed solution is then verified by comparison with experimental data.

3.1. Basic Rate Equations

Consider the six-reaction system (3) and define the following variables to represent species concentrations (g-atom/dm³):

$$C_1 = [\text{I}^-], \quad C_2 = 2[\text{I}_2], \quad C_3 = [\text{HIO}], \quad C_4 = [\text{HIO}_2], \quad C_5 = [\text{HIO}_3], \quad C_H = [\text{H}^+].$$

Using the notation $\dot{C} = dC/dt$, the following rate expressions can be derived for the six reactions (3 forward and 3 reverse):

$$\dot{C}_1 = -\frac{1}{2}\dot{C}_2 = \dot{C}_3 = \frac{1}{2}k_1C_2 \qquad -\dot{C}_1 = \frac{1}{2}\dot{C}_2 = -\dot{C}_3 = k_2C_1C_3C_H$$

$$\dot{C}_1 = -\frac{1}{2}\dot{C}_3 = \dot{C}_4 = k_3C_3^2 \qquad -\dot{C}_1 = \frac{1}{2}\dot{C}_3 = -\dot{C}_4 = k_4C_1C_4C_H$$

$$\dot{C}_1 = -\dot{C}_3 = -\dot{C}_4 = \dot{C}_5 = k_5C_3C_4 \qquad -\dot{C}_1 = \dot{C}_3 = \dot{C}_4 = -\dot{C}_5 = k_6C_1C_5C_H$$

These rates for individual reactions can be combined to form overall rate equations

$$\dot{C}_1 = \frac{1}{2}k_1C_2 - k_2C_1C_3C_H + k_3C_3^2 - k_4C_1C_4C_H + k_5C_3C_4 - k_6C_1C_5C_H \quad (6a)$$

$$\dot{C}_2 = -k_1C_2 + 2k_2C_1C_3C_H \quad (6b)$$

$$\dot{C}_3 = \frac{1}{2}k_1C_2 - k_2C_1C_3C_H - 2k_3C_3^2 + 2k_4C_1C_4C_H - k_5C_3C_4 + k_6C_1C_5C_H \quad (6c)$$

$$\dot{C}_4 = k_3C_3^2 - k_4C_1C_4C_H - k_5C_3C_4 + k_6C_1C_5C_H \quad (6d)$$

$$\dot{C}_5 = k_5C_3C_4 - k_6C_1C_5C_H. \quad (6e)$$

These equations constitute a system of five ordinary differential equations which is highly nonlinear in the five unknown functions $C_i(t)$, $i = 1, \dots, 5$. An analogous system of seven equations in seven unknown functions can be obtained for the ten reaction model (4); details are provided in Appendix A.

If Eq. (5a) is used in place of the forward reaction in Eq. (3a), then the appropriate rate expression is

$$\dot{C}_1 = -\frac{1}{2}\dot{C}_2 = \dot{C}_3 = \frac{1}{2}k_1C_2[\text{OH}^-] = \frac{1}{2}k_1K_wC_2/C_H,$$

and the resulting system of equations is the same as system (6) with k_1 replaced by

$$\tilde{k}_1 = k_1K_w/C_H. \quad (7)$$

The implementation of alternative reactions (5e) and (5g) is discussed in Appendix B.

3.2. Dimensional Analysis and Scaling

Each equation in system (6) contains terms with different products of various concentrations, which may vary by many orders of magnitude. For example, Eq. (6b) contains two terms; the concentration C_2 may be 10^{-5} - 10^{-4} g-atom/dm³, whereas the product $C_1C_3C_H$ might be 10^{-20} - 10^{-25} (g-atom/cm³)³. This implies that the rate constants may also vary by many orders of magnitude. This great variation presents problems during numerical computation, and it is advantageous

to scale the concentrations so that all terms and all rate constants are within a few orders of magnitude.

In addition, several factors are present which may drastically affect both the speed and extent of reaction. Most notable are the pH and temperature, with lesser effects from concentration (at least within the range considered here) and pressure. For example, at 25°C and pH = 7 the forward hydrolysis reaction takes roughly 16 hours to reach 70% completion. At pH = 9, the reaction is 85% complete in just 5 minutes. For ease of analysis, it is convenient to place each transient on a dimensionless time scale relative to its maximum observation time T . This represents the final time for an experimental transient, which generally corresponds to a time sufficient for equilibrium to be attained.

The scaling procedure to address these concerns consists of a change in the independent time variable and each dependent concentration variable as follows:

$$s = \frac{t}{T}, \quad y_i(s) = \frac{C_i(sT)}{C_T} = \frac{C_i(t)}{C_T}, \quad (8)$$

where C_T is the total iodine concentration (g-atom/dm³) in the system, which is constant over time:

$$C_T = \sum_{i=1}^5 C_i(t).$$

Using the notation $\dot{y} = dy/ds$, straightforward differentiation gives

$$\dot{y}_i = \frac{T}{C_T} \dot{C}_i. \quad (9)$$

Substituting Eqs. (8) and (9) into system (6) yields the dimensionless system

$$\dot{y}_1 = \frac{1}{2}u_1y_2 - u_2y_1y_3 + u_3y_3^2 - u_4y_1y_4 + u_5y_3y_4 - u_6y_1y_5 \quad (10a)$$

$$\dot{y}_2 = -u_1 y_2 + 2u_2 y_1 y_3 \quad (10b)$$

$$\dot{y}_3 = \frac{1}{2}u_1 y_2 - u_2 y_1 y_3 - 2u_3 y_3^2 + 2u_4 y_1 y_4 - u_5 y_3 y_4 + u_6 y_1 y_5 \quad (10c)$$

$$\dot{y}_4 = u_3 y_3^2 - u_4 y_1 y_4 - u_5 y_3 y_4 + u_6 y_1 y_5 \quad (10d)$$

$$\dot{y}_5 = u_5 y_3 y_4 - u_6 y_1 y_5, \quad (10e)$$

where u_i are the dimensionless rate coefficients

$$u_1 = k_1 T, \quad u_3 = k_3 T C_T, \quad u_5 = k_5 T C_T, \quad (11a)$$

$$u_i = k_i T C_T C_H, \quad i = 2, 4, 6. \quad (11b)$$

If Eq. (5a) is used in place of the forward reaction in Eq. (3a), then Eq. (7) implies that the form of system (10) is unchanged, but u_1 in Eq. (11a) is replaced by

$$\tilde{u}_1 = \tilde{k}_1 T = k_1 K_w T / C_H. \quad (11c)$$

The development of rate equations and scaling for the ten reaction model (4) is done in an analogous manner, and details are given in Appendix A. The resulting dimensionless system is:

$$\dot{y}_1 = \frac{1}{2}u_1 y_2 - y_1(u_6 y_7 + u_7 y_6 + u_8 y_5 + u_9 y_5 + u_9 y_4 + u_{10} y_3) \quad (12a)$$

$$\dot{y}_2 = -u_1 y_2 + u_2 y_3^2 + y_1(u_6 y_7 + u_7 y_6 + u_8 y_5 + u_9 y_4 + 2u_{10} y_3) \quad (12b)$$

$$\dot{y}_3 = \frac{1}{2}u_1 y_2 - 2u_2 y_3^2 + u_3 y_4^2 + y_1(u_9 y_4 - u_{10} y_3) \quad (12c)$$

$$\dot{y}_4 = u_2 y_3^2 - 2u_3 y_4^2 + u_4 y_5^2 + y_1(u_8 y_5 - u_9 y_4) \quad (12d)$$

$$\dot{y}_5 = u_3 y_4^2 - 2u_4 y_5^2 + u_5 y_6^2 + y_1(u_7 y_6 - u_8 y_5) \quad (12e)$$

$$\dot{y}_6 = u_4 y_5^2 - 2u_5 y_6^2 + y_1(u_6 y_7 - u_7 y_6) \quad (12f)$$

$$\dot{y}_7 = u_5 y_6^2 + y_1(-u_6 y_7) \quad (12g)$$

where the dimensionless rate coefficients are

$$u_1 = k_1 T, \quad (13a)$$

$$u_i = k_i T C_T, \quad i = 2, \dots, 5 \quad (13b)$$

$$u_i = k_i T C_T C_H, \quad i = 6, \dots, 10. \quad (13c)$$

If Eq. (5a) is used in place of Eq. (4a), then again u_1 in system (12) is replaced by $\tilde{u}_1$ from Eq. (11c). In addition, the extensions represented by reactions (5e) and (5g) easily fit the form of the above system, and are developed in Appendix B. In practice both of these extensions enhance and improve the models; hence, they are used in all calculations performed and reported in succeeding sections.

3.3. Application

The previous developments have yielded systems of differential exit equations (10) and (12), corresponding to the chemical reaction models (3) and (4), respectively. Given the pH, initial concentrations, and values for the rate coefficients k_i , either system could be used in calculating a transient distribution of species. In practice, the pH and initial values could be chosen to correspond to actual experimental transients. However, as mentioned in Section 2, the chemical models (and therefore their mathematical representations) are not well understood or verified; hence, reliable values for the rate coefficients are not available. Thus, it is necessary to determine values for each so that systems (10) and (12) accurately simulate iodine hydrolysis reactions. In fact, it is the major goal of this work, and a formal methodology will be developed in Section 5.

4. HYDROLYSIS DATA

4.1. Experimental Procedure

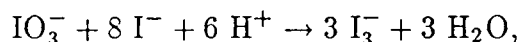
Aqueous iodine chemistry has been studied extensively in the past, and hundreds of experimental results have been obtained. Invariably, these involve much higher concentrations than would occur in reactor containments, and often contain salts or other "impurities" that might affect hydrolysis processes. To provide more useful data for nuclear accident analysis, a number of batch experiments were performed under more appropriate conditions. These generally involve pH values of 3-9, temperatures between 25°C and 90°C, atmospheric pressure, and total iodine concentrations of 10^{-4} - 10^{-3} M. The results form a sufficient data base for the quantitative selection of rate constants for models discussed in previous sections.

Each batch experiment involves the dissolution of various iodine species in water, and the measurement of I^- , I_2 , and IO_3^- after a specified time. For the forward (hydrolysis) reaction, a small quantity of 10^{-3} M I_2 solution is added to a borated water solution. For the reverse (Dushman) reaction, CsI, HI, and HIO_3 are added to borated water, the pH being adjusted by the distribution of species.

Borate is typically used in experimental studies supporting reactor safety, since virtually every light water reactor uses borated water as a neutron absorber. Borate is also an effective buffer, allowing the reactions to proceed at nearly constant pH. Reactions between borate and the various iodine species have not been noted in any of the literature reviewed for this work, although a slight ionic strength effect may occur at neutral pH.

The measurement procedure itself is similar to that described in ref. 7 and is summarized as follows:

1. The I_2 is extracted into isooctane, and then back stripped into 0.2 M NaOH solution containing hydroxylamine-hydrochloride. The resulting iodide is measured with a specific ion electrode.
2. The I^- occurring in the original solution is measured using a specific ion electrode.
3. After removing I_2 with phenol, the IO_3^- is converted into I_3^- by the addition of excess acidified I^- via the reaction



and measured by UV-visible spectroscopy.

Typically, these steps are performed simultaneously by three different people. The concentrations of measured species are frozen in each step within 30–60 s (by isooctane extraction, quick measurement, and addition of acidified I^- , respectively). For long transients, this measurement time is immaterial; however, for the most rapid transients (e.g., Transient 15) the time delay of measurement may impose experimental error of 5–10% for some of the data points.

4.2. Results and Analysis

Raw data from over 100 batch reactor experiments are shown in Tables 2 and 3, grouped according to similarities in initial concentrations, temperature, and pH. Each of these groupings is then considered as a single transient, i.e., the description of time varying behavior of a single system. The initial inventories are shown for the two different types of experiments—the Dushman reaction of I^- and IO_3^- , and

the hydrolysis of I_2 . Also shown are the final inventories of each measured species, normalized by dividing by the total initial inventory. Where more than one entry appears at the same time within the same group, a simple arithmetic average is used to yield a single value.

Each transient has an initial nonzero concentration either for I_2 , or for I^- and IO_3^- . These are shown in columns 3-5 of Table 2 and column 3 of Table 3. Entries showing one or two significant digits were estimated using assumed compositions of initial solution ingredients; those with three or four digits were actually measured by the same procedures used for final reaction products (i.e., using either step 1, or steps 2 and 3 in Section 4.1). The uncertainty in either approach is probably 5-10%.

While the pH and temperature are constant, the initial species concentrations for each batch experiment within a transient may vary slightly. Hence, an average value is used for the group as a whole. Table 4 depicts these mean values, together with the standard deviations, which in most cases are less than 10% of the mean values. This error in averaging is probably no greater than the uncertainty in the individual initial concentrations. For transients whose initial inventories were estimated, rather than measured, the average is equal to this estimated value and the deviation is zero.

The 15 transients in Tables 2 and 3 can each be simulated by solving system (10) or (12), using the average initial values shown in Table 4. As previously discussed, the uncertainty in these values arises both from the averaging process (quantified by the standard deviation) and from errors in the initial concentrations of the batch reactor data. The total uncertainty is generally in the range of 10-25%, indicating that variations in these initial values should also be considered in simulating

Table 2. Dushman Reaction Data

	Time (min)	Initial inventory (g-atom/dm ³) × 10 ³			Final inventory fraction (g-atom basis)			Run #
		I ⁻	IO ₃ ⁻	Total	I ⁻	I ₂	IO ₃ ⁻	
Transient 1	5	0.1	1.412	1.512	.0339	.0325	.9074	69
pH = 2.8	15	0.1	1.635	1.735	.0130	.0515	.9193	70
T = 25°C	30	0.1	1.654	1.754	.0101	.0573	.9019	108
	60	0.1	1.451	1.551	.0100	.0629	.8214	112
Transient 2	5	0.1	0.900	1.000	.0887	.0157	.7198	64
pH = 3.7	15	0.1	0.850	0.950	.0781	.0272	.7183	61
T = 25°C	15	0.1	0.787	0.887	.0827	.0282	.7665	77
	30	0.1	0.917	1.017	.0562	.0468	.7166	65
	60	0.1	1.118	1.218	.0216	.0710	.7342	63
	120	0.1	0.963	1.063	.0212	.0836		62
Transient 3	30	0.507	0.104	0.611	0.659	0.105	0.171	101
pH = 3.2	60	0.555	0.099	0.654	0.642	0.152	0.125	104
T = 25°C	120	0.524	0.095	0.618	0.539	0.235	0.106	103
	240	0.503	0.094	0.597	0.435	0.245	0.122	111
Transient 4	15	.5693	.0200	.5893	0.852	.0012	.0289	78
pH = 3.2	30	.5528	.0190	.5718	0.871	.0191	.0269	93
T = 25°C	30	.5587	.0192	.5779	0.846	.0167	.0276	96
	60	.4734	.0191	.4925	0.864	.0311	.0312	91
	60	.5311	.0190	.5501	0.810	.0330	.0284	98
	90	.5552	.0198	.5750	0.765	.0442	.0230	94
	120	.5320	.0186	.5507	0.793	.0716	.0169	92
	120	.5300	.0193	.5493	0.803	.0547	.0224	97
	150	.5334	.0193	.5527	0.740	.0550	.0194	95
	150	.5468	.0191	.5659	0.721	.0713	.0186	99
	240	.5088	.0182	.5271	0.724	.0855	.0201	100
	360	.5899	.0202	.6101	0.649	.1203	.0177	109
	360	.5596	.0204	.5800	0.652	.0644	.0240	114
Transient 5	5	0.1	.823	.923	.050	.026	.751	81
pH = 3.0	15	0.1	.894	.994	.043	.054	.806	79
T = 50°C	30	0.1	.911	1.011	.035	.068	.793	80
	60	0.1	.880	.980	.017	.078	.789	82
Transient 6	15	.500	.0191	.519	.795	.0198	.0287	87
pH = 3.2	15	.477	.0200	.497	.733	.0227	.0312	115
T = 50°C	30	.535	.0196	.555	.915	.0384	.0247	118
	60	.490	.0199	.510	.866	.0536	.0227	117
	120	.493	.0198	.513	.637	.1000	.0170	116
	240	.536	.0198	.556	.788	.1330	.0029	119
Transient 7	15	1.998	.0729	2.071	.886	.072	.020	121
pH = 3.2	30	2.056	.0716	2.128	.856	.099	.014	120
T = 90°C	60	2.021	.0746	2.096	.573	.120	.010	122

Table 3. Hydrolysis Data

	Time (min)	Initial I ₂ (g-atom/dm ³) × 10 ⁴	Final inventory fraction (g-atom basis)			Run #
			I ⁻	I ₂	IO ₃ ⁻	
Transient 8	15	1.2	0.082	0.893	0.013	4
pH = 7	15	1.2		0.847	0.012	6
T = 25°C	30	1.2		0.666	0.021	2
	30	1.2	0.206	0.674	0.021	48
	60	1.2	0.228	0.730	0.031	26
	60	1.2	0.218	0.692	0.049	28
	120	1.2	0.239	0.553	0.050	27
	240	1.2	0.313	0.561	0.063	47
	360	1.2	0.356	0.475	0.089	31
	420	1.2	0.342	0.399	0.093	51
	960	1.2	0.486	0.274	0.111	30
Transient 9	5	1.2	0.403	0.510	0.056	38
pH = 8	5	1.2		0.514	0.049	18
T = 25°C	10	1.2	0.460	0.539	0.065	44
	10	1.2	0.457	0.531	0.068	45
	15	1.2	0.483	0.328	0.096	37
	15	1.2		0.400	0.093	1
	30	1.2		0.225	0.114	20
	30	1.2	0.596		0.107	35
	30	1.2	0.578	0.241	0.108	36
	60	1.2		0.155	0.149	16
	60	1.2	0.661	0.202	0.119	34
	120	1.2	0.675	0.112	0.128	46
Transient 10	5	1.2	0.534	0.293	0.082	43
pH = 8.5	15	1.2	0.604	0.155	0.116	42
T = 25°C	30	1.2	0.673	0.091	0.122	41
	60	1.2	0.719	0.049	0.125	40
	120	1.2	0.735	0.026	0.128	39
Transient 11	3	1.2	0.671	0.165	0.126	50
pH = 9	5	1.2	0.606	0.133	0.119	24
T = 25°C	10	1.2	0.863	0.062	0.145	49
	15	1.2	0.676	0.050	0.131	23
	30	1.2	0.733	0.027	0.148	22
	60	1.2	0.807	0.016	0.160	21
Transient 12	15	1.2	0.336	0.421	0.079	58
pH = 7	30	1.2	0.401	0.253	0.090	57
T = 50°C	60	1.2	0.462	0.222	0.059	59
Transient 13	15	1.0508	0.274	0.529	0.067	130
pH = 5.5	30	1.1589	0.445	0.409	0.108	128
T = 90°C	60	1.0143	0.429	0.282	0.116	131
	120	0.8842	0.566	0.247	0.115	129
Transient 14	10	1.0203	0.440	0.407	0.048	127
pH = 6.0	15	1.1590	0.542	0.355	0.091	124
T = 90°C	30	0.9548	0.592	0.228	0.095	125
	60	0.9432	0.735	0.153	0.119	126
Transient 15	5	1.2	0.522	0.120	0.121	54
pH = 7.0	10	1.2	0.699	0.068	0.133	56
T = 90°C	15	1.2	0.595	0.045	0.133	53
	30	1.2	0.741	0.017	0.133	55

Table 4. Mean ($\bar{X}$) and Standard Deviation (σ) of Initial Species Inventory
(g-atom/dm³) $\times 10^4$

Transient	I ⁻		IO ₃ ⁻		I ₂	
	$\bar{X}$	σ	$\bar{X}$	σ	$\bar{X}$	σ
1	1.0	0	15.38	1.24		
2	1.0	0	9.43	1.11		
3	5.22	0.24	0.980	0.0067		
4	5.42	0.27	0.193	0.0455		
5	1.0	0	8.77	0.38		
6	5.05	0.25	0.197	0.0031		
7	5.40	0.078	0.195	0.0039		
8					1.2	0
9					1.2	0
10					1.2	0
11					1.2	0
12					1.2	0
13					1.027	.113
14					1.019	.099
15					1.2	0

each transient. This is done by treating the nonzero initial concentrations as parameters which can be varied so as to optimize the system. The mathematical details of this procedure follow the development of the next section, and are given in Appendix E. Thus, the final result will be optimized with respect to both the nonzero initial values and the reaction rate coefficients.

Finally, note that these results involve only the end products I⁻, I₂, and IO₃⁻, which are stable and measurable with reliable laboratory methods. Any intermediate species are assumed to be present in small enough amounts so that (1) failure to measure them will not significantly alter the total recorded iodine inventories, and (2) only negligible quantities of volatile intermediates will evaporate. Initial concentrations of these species are assumed to be zero.

5. OPTIMIZATION THEORY FOR PARAMETER ESTIMATION

5.1. Solution of Basic Rate Equations

The data shown in Section 4 must be used in some meaningful way to select rate coefficients for the different reaction models developed in Section 2. Once this is done, the mathematical models in Section 3 can be used to simulate actual reaction kinetics, so that calculated solutions of the system (10) or (12) will yield good approximations to the experimental transients in Tables 2 and 3. It is thus necessary to develop a quantitative procedure to select optimal values for chemical rate coefficients.

Either system (10) or (12) [with or without the modification of Eq. (11c) and those described in Appendix B] can be expressed in the form

$$\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}, \mathbf{u}, s), \quad (14)$$

where $\mathbf{y} = (y_1, y_2, \dots, y_I)^T$ is the vector of dimensionless species concentrations at time s , and $\mathbf{f}$ is a vector function. In order to simulate overall hydrolysis behavior, Eq. (14) must be solved to obtain the time-varying values of $\mathbf{y}$, given an estimate of the parameter vector $\mathbf{u}$. This task is not trivial, since system (14) is generally stiff.* However, it is not intractable either, and major scientific software libraries contain routines that can handle this type of problem. In this study, the routine LSODE⁸ was used in all cases to solve Eq. (14), and no difficulties were encountered.

If a very good estimate of $\mathbf{u}$ is available, then it remains only to solve Eq. (14) in order to simulate the reaction kinetics involved. However, as mentioned in Section 3, most of the reactions leading to Eq. (14) are speculative or not well understood.

* A problem is considered stiff if there are great differences (many orders of magnitude) in the time constants of the solutions y_i .

Virtually no data exist to estimate concentrations of various intermediate species or rates of the reaction substeps. This implies that the information needed to supply a good estimate of $\mathbf{u}$ is not available *a priori*, but must somehow be determined using the gross system data presented in Section 4.

5.2. Formulation of the Optimization Problem

Consider the simulation of a single transient in which initial species concentrations are known

$$\mathbf{y}(0) = \mathbf{y}_0 = (y_1(0), y_2(0), \dots, y_I(0))^T, \quad (15)$$

and experimental values at certain time points have been measured. Denoting the latter as

$\mathbf{x}(s_n) \equiv$ vector of experimentally measured concentrations
at dimensionless time s_n ,

where $s_n = t_n/T$, it is desired to pick rate coefficients so that the solution $\mathbf{y}$ of Eq. (14) will closely match $\mathbf{x}$ at the time points s_n . This is most commonly accomplished by minimizing the least squares residual

$$\min_{\mathbf{u} > 0} R = \min_{\mathbf{u} > 0} \sum_n \mathbf{e}_n^T \mathbf{e}_n, \quad (16)$$

where $\mathbf{e}_n \equiv \mathbf{y}(s_n) - \mathbf{x}(s_n)$ is the error at time s_n . Other residuals might also be used, by replacing $\mathbf{e}_n$ with relative or logarithmic errors:

$$\hat{\mathbf{e}}_n \equiv \frac{\mathbf{y}(s_n) - \mathbf{x}(s_n)}{\mathbf{x}(s_n)}, \quad \tilde{\mathbf{e}}_n \equiv \ln \mathbf{y}(s_n) - \ln \mathbf{x}(s_n) = \ln \frac{\mathbf{y}(s_n)}{\mathbf{x}(s_n)}.$$

Both of these tend to refine the prediction of smaller concentrations at the expense of larger ones. In the current work, concentrations are significant relative to their magnitude; hence, the absolute error residual (16) was used.

In practice, it is necessary to pick rate coefficients so that many experimental transients are simulated adequately. Assuming there are M such transients, and letting $\mathbf{y}^m$ and $\mathbf{x}^m$ denote the calculated and experimental concentrations, respectively, then Eq. (14) must be satisfied for each transient:

$$\dot{\mathbf{y}}^m = \mathbf{f}(\mathbf{y}^m, \mathbf{u}, s), \quad m = 1, 2, \dots, M. \quad (17)$$

Each experiment has its own initial values $\mathbf{y}^m(0) = \mathbf{y}_0^m$, but the parameter vector $\mathbf{u}$ and the function vector $\mathbf{f}$ are the same for all experiments. Furthermore, it is desired to optimize the choice of $\mathbf{u}$ over all experiments, so Eq. (16) is expanded to get

$$\min_{\mathbf{u} > 0} \sum_{n,m} (\mathbf{e}_n^m)^T \mathbf{e}_n^m, \quad \mathbf{e}_n^m \equiv \mathbf{y}^m(s_n) - \mathbf{x}^m(s_n). \quad (18)$$

Defining the new vectors

$$\mathbf{Y} \equiv \begin{bmatrix} \mathbf{y}^1 \\ \mathbf{y}^2 \\ \vdots \\ \mathbf{y}^M \end{bmatrix}, \quad \mathbf{X} \equiv \begin{bmatrix} \mathbf{x}^1 \\ \mathbf{x}^2 \\ \vdots \\ \mathbf{x}^M \end{bmatrix}, \quad \mathbf{F} \equiv \begin{bmatrix} \mathbf{f}(\mathbf{y}^1, \mathbf{u}, s) \\ \mathbf{f}(\mathbf{y}^2, \mathbf{u}, s) \\ \vdots \\ \mathbf{f}(\mathbf{y}^M, \mathbf{u}, s) \end{bmatrix}, \quad \mathbf{E}_n = \begin{bmatrix} \mathbf{e}_n^1 \\ \mathbf{e}_n^2 \\ \vdots \\ \mathbf{e}_n^M \end{bmatrix}$$

then Eqs. (17) and (18) can be expressed in the final form:

$$\min_{\mathbf{u} > 0} R = \sum_n \mathbf{E}_n^T \mathbf{E}_n, \quad (19a)$$

$$\text{subject to:} \quad \dot{\mathbf{Y}} = \mathbf{F}(\mathbf{Y}, \mathbf{u}, s) \quad (19b)$$

$$\mathbf{Y}(0) = \mathbf{Y}_0 \quad (19c)$$

The system (19) is similar to a constrained optimization problem, with a quadratic objective function (19a), and the differential system (19b, c) as a constraint. Such systems commonly arise in the theory of optimal control, where $\mathbf{u}$ is the control vector and $\mathbf{Y}$ is the system solution.

5.3. Mathematical Solution of the Optimization Problem

Although solutions to Eq. (19) are not easily obtained, solutions for simple versions have been published. If the differential equations (19b) are linear with constant coefficients, then an optimal solution can be determined from the solution of a single matrix equation by several different means.⁹ In reality, though, Eq. (19b) is quite nonlinear, and some iterative procedure must be used.

A local minimum for Eq. (19a) must satisfy the standard criteria

$$\frac{\partial R}{\partial u_j} = 0, \quad j = 1, 2, \dots, J. \quad (20)$$

All iterative techniques begin with some initial guess $\mathbf{u}^0$ for which Eq. (20) does not generally hold. The first estimate $\mathbf{Y}^1$ is then calculated from Eqs. (19b,c). The derivatives $\partial R/\partial u_j$ are calculated or estimated in various ways, and used to calculate a new estimate $\mathbf{u}^1$. The cycle is repeated until Eq. (20) is satisfied within acceptable tolerance and the residual R appears to be a minimum.

Because of the complexities in system (19), many local minima may exist, and locating them may depend strongly on the initial guess. While some procedure for determining a global minimum is obviously desired, no analytic method is generally available.¹⁰ The only recourse is to examine a number of local minima, and use the smallest as a good estimate of the global minimum. It is also important to realize that uncertainties of the data points will produce some uncertainty in both local and global minima. This in turn implies that the global minimum of the mathematical formulation may not actually be a global optimum for the physical problem. It is thus pointless to search exhaustively for a true global extremum, but is sufficient to find a nearly optimal solution that is reasonable from a chemical kinetic point of view.

As mentioned above, obtaining a good global solution is likely to depend strongly on finding a good initial guess for the parameter vector $\mathbf{u}$. This is a critical problem in the current situation, since an initial guess is not obvious. However, elimination of many unlikely candidates can be accomplished through close screening and application of basic chemical understanding. These matters will be discussed more fully in the next section, where they form an important part of the overall process for obtaining an optimal solution.

Several mathematical approaches have been suggested for solving the parameter estimation problem (19). The method of quasilinearization¹¹ was developed for problems such as this, and has even been used to estimate rate constants in a simple chemical system.¹² In this instance, the forward and reverse rate constants for a single reaction were refined by using an initial guess obtained from graphical analysis. Furthermore, it has been shown that if the initial guess is close enough to an optimal point, that convergence will occur.¹³ Another approach is the Gauss-Newton method, which also has been demonstrated on small problems.^{9,13} A formidable difficulty with both of these procedures is the possibility of divergence unless the initial guess is "close enough" to the optimal solution.^{10,13} In addition, the complexity of system (19) is considerably greater than the examples which have been successfully addressed by either quasilinearization or the Gauss-Newton scheme; there is no assurance that these methods will be adequate. In fact, both were tried with a variety of initial guesses for the parameter vector $\mathbf{u}$, and divergence of the iteration scheme was always the result.

The only constraint imposed by Eqs. (19b,c) is the restriction of parameter choices to those which produce a well-posed solution to this differential system. In practice, this is a trivial constraint, since a solution exists for virtually all reasonable choices of $\mathbf{u}$. Thus, the total system (19) is analogous to a nonlinear algebraic

objective function, and similar solution methods can be used for each. All such methods require an initial parameter vector $\mathbf{u}^0$, and use it to start an iterative procedure whereby $\mathbf{u}^{k+1}$ is calculated if $\mathbf{u}^k$ is known. This involves the following basic steps,¹⁰ assuming the current parameter estimate is $\mathbf{u}^k$:

1. Test for optimality. If conditions for minimization are satisfied sufficiently, then stop.
2. Search direction. Compute a non-zero vector $\mathbf{p}^k$, the direction of search from the current estimate $\mathbf{u}^k$.
3. Step length. Compute or specify a positive scalar α which assures improvement in the objective function: $R(\mathbf{u}^k + \alpha\mathbf{p}^k) < R(\mathbf{u}^k)$.
4. Update. Set $\mathbf{u}^{k+1} = \mathbf{u}^k + \alpha\mathbf{p}^k$, go back to step 1.

As seen from step 4, the product $\alpha\mathbf{p}^k$ is the correction to $\mathbf{u}^k$ that produces $\mathbf{u}^{k+1}$.

Selection of the step length α is often more art than science, with only general criteria available for guidance. With convergence to a solution as the driving force, various procedures have been formulated to compute an appropriate value. Maintaining the descent of the objective function R is usually an important criterion, although it is not always necessary. In practice, an algorithm by Fletcher¹⁴ has proved to be quite useful, and has been used for all calculations presented in this study. A brief description is given in Appendix C.

A number of different procedures have been developed to select a correction $\mathbf{p}^k$, and they almost always depend on the gradient $dR/d\mathbf{u}^k$. Probably the simplest is the assignment

$$\mathbf{p}^k = -\frac{dR}{d\mathbf{u}^k}, \quad (21)$$

which is appropriately called the steepest-descent method, since $\mathbf{p}^k$ is oriented so that the instantaneous reduction in $R(\mathbf{u}^k)$ is the greatest. While this approach may

be quite successful initially, it is very inefficient as the optimal point is approached, and may not even converge in a finite number of iterations.^{10,15}

A number of improvements to the selection of $\mathbf{p}^k$ have been suggested which exhibit enhanced convergence under certain conditions. Two classes of methods that have received considerable attention are the conjugate gradient and quasi-Newton (or variable metric) methods.^{10,14-16} In particular, the quasi-Newton method labelled BFGS (for the independent developers Broyden-Fletcher-Goldfarb-Shanno) has superior convergence properties and is usually preferred by most practitioners. Several other methods were tried, and found to give inferior convergence results when compared with the BFGS procedure. Details of the method and its implementation are given in Appendix C.

Because the BFGS procedure depends at each step on the gradient dR/du , it is advantageous to calculate this quantity as efficiently as possible. A procedure is developed in Appendix D to accomplish this by solving one additional system of differential equations, similar to Eqs. (19a, b).

A computer program has been written to perform optimization calculations according to the procedure outlined in this section. A brief description and several flow charts are given in Appendix I. The initial estimate of the parameter vector $\mathbf{u}^0$ is supplied by the user in addition to various other controlling parameters. The first correction steps use the steepest descent method (21), with an automatic switch to the BFGS method occurring when the steepest descent loses efficiency. For both methods, the steplength α is chosen using the algorithm in Fletcher.¹⁴

The BFGS method is fairly successful at converging to local minima, however in certain situations it will converge to a value that is not an optimum. Thus, when the method appears to be nearly converged, it is necessary to check for true optimality

by restarting the algorithm. In some cases, either the initial steepest descent steps or the revised BFGS usage will produce further optimization.

5.4. Parameter Scaling

One further note regards the scaling of the parameters u_i , which was not addressed in Section 3, where scaling of the dependent and independent variables was discussed. If the u_i vary by many orders of magnitude, then $\partial R/\partial u_i$ might also. A change ΔR in the residual is then approximated by

$$\Delta R = \sum_i \frac{\partial R}{\partial u_i} \Delta u_i, \quad (22)$$

where the sum will be dominated by terms of highest magnitude. The really important question is not the absolute change in R related to the absolute change in a parameter, but rather the relative change $\Delta R/R$ due to the relative parameter change $\Delta u_i/u_i$. This can be accounted for in Eq. (22) by consideration of the sensitivity derivative $\frac{\partial R/R}{\partial u_i/u_i}$ instead of the partial derivative $\partial R/\partial u_i$:

$$\frac{\Delta R}{R} = \sum_i \frac{\partial R/R}{\partial u_i/u_i} \cdot \frac{\Delta u_i}{u_i}. \quad (23)$$

An equivalent way to approach this matter is to define the new variables

$$v_j = \ln u_j, \quad (24)$$

for which it follows that

$$\Delta R = \sum_j \frac{\partial R}{\partial v_j} \Delta v_j = \sum_j \frac{\partial R}{\partial u_j} \frac{\partial u_j}{\partial v_j} \Delta(\ln u_j) = \sum_j \frac{\partial R}{\partial u_j/u_j} \Delta(\ln u_j). \quad (25)$$

Now, from the series expansion

$$\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - + \dots,$$

the logarithmic term can be written as

$$\Delta(\ln u_j) = \frac{\Delta u_j}{u_j} + O\left(\frac{\Delta u_j}{u_j}\right)^2.$$

Substituting this into Eq. (25) and dividing by R produces

$$\frac{\Delta R}{R} = \sum_j \frac{\partial R/R}{\partial u_j/u_j} \left[\frac{\Delta u_j}{u_j} + O\left(\frac{\Delta u_j}{u_j}\right)^2 \right],$$

which is a first order approximation to Eq. (23). Thus, the logarithmic scaling (24) allows optimization on the new variable v_j to produce changes in the response that are nearly the same as if the relative changes $\Delta u_j/u_j$ and sensitivity derivative were used. In effect, the transformation (24) allows more uniform treatment of parameters within the optimization process, and is a necessary step to allow optimization relative to all parameters (see ref. 14).

6. APPLICATION AND RESULTS OF OPTIMIZATION PROCEDURES

6.1. Obtaining Initial Parameter Estimates

The mathematical optimization procedures described in Section 5 are precisely formulated and can be rigorously applied to find local optimal values in the neighborhood of the initial guesses. To achieve global convergence, it is important that the initial guesses themselves be reasonable, and their selection is in many ways the most difficult part of the process. The initial values ultimately used in this study resulted from considerable effort, including the use of trial-and-error, heuristic mathematical reasoning, and an understanding of the underlying chemistry.

Most important in refining the choices for parameter values are the principles of chemical kinetics. The Arrhenius form is commonly used to describe the temperature dependence of rate coefficients:

$$k = A e^{-E/RT}, \quad (26)$$

where A is the frequency (or pre-exponential) factor, and E is the activation energy for the reaction. For a bi-molecular reaction in aqueous solution, transition state theory suggests that an approximate value for A is $10^{13} \text{ M}^{-1}\text{s}^{-1}$, provided that both interacting species are not charged.¹⁷ For species of like charge, this value should decrease by one order of magnitude for each charge placed on a molecule; for species of opposite charge, A should increase analogously. In practice, actual frequency factors may vary from this ideal by several orders of magnitude. For tri-molecular reactions, further deviation from this theory should be expected. The activation energy is a measure of likelihood that the reaction will occur once the two molecules do interact. A value of $E=0$ suggests that the reaction will occur spontaneously, while a value of $E=100 \text{ kJ}$ implies that reaction is unlikely. In

general, it is reasonable to expect activation energies between 10 and 60 kJ, and most likely occurring between 20 and 40 kJ.

If the frequency factor is considered to lie in the range $10^{11} \leq A \leq 10^{14}$ and activation energy in the range $20 \text{ kJ} \leq E \leq 40 \text{ kJ}$, then at 25°C the rate coefficient will, satisfy:

$$9.2 \leq \ln k \leq 24.2 \quad (27)$$

This gives some idea of the likely range, although it is still quite large, and will require scoping calculations to reduce it further. Furthermore, this range should not be enforced too strictly, since it is possible that some rate constants will fall outside of it. Nevertheless, it is a good place to start, and trial calculations for both systems (10) and (12) use values obtained from this range.

6.2. Optimization at 25°C

As a first step, optimization is performed using only data at 25°C. In this case, the rate coefficients are actually rate *constants*, *i.e.*, there is only one free parameter for each reaction. [The form of Eq. (26) indicates there are two free parameters— A and E —when temperature is not constant.] From Tables 2 and 3, it is seen that much more data exists at 25°C than at any other single temperature. These room temperature data should also be considered more reliable, since they are less likely to include significant evaporative losses or other high temperature anomalies.

Using a variety of initial values, mathematical optimization of systems (10) and (12) was performed, as detailed in Appendixes F and G, respectively. After convergence was nearly achieved for rate constants alone, optimal initial inventories were also included according to the procedure outlined in Appendix E. The final results at 25°C are shown in Tables 5–7.

Table 5. Optimal Rate Constants at 25°C

i	Reaction	$k_i(\underline{M}^{-1}s^{-1})^a$	$\Delta k_i/k_i$
(a) Rate Constants for Six-Constant Model			
1	Eq.(5a)	707.2	.088
2	Eq.(3a) reverse	4.394×10^4	.264
3	Eq.(3b) forward	2.408×10^5	.664
4	reverse	3.086×10^4	.680
5	Eq.(3c) forward	1.180×10^4	.635
6	reverse	8.989×10^6	.728
7	Eqs. (5g), (B.4)	.5383	.392
(b) Rate Constants for Ten-Constant Model			
1	Eq.(5a)	1883	.069
2	Eq.(4a)	158	.25
3	Eq.(4c)	1.62×10^{10}	8.3
4	Eq.(4d)	1.62×10^{10}	4.9
5	Eq.(5e)	4.42×10^{10}	3.9
6	Eq.(4f)	7.61×10^4	0.16
7	Eq.(4g)	6250	0.89
8	Eq.(4h)	1.81×10^7	3.0
9	Eq.(4i)	1.81×10^7	5.2
10	Eq.(4j)	5.75×10^8	0.15
11	Eq.(5g)	0.216	0.27

^aunits for $i=2,4,6$ in (a) and $i=5, \dots, 10$ in (b) are $\underline{M}^{-2}s^{-1}$

Table 6. Optimal Initial Inventories at 25°C Using Ten-Reaction Model
(g-atom/dm³)×10⁴

Transient	I ⁻		IO ₃ ⁻		I ₂	
	Expt	Calc	Expt	Calc	Expt	Calc
1	1.0	0.99	15.38	15.20		
2	1.0	1.03	9.43	8.13		
3	5.42	4.95	0.193	0.234		
4	5.22	5.03	0.979	0.836		
8					1.2	1.117
9					1.2	1.177
10					1.2	1.135
11					1.2	1.201

Table 7. Error Residual Comparison*

	Residual	<i>I</i>	Residual × <i>I</i>	Residual × <i>I</i> ²
This work (10-reaction model)	.3586	8	2.868	22.95
This work (6-reaction model)	.9958	6	5.975	35.85
Ref. 5	12.98			
Ref. 6	12.93			
Modified 6-reaction model**	6.139			

*The number of dependent variables in the final model is *I*.

**Using forward reaction of Eq. (3a) instead of Eq. (5a).

From Table 5(a), it is clear that the rate constants are generally quite different from those reported by other researchers (cf. Table 1 on page 5). Due to the use of Eq. (5a) instead of Eq. (4a), k_1 is not comparable between the two tables; however, other values $k_2 \dots, k_6$ are. It is also possible to compare certain values in Table 5(a) with their analogues in Table 5(b), specifically those noted below:

Six-Reaction ModelTen-Reaction Model

$k_1=707.2$

$k_1=1883$

$k_2=4.394 \times 10^4$

$k_{10}=5.75 \times 10^8$

$k_7=0.538$

$k_{11}=0.216$

The first two constants in each column represent the same forward and reverse reactions of the primary hydrolysis process. The last in each column represents the tri-iodide equilibration rate constant. The first and third are reasonably close (certainly within the same order of magnitude), but their difference does illustrate the effects of other equations in the two different models. The second entries differ markedly, due to fundamentally different contributions made to the overall models. This can be seen from the original reaction systems (3) and (4) on pp. 4-6, where the reverse reaction of Eq. (3a) is compared to Eq. (4j). While both represent reverse hydrolysis reactions, Eq. (3a) is the *only* path in the six-reaction model for the formation of I_2 from other species. However, in the ten-reaction model, five different reactions can form I_2 from I^- and various intermediates. Thus, Eq. (4j) is not in such a unique position as Eq. (3a, reverse), and it is therefore reasonable to expect their rate constants to differ considerably.

The relative errors $\Delta k_i/k_i$ shown in the last column of Tables 5(a) and 5(b) were calculated according to the procedures described in Appendixes F and G. These generally represent the fractional change in parameter value necessary to produce a 1% change in the error residual. The value of 1% change was selected since this is the estimated uncertainty in the residual due to calculational and numerical error (the "region of indifference," see ref. 20). Such relative errors can be regarded as parameter uncertainties for the particular optimal solution shown. Other optima

may also exist, but these parameter uncertainties are not necessarily related to either the location or the uncertainties of other optimal solutions.

The parameter uncertainties are a measure of the ability of the data to adequately predict parameter values. Thus, for the data used in this study, certain parameters were predicted well, while others were not. This is most obvious for the ten-reaction model, where rate constants k_3 , k_4 , k_5 , k_8 , and k_9 have much larger uncertainties than the other constants do. Changes in these values of several hundred percent will not alter the error residual appreciably. This can be explained by referring to the reactions in system (4) and recalling that data measurements were made only for I^- , I_2 , and IO_3^- . Rate constants k_3 and k_4 involve reactions of species for which no data are available. For k_5 , only the product HIO_3 (in practice, the ionized form IO_3^-) was measured. Although all reactions (4f)–(4j) involve I^- as a reactant and I_2 as a product, these species are generally present in great excess compared to the other iodine species [except for iodate in Eq. (4f)]. This is increasingly the case as stepwise reduction of IO_3^- occurs. Thus, the dominant reactions in the Dushman reaction process are (4f) and (4g), with the others playing a minor role. Except for Eq. (4j), which represents the natural reversal of Eq. (4a), none of these five reactions plays a dominant role in forward hydrolysis. Evidently Eqs. (4h) and (4i) are not very important in either case. Furthermore, their rates are controlled by intermediate species which are not measured. In summary then, the presence of large parameter uncertainties is directly related to the absence of data which would verify both the concentrations of reacting intermediate species and the validity of the reactions in which they take part. Without data for these intermediate forms, nothing more can be done to verify reaction models or determine certain rate constants with high degrees of confidence.

Optimal calculated values for initial inventories of the ten-constant model are given in Table 6 in the columns headed by 'Calc.' These can be compared with the average experimental values, which appear in Table 4 and also in Table 6 under columns marked 'Expt.' All changes are less than 20%, and most fall in the range 5–12%. It should be noted that the error residuals generally fell 10–40% when initial inventories were included in the optimization.

6.3. Calculation of Species Concentrations

Using the optimal rate constants given in Table 5b, iodine species for the ten reaction model [representing reactions (4), with modifications (5a, e, g)] are calculated by solving the rate equations (12). Calculated dimensionless concentrations for all species (except I_3^-) are shown in Figure 1 for three of the transients described in Table 2-3 on pp. 16-17. Comparison of the three stable species I^- , I_2 , and IO_3^- with data is made in Sect. 6.5 after optimization of activation energies has also been accomplished. (The calculated values of these principal species are very nearly identical in two cases.)

While principal concern is rightly focussed on the three stable species for which data exist, the intermediate species occupying oxidation states 1, 2, 3, and 4 do play an important role. As seen for each transient in Figure 1, curves for certain intermediates are nearly parallel to each other after an initial transient period. Since these are logarithmic plots, this indicates that the ratio of concentrations for such species is constant. This is reasonable, since it is expected that highly reactive intermediates may quickly reach quasi-equilibrium with each other and with the major species. In certain cases, it might be possible to simplify calculations by assuming these equilibria hold throughout, thus reducing the number of species

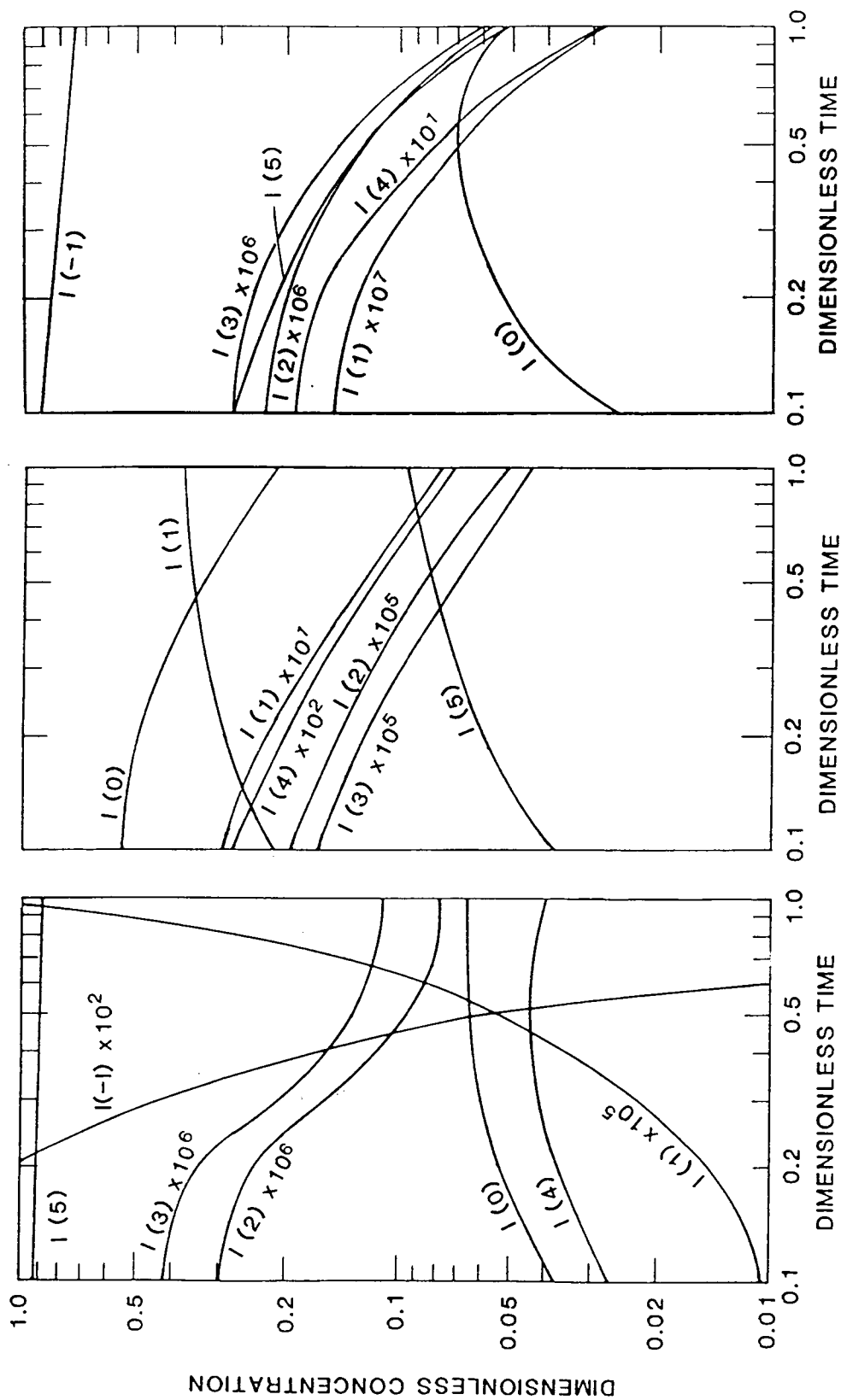


Figure 1. Species concentration for three representative transients

and reactions to be considered. In general, however, this approximation is not appropriate, since

1. No pair of species concentrations always forms a constant ratio over all conditions studied. While all intermediates (*i.e.*, oxidation states 1-4) form constant ratios at pH=7 (*cf.* Fig. 1b), only states 2 and 3 do at pH=2.8 (Fig. 1a). In Transient 3 (pH=3.2), all intermediates exhibit the same general shape, but none is really parallel to another.
2. Even when two species concentrations do have constant ratio, this ratio itself changes from transient to transient. It is a complex function of pH and concentrations of other species present.

Thus, the transient roles of intermediates are important in characterizing the overall modelling results. While they could be replaced by equilibria in some cases, they should be retained in the kinetic model for the general case. This is especially true if the model is operated under conditions of changing pH, temperature, and even total iodine concentration. Similar conclusions can be drawn for the six-reaction model.

6.4. Comparison of Different Models

At this point, it is useful to compare the abilities of various models to match data. This is done quantitatively in Table 7, where error residuals for five different models measure their simulations of data at 25°C (transients 1-4 and 8-11 on pages 16-17). The first two lines represent the best results obtained for the two models considered in this work. The last two lines utilize the hydrolysis models appearing in refs. 5 and 6, namely the reaction system (3) using rate constants from Table 1

on page 5. Simulation of the various room-temperature transients by each of these models yielded the error residuals shown. The magnitude of these last two residuals (in comparison to the first two) indicates their ineffectiveness in simulating the data.

The first two lines of Table 7 represent models that have been optimized relative to the data of Section 4; lines 3 and 4 represent models whose parameters were chosen by other criteria, perhaps by comparison with other data at different conditions. Since the residual error in Table 7 is determined by comparison with data in Section 4, it is appropriate to attempt improvement for these two models relative to the new data in Tables 2 and 3. To accomplish this, the parameter values in Table 1 were used as initial guesses, and the optimization procedures of Section 5 applied to each model in turn. The resulting rate constants were not very different from the initial values, and the error residuals improved by less than 5%. This is likely because the rate constant values in Table 1 are poor initial guesses, although the difference in the actual models [use of original Eq. (3a) instead of Eq. (5a)] may also be a factor.

In order to test the latter possibility, numerous trials were conducted using the forward reaction (3a) instead of Eq. (5a). The best residual after optimization is shown in the final row of Table 7. This residual error is considerably higher than that of the preferred six-reaction model shown in row 2. Thus, the model using Eq. (5a) appears to be superior to that using the forward step of Eq. (3a); when used with the rate constants in Table 5(a), this model exhibits a considerable improvement over previous models represented by Eq. (3a) and parameter values in Table 1.

From Section 2.4, several variations were also possible in the ten-reaction model. Calculations using reactions (5a) and (5e) gave consistently better results than

those using reactions (4a) and (4e). For this reason, the former were regarded as producing a superior model and used to produce all results shown in this paper.

Also from Table 7, it is possible to compare the six- and ten-reaction models whose rate constants have been optimized in this study. As one would expect, the ten-reaction model has a lower error residual, due to the presence of more free parameters. However, even the six-reaction model gives a reasonable description of system behavior, which qualitatively reflects the data quite well. The judgement of which is "better" is quite subjective, and will depend in part on the particular use. For example, in a radiation environment, the ten-reaction model provides a more complete list of intermediate species; certain of these may be necessary as reactants or products in reactions with water radiolysis products.

If computational complexity (i.e., time and expense) is proportional to the dimension I of the system to be modelled, then one measure of a method's "goodness" is the product

$$I \times \text{Residual.}$$

Where matrix operations are concerned (such as ours), computational complexity may be proportional to I^2 , implying that "goodness" is measured by

$$I^2 \times \text{Residual}$$

For the two systems considered here, $I=8$ and $I=6$, representing the number of dependent variables in Systems (12) and (10), respectively. Under either criterion, the last two columns of Table 7 indicate that the ten-reaction model is superior.

6.5. Temperature Dependence of Rate Coefficients

Optimization in the previous section utilized data only at room temperature to determine a single value for each chemical rate coefficient. From Eq. (26), it can

be seen that if temperature is not constant, then the value of k is not constant, but obtained from the Arrhenius expression once A and E are known. Thus, in general, two free parameters must be determined for each chemical reaction.

Optimization of two parameters for each rate coefficient is accomplished using the theoretical basis outlined in Section 5. Details of the extension are described in Appendix H, where it is seen that the problem is closely related to the case at constant temperature.

The choice of initial parameter guesses is discussed generally in Section 6.1 and the actual process of finding good estimates at 25°C is discussed in Appendixes F and G. Because this was a laborious process even at constant temperature, it might prove to be overwhelming to find initial guesses when the number of free parameters is doubled. For this reason, it is helpful to use the rate coefficient values found at 25°C in order to provide initial estimates for the free parameters A and E in Eq. (26).

As mentioned in part A, it is reasonable to assume that frequency factors ($\text{M}^{-1}\text{s}^{-1}$) satisfy

$$10^{11} \leq A \leq 10^{14}. \quad (27)$$

Assigning a specific value to A and utilizing the rate coefficients calculated at 25°C, a corresponding value for activation energy can be determined by rearranging Eq. (26):

$$E = -RT \ln(k/A) \quad (28)$$

Shown in Table 8 are the results of several calculations, using various choices of A and the rate coefficients k_1 of the ten-reaction model. Repeating this process for each rate coefficient gives a set of reasonable candidates for initial guesses. Various

combinations can then be tried and used to calculate optimal parameters for each rate coefficient.

Table 8. Choices of Pre-Exponential Factor and Corresponding Activation Energy^a

$A(\text{M}^{-1}\text{s}^{-1})$	$E \text{ (kJ/mol)}$
10^{11}	44.1
10^{12}	49.8
10^{13}	55.5
10^{14}	61.2

^a assuming $k_1 = 1884 \text{ M}^{-1}\text{s}^{-1}$ at 25°C

An alternative approach for obtaining Arrhenius parameters would be to independently optimize at each temperature. Values for A and E could then be determined by estimating or fitting Eq. (26) to the three data points. This procedure was tried initially but proved to be quite unsatisfactory due to the empirical nature of the models and the imperfection of the data. In most cases, a successfully optimized solution at a higher temperature would yield some rate coefficients with values lower than their respective values at 25°C . This decrease with increasing temperature implied $E < 0$, which is not physically realistic. Using the values in Table 8 forced the optimization process to search for solutions which were consistent with the principles of chemical kinetics.

The best results for the ten-reaction model were achieved when an initial value of $A=10^{12}$ was used for all rate coefficients. This resulted in the lowest residual for which all values of A remained in the acceptable range (27) after the optimization process. The fully optimized parameter values are shown in Table 9, together with the initial guesses which produced them. Also shown in Table 9 are the rate

coefficients at 25°C calculated from the optimal values of A and E using Eq. (26). Most have not deviated appreciably from their values in Table 5(b).

Table 9. Optimized Arrhenius Parameters for Ten-Constant Model

i	Initial Guesses		Final Values		$k(25^\circ\text{C})$
	$A \times 10^{11}$	E	$A \times 10^{11}$	E	
1	10.0	49.77	1.50	45.40	1652
2	10.0	55.92	2.41	54.87	58.3
3	10.0	8.94	10.11	8.938	2.74×10^{10}
4	10.0	8.94	9.52	8.964	2.55×10^{10}
5	10.0	6.46	6.91	6.525	4.96×10^{10}
6	10.0	40.56	7.12	40.62	5.40×10^4
7	10.0	46.79	4.19	46.85	2570
8	10.0	26.14	10.95	26.09	2.92×10^7
9	10.0	26.21	9.81	26.21	2.50×10^7
10	10.0	18.48	7.43	18.86	3.63×10^8
11	10.0	72.25	0.545	72.69	0.00988

Using the Arrhenius parameters in Table 8, rate coefficients calculated at temperatures of 25, 50, and 90°C can be used to solve System (12) (the dimensionless rate equations for the ten-reaction model) and determine the inventories y_i for each species. This is done for each of the transients described in Section 4 where data exists.

The calculated inventories of I^- , I_2 , and IO_3^- are then shown in Figs. 1-15 as solid, broken, and dotted lines, respectively. The experimental data for each case are shown as distinct points. These figures illustrate that the model is generally quite good at reflecting iodine speciation under a variety of conditions. Calculations using System (10) (i.e., the six-reaction model) at 25°C were similar, although not quite as close to the data.

pH = 2.8, T = 25 C

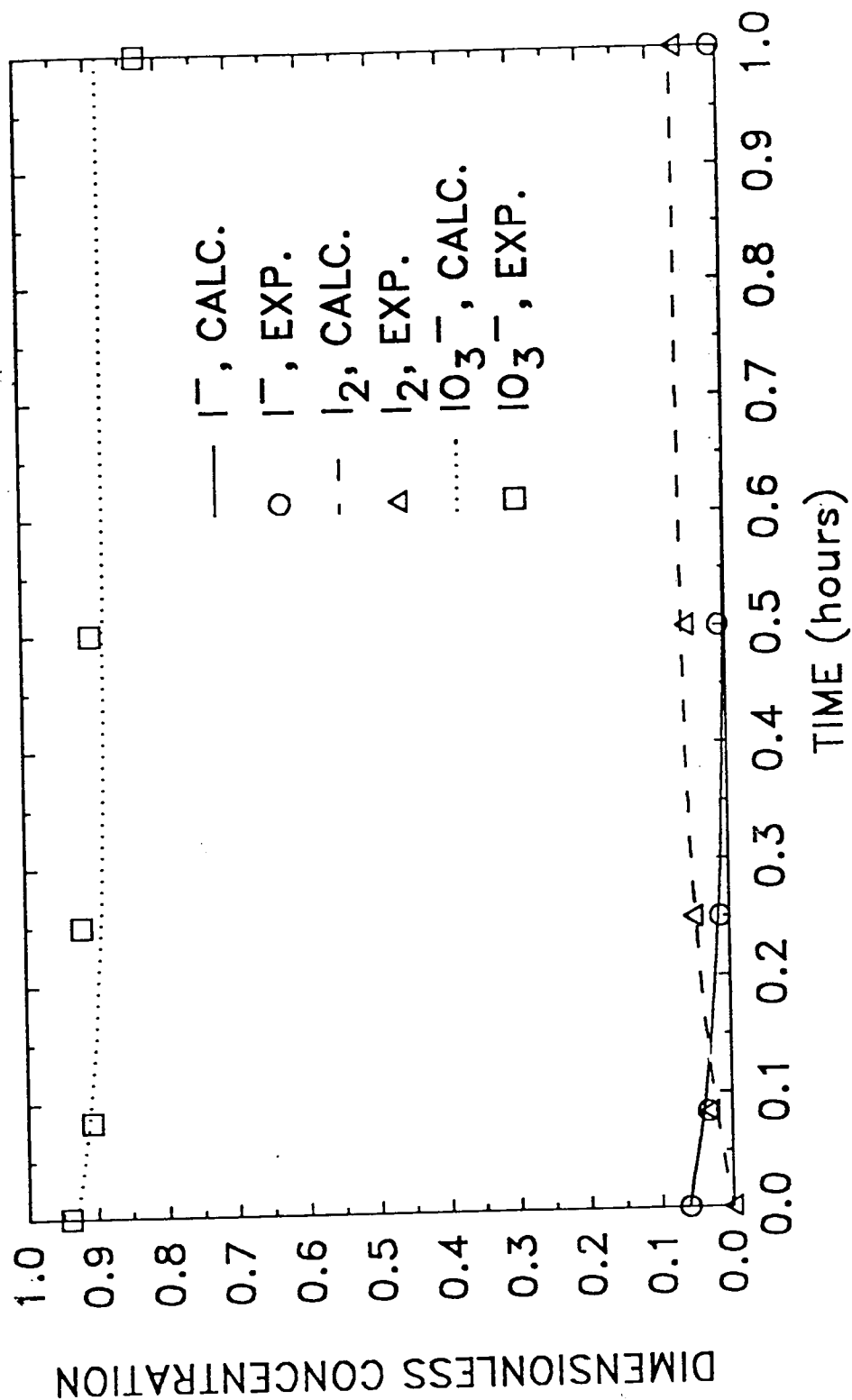


Figure 2. Transient 1 data and calculations.

pH = 3.0, T = 25 C

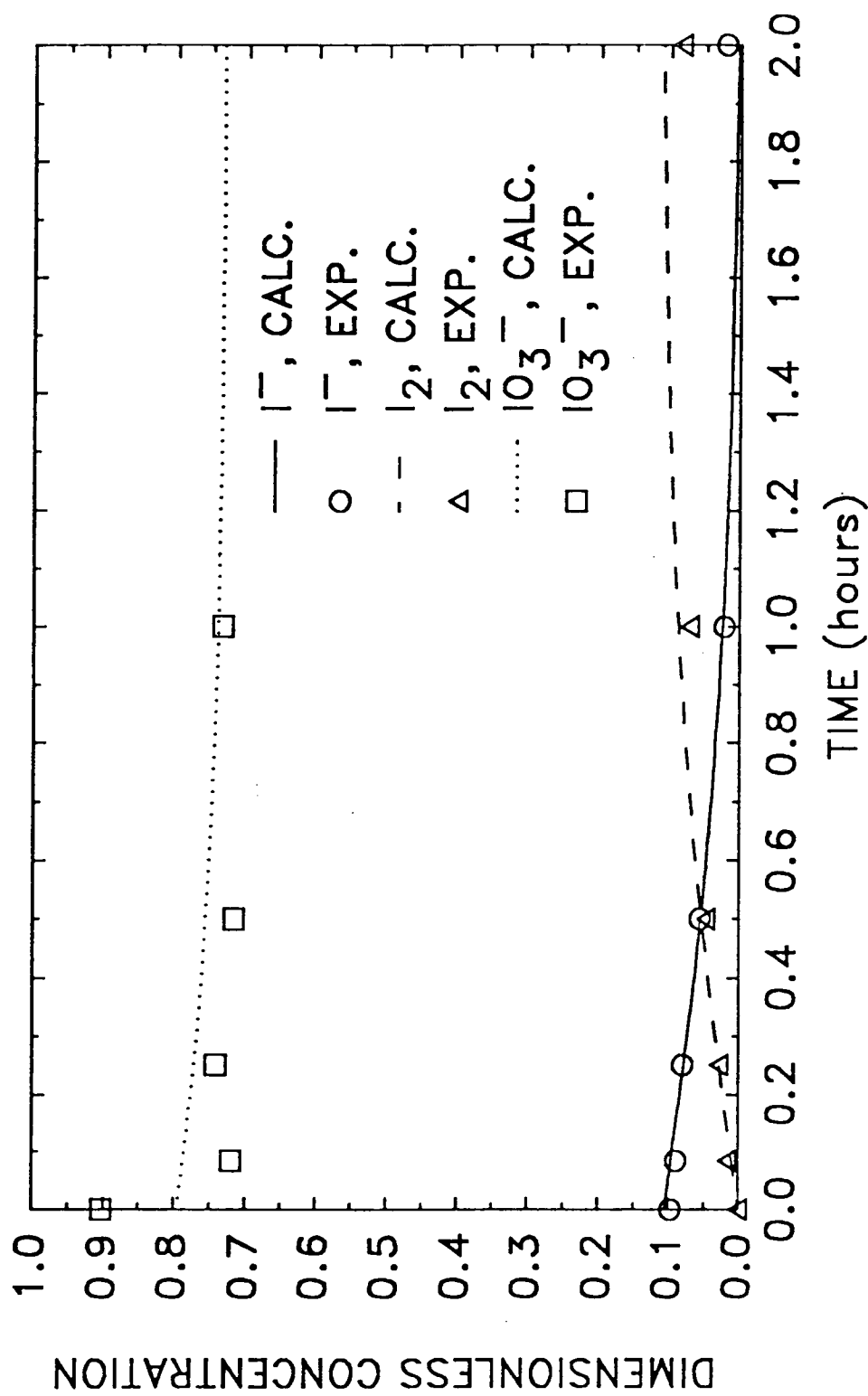


Figure 3. Transient 2 data and calculations

pH = 3.2, T = 25 C

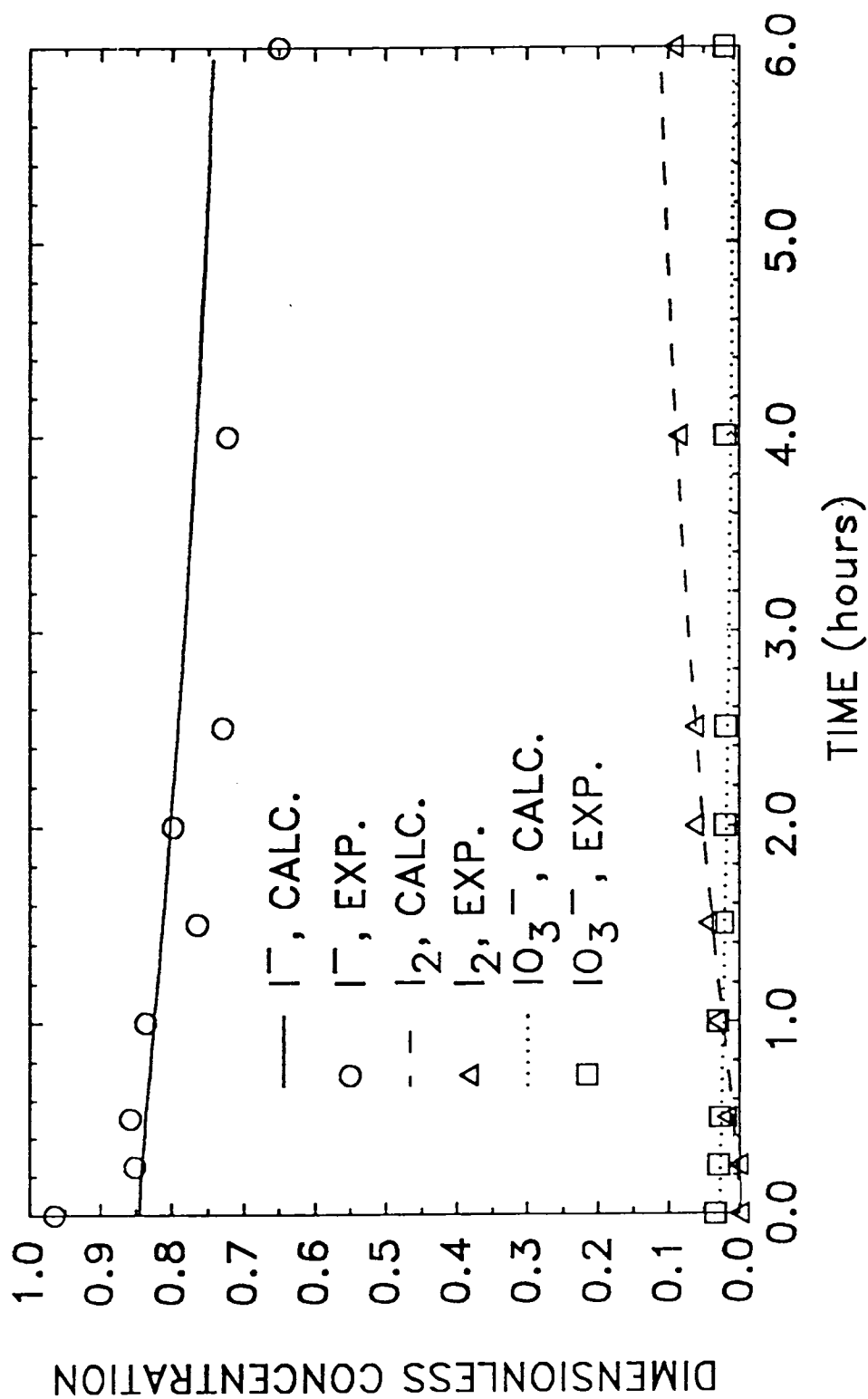


Figure 4. Transient 3 data and calculations.

pH = 3.2, T = 25 C

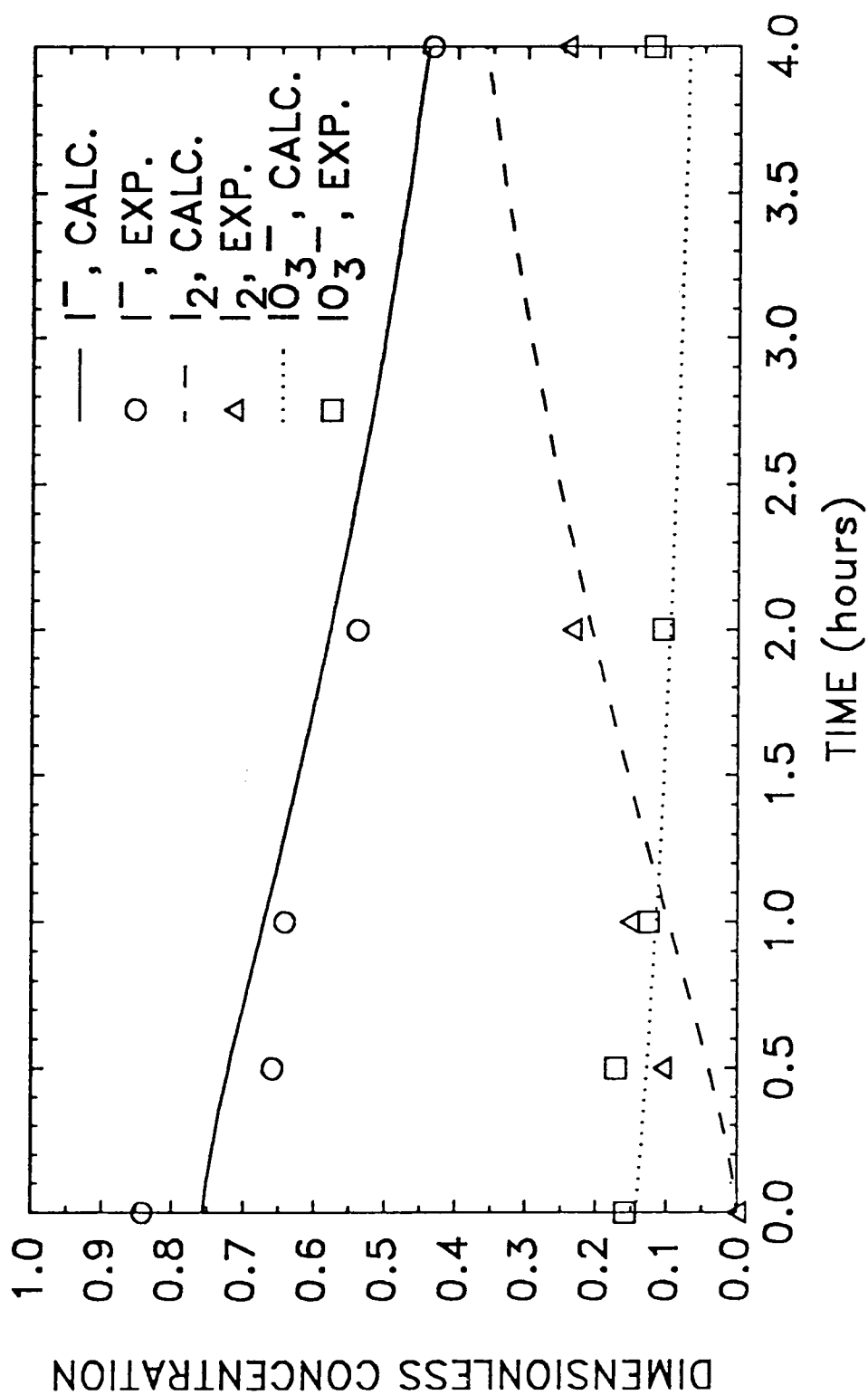


Figure 5. Transient 4 data and calculations

pH = 7, T = 25 C

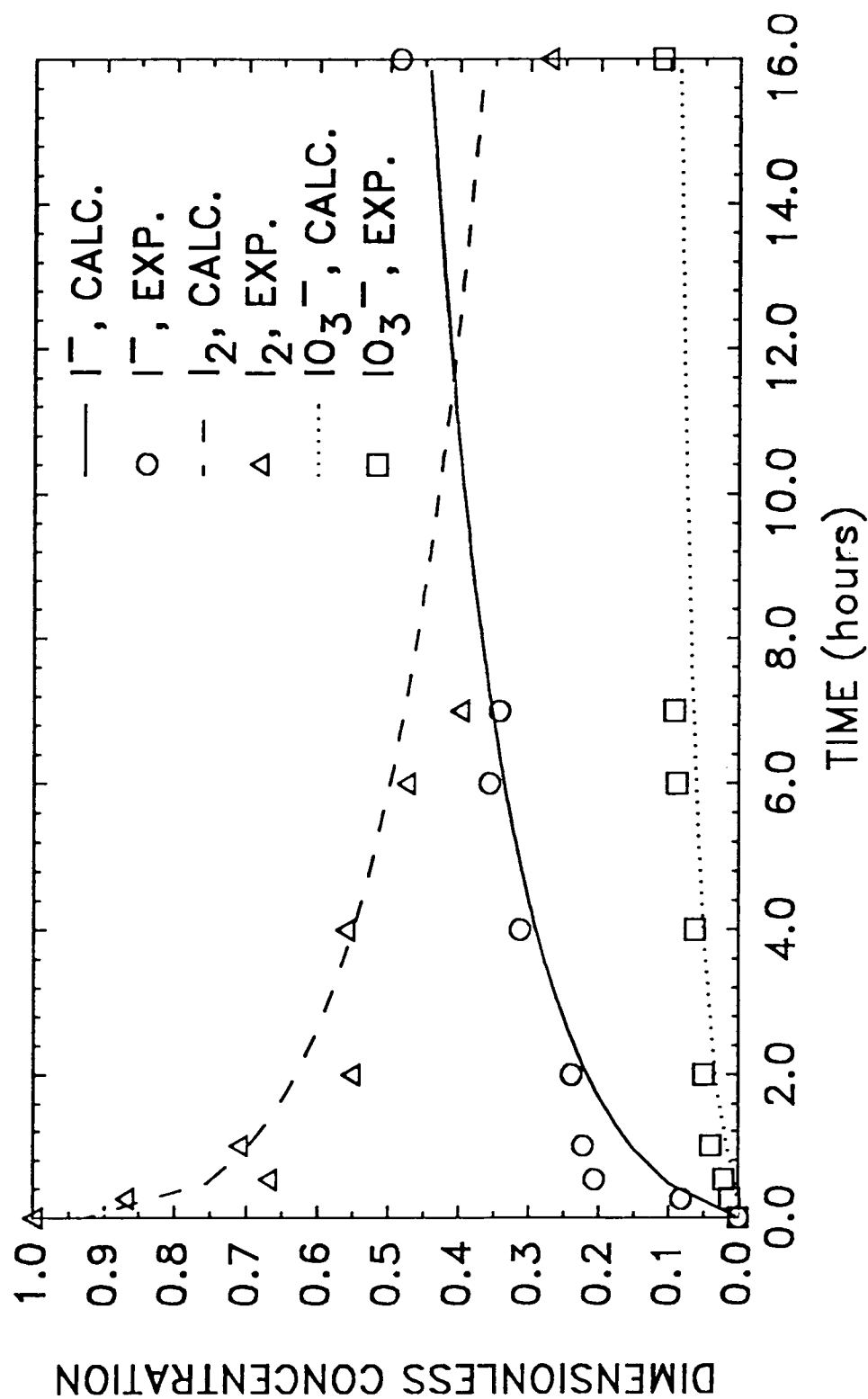


Figure 6. Transient 5 data and calculations.

pH = 8, T = 25 C

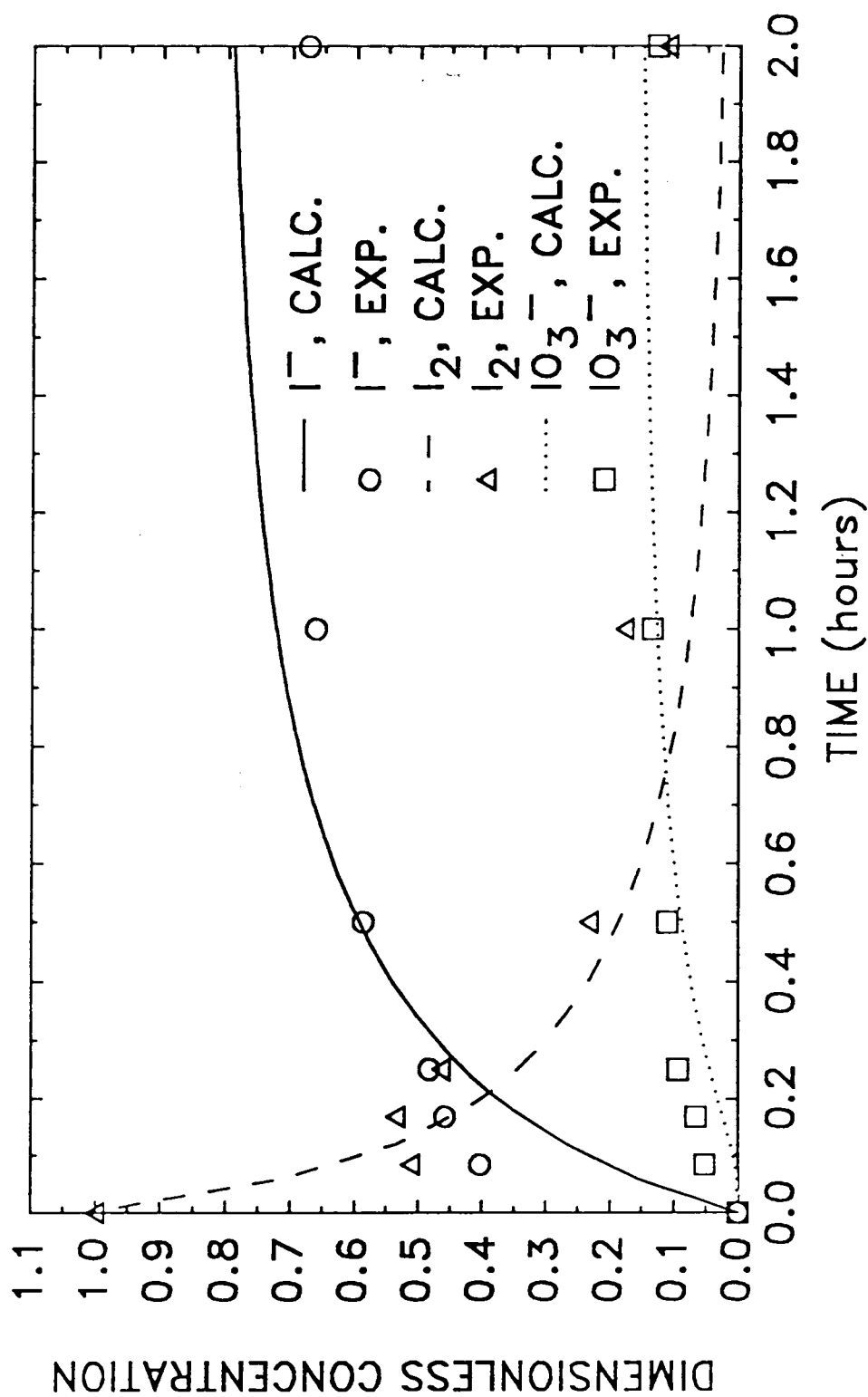


Figure 7. Transient 6 data and calculations

pH = 8.5, T = 25 C

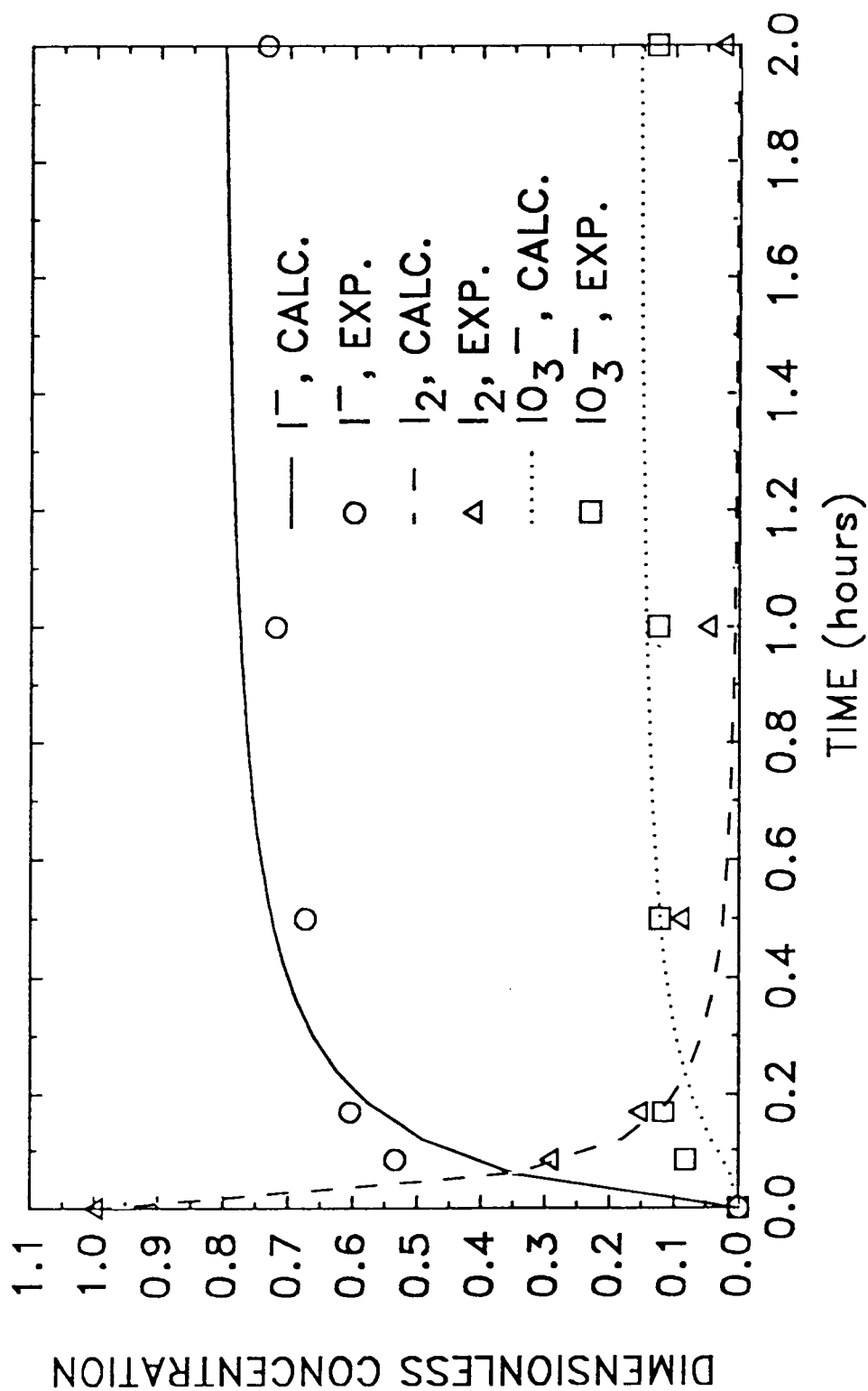


Figure 8. Transient 7 data and calculations.

pH = 9, T = 25 C

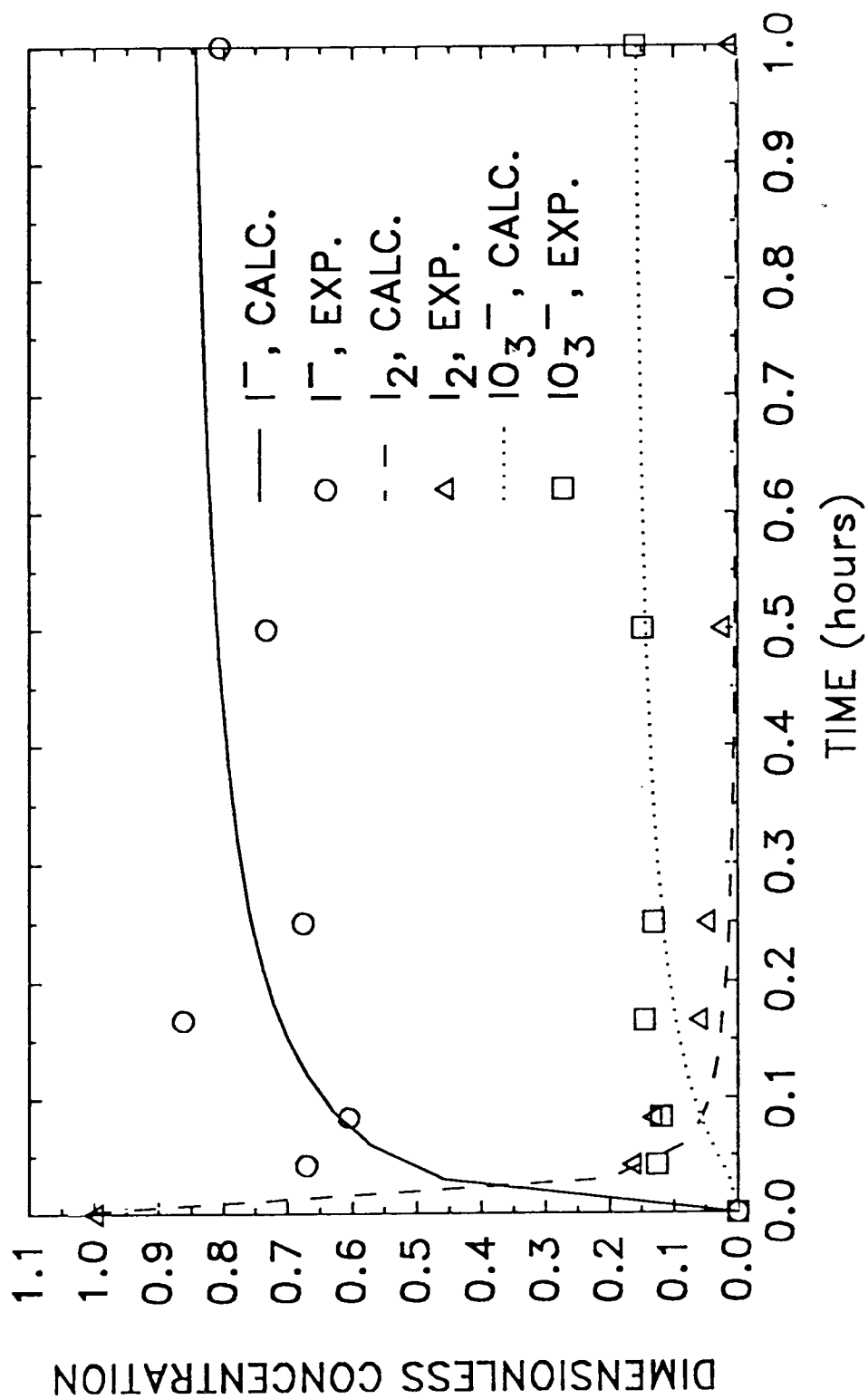


Figure 9. Transient 8 data and calculations

pH = 3, T = 50 C

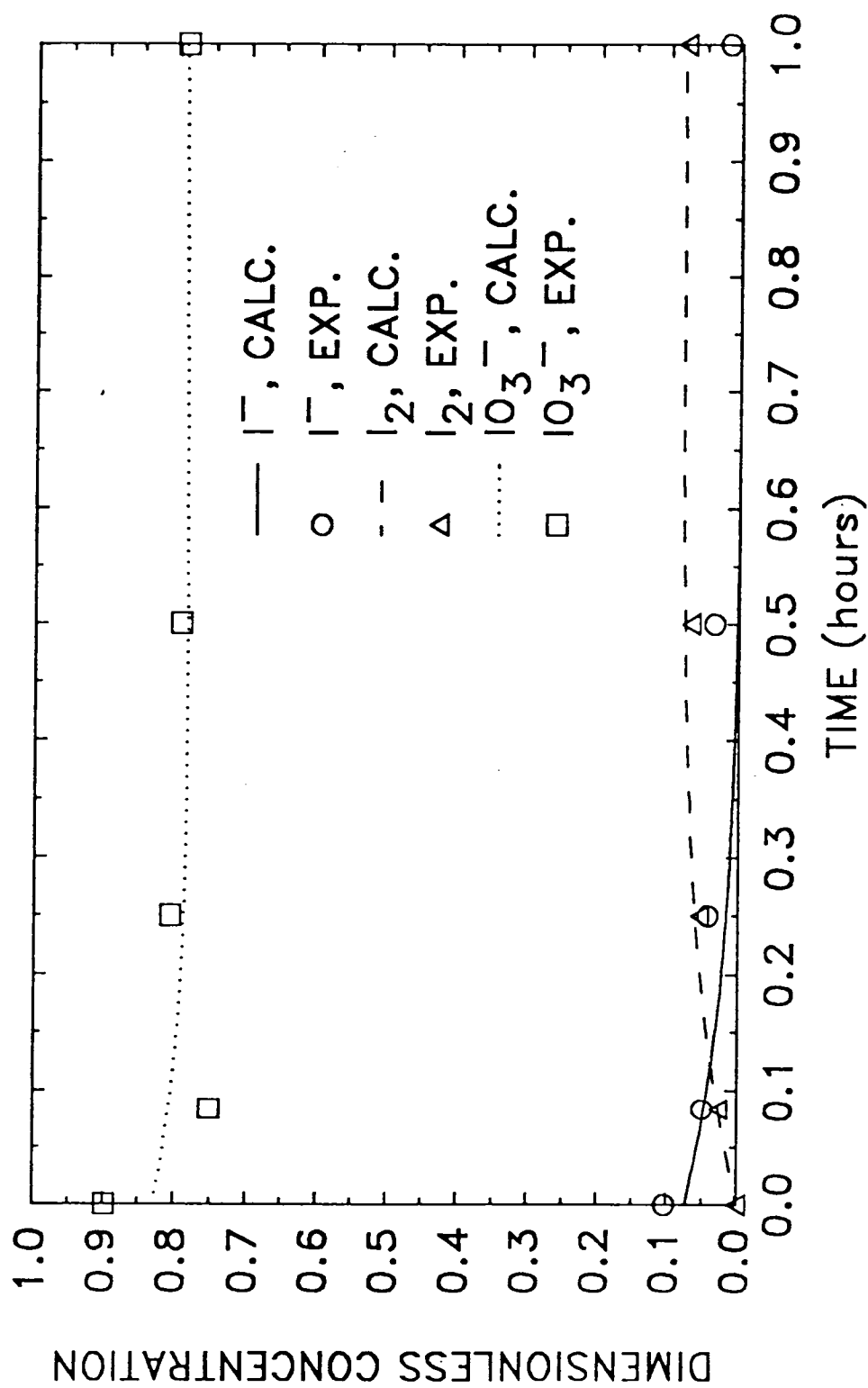


Figure 10. Transient 9 data and calculations.

pH = 3.2, T = 50 C

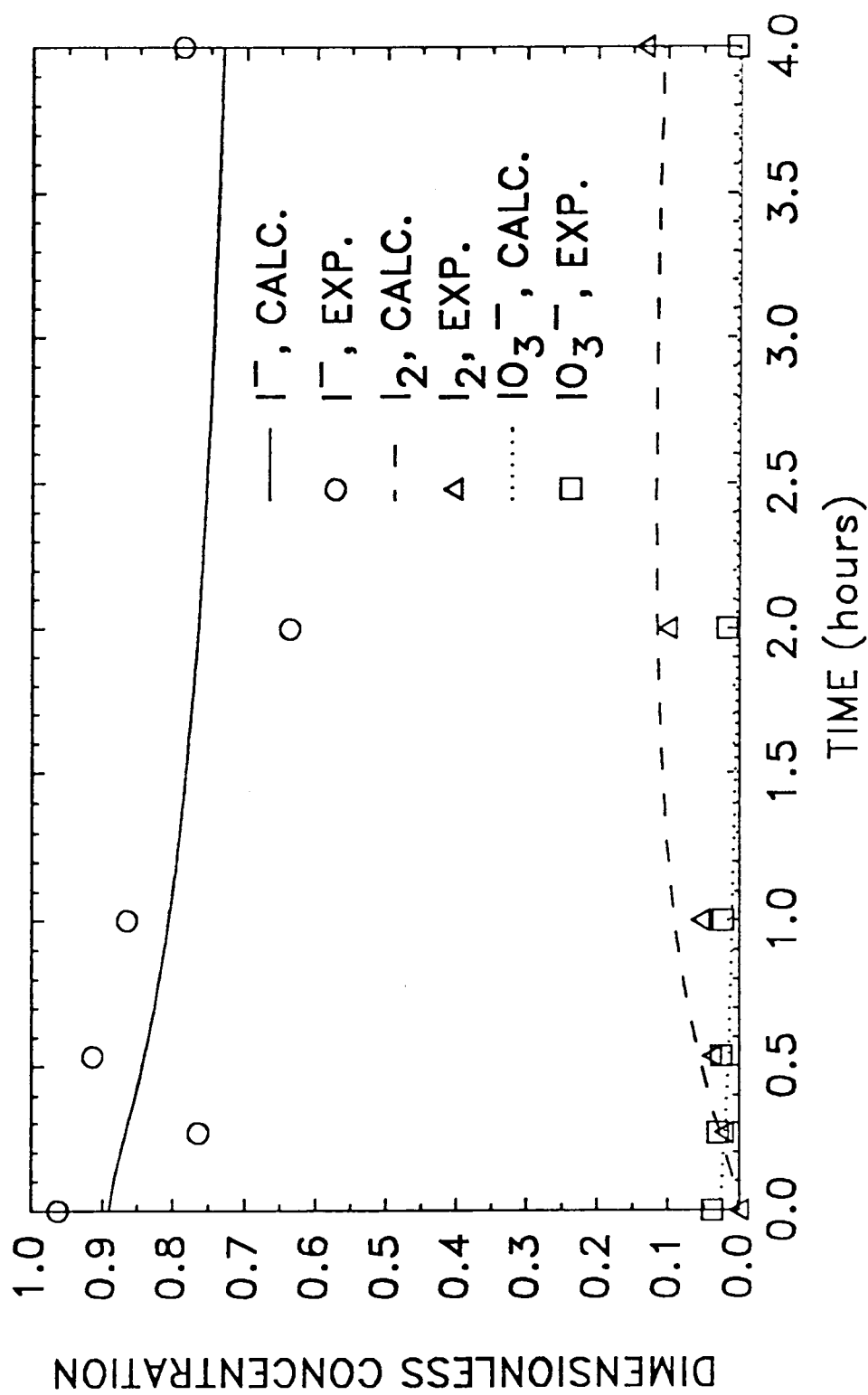


Figure 11. Transient I_0 data and calculations

pH = 7, T = 50 C

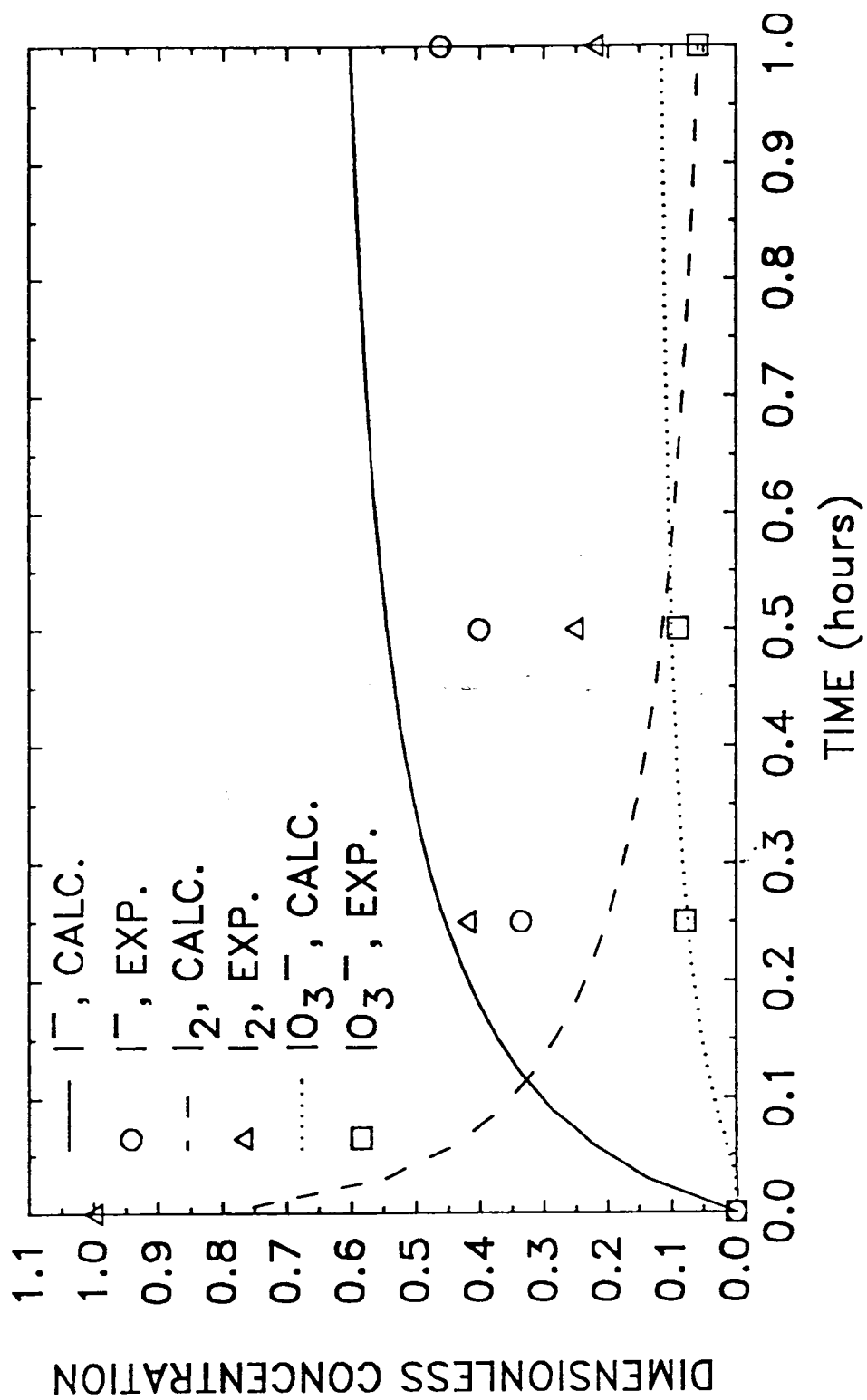


Figure 12. Transient 11 data and calculations.

pH = 3.2, T = 90 C

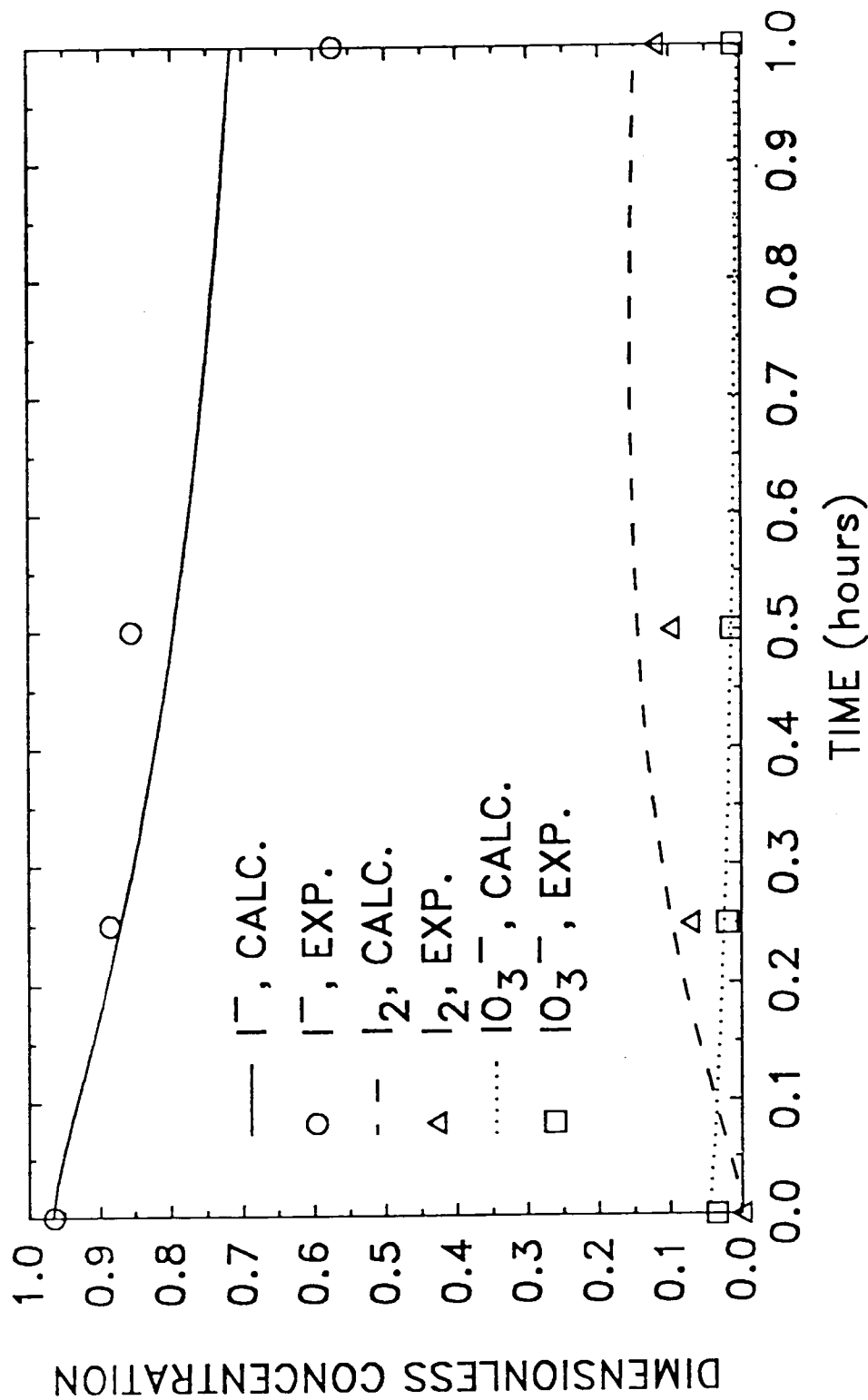


Figure 13. Transient I_2 data and calculations

pH = 5.5, T = 90 C

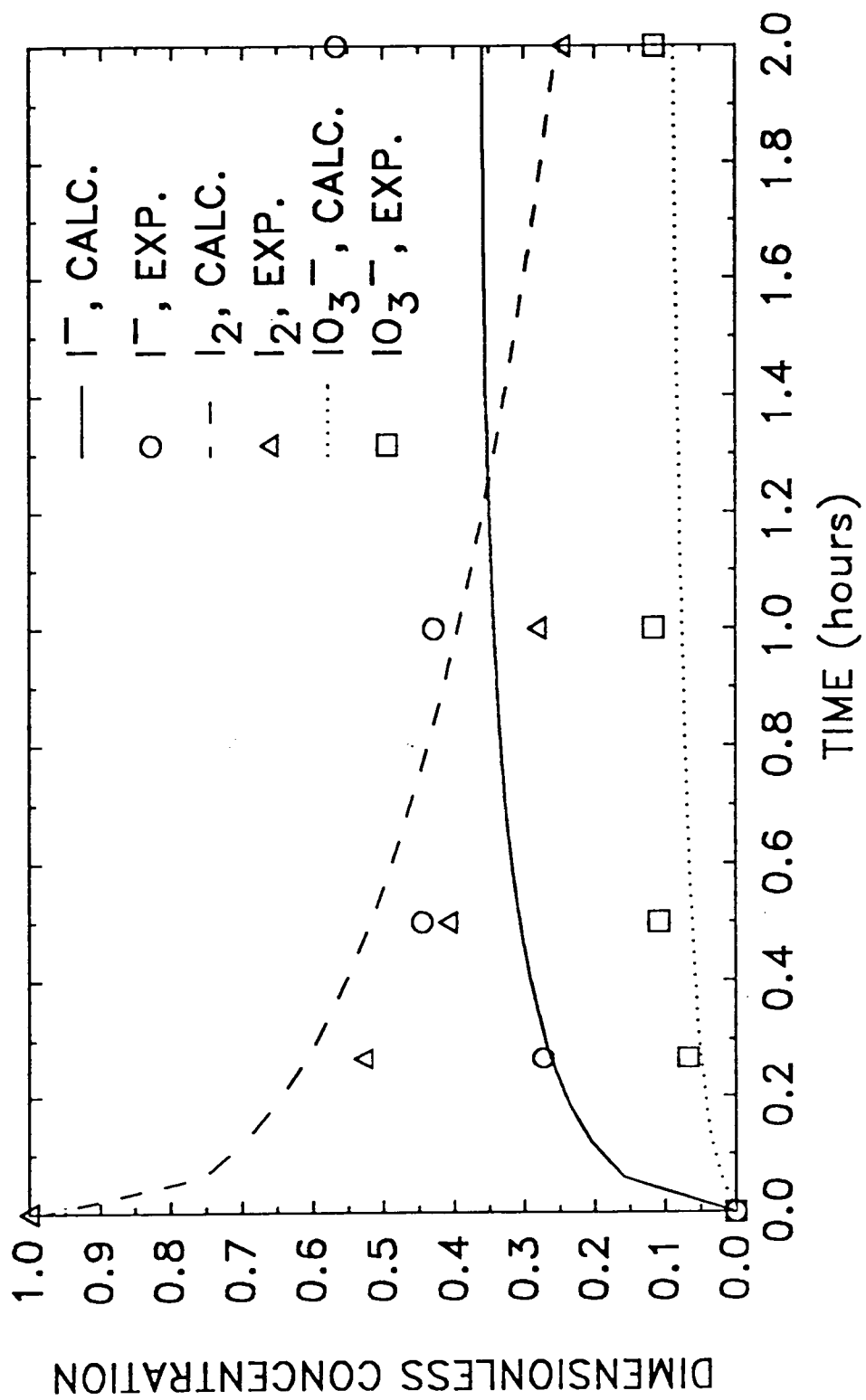


Figure 14. Transient 13 data and calculations.

pH = 6, T = 90 C

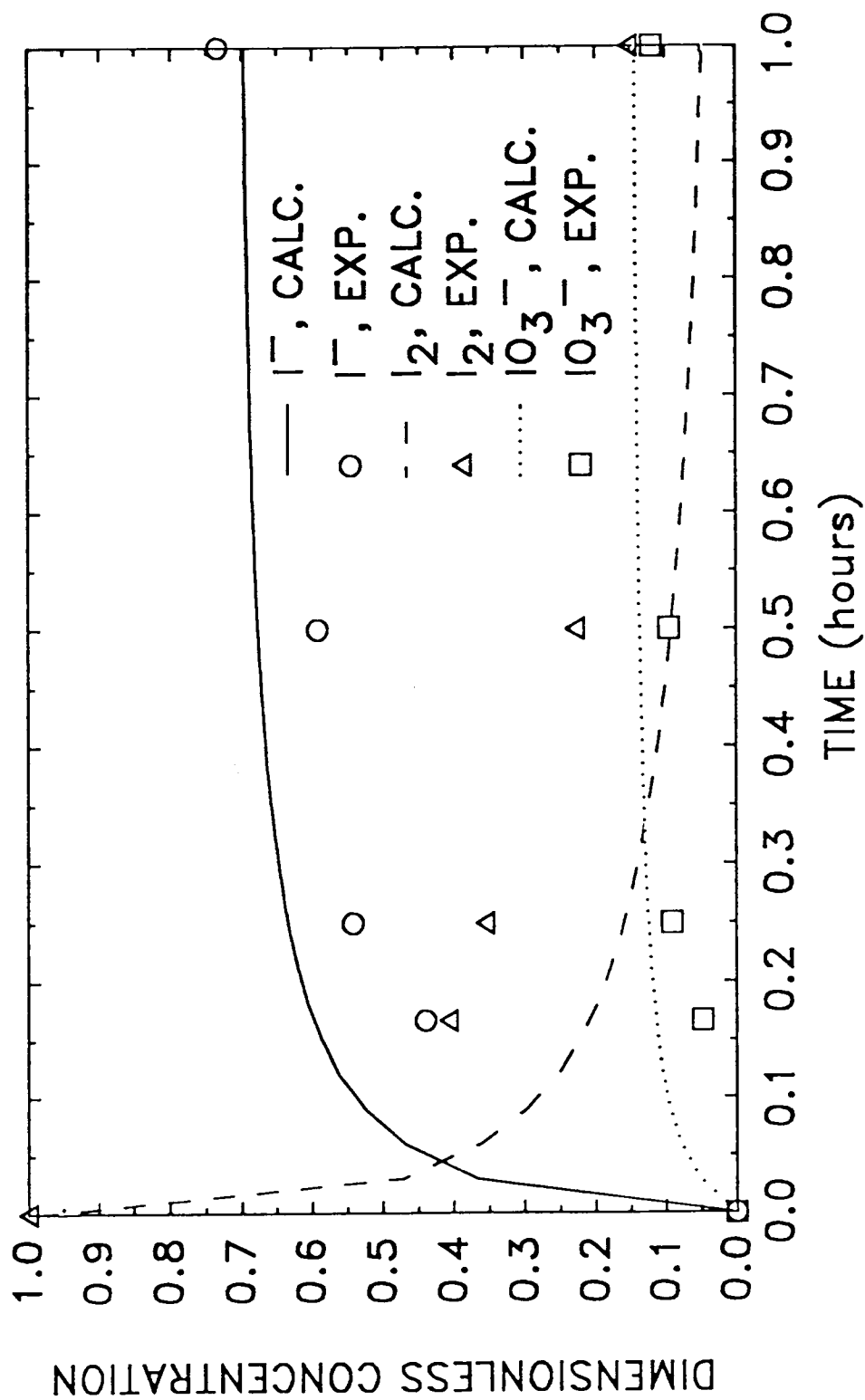


Figure 15. Transient 14 data and calculations

pH = 7, T = 90 C

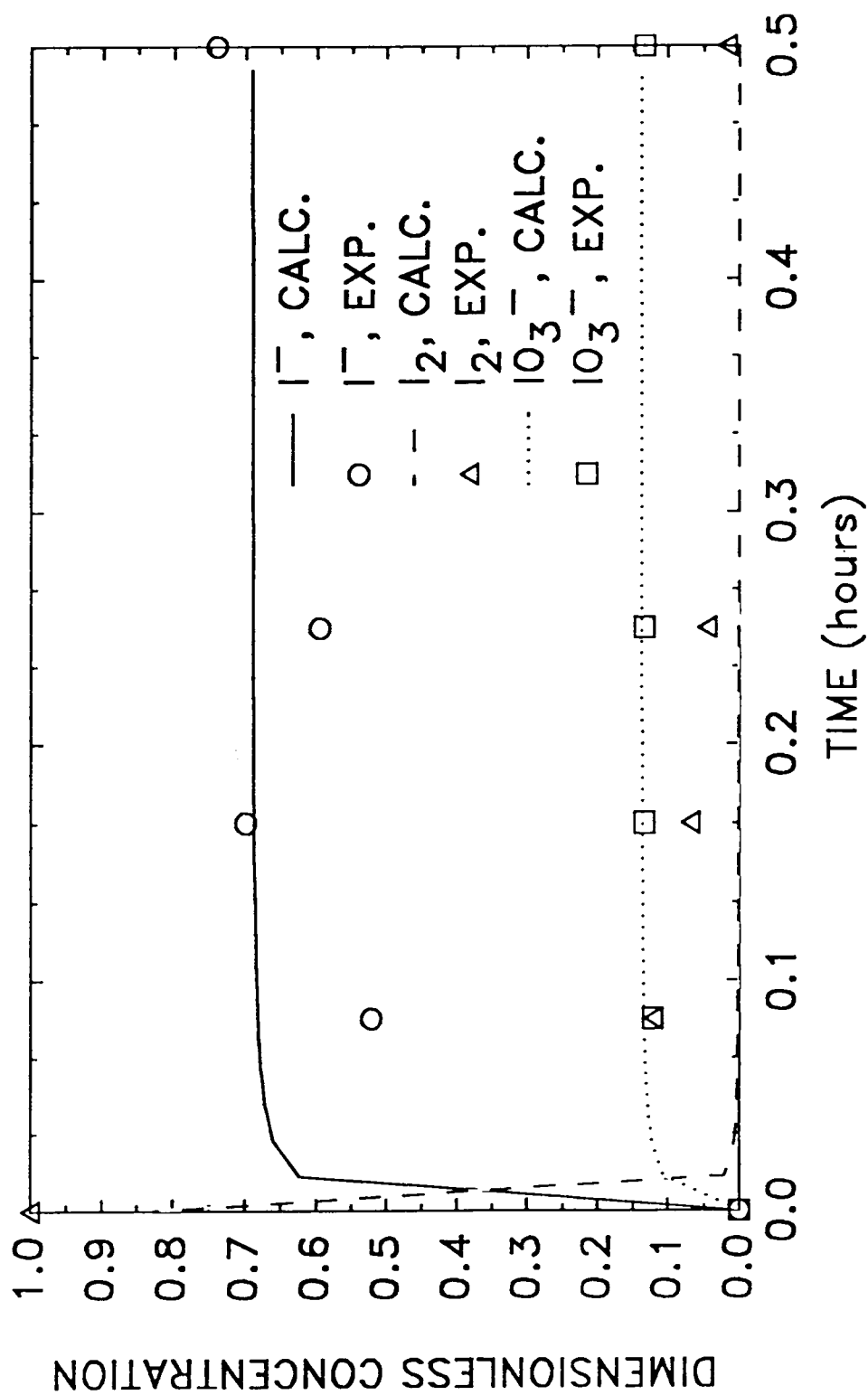


Figure 16. Transient 15 data and calculations.

7. SUMMARY AND CONCLUSION

The results of the previous section indicate that reactions sets (3) and (4) on pp. 4-6 can both be used as reasonable descriptions of aqueous iodine behavior. Since each represents a credible model, neither can be assumed to represent an actual mechanism, but rather a semi-empirical approach, whose rate coefficients are viewed as optimizable parameters. Furthermore, the various modifications such as Eqs. (5a), (5e), and (5g) indicate that the best form to use is far from certain, but must be continuously re-evaluated. Thus, the problem of modeling iodine hydrolysis is still far from resolved.

In spite of the many questions that still exist, the procedures described in this work do result in useful models. Although empirical in nature, the models do reflect reasonable expectations of chemical interactions. The rate coefficients obtained are both reasonable and useful, and the species represented are conducive to further development of radiolysis reactions. The overall models developed within this study were demonstrably superior to previous models in simulating the data of Section 4.

In comparing the usefulness of systems (10) and (12), it is expected that system (12) would give a better approximation to data, i.e., a lower error residual. Were this not the case, this model should be discarded entirely. However, it is legitimate to ask whether or not it is sufficiently more accurate to justify the expense and complication of four additional free parameters. As discussed previously, this is a subjective point, and may depend on application. In some cases, the lower accuracy provided by system (10) may be tolerable in order to gain advantage in computation. However, based on the analysis in Section 6.2, the extra accuracy of system (12) should sufficiently justify its use. Furthermore, if radiolysis is to be considered, it may be essential to use system (12), with its full complement of species and reactions.

To obtain the results required an understanding of chemical kinetics as well as fairly complicated mathematical techniques. The former kept the problem in proper perspective and limited the field of possible parameter values. This allowed for a relatively simple approach to finding initial guesses for mathematical optimization.

The actual optimization process was developed fully, resembling an optimal control problem in its final formulation. The method of solution was a quasi-Newton iteration scheme, as commonly used in unconstrained optimization. The use of adjoint functions allowed for an efficient procedure for calculating the residual gradient, necessary for revising parameter values in minimizing the least squares residual.

In practice, it appeared that several parameters (rate coefficients) were not fully determined, but could rather take on a range of values without altering the residual appreciably. These were invariably associated with species and reactions for which no data were available. The rate coefficients associated with I^- , I_2 , and IO_3^- (for which data were available) were more securely determined by the optimization process and are therefore more reliable. Each reacting system is somewhat empirical and therefore represents an overall model whose parameters depend heavily on each other. With only a few exceptions, the rate coefficients may be quite unreliable in representing single reactions separate from the rest of the system. This would work only if each reaction were known to represent an actual interaction mechanism. It would require much greater understanding of reaction processes and considerably more data, including at least some transient data for each species.

REFERENCES

REFERENCES

1. C. C. Lin, *J. Inorg. Nucl. Chem.* **43**, 3229-3238 (1981).
2. T. R. Thomas, D. T. Pence, and R. A. Hasty, *J. Inorg. Nucl. Chem.* **42**, 183, (1980).
3. J. T. Bell, M. H. Lietzke, and D. A. Palmer, "Predicted Rates of Formation of Iodine Hydrolysis Species at pH Levels, Concentrations, and Temperatures Anticipated in LWR Accidents," NUREG/CR-2900 (ORNL-5876), 1982.
4. L. M. Toth, K. D. Pannell, and O. L. Kirkland, "The Chemical Behavior of Iodine in Aqueous Solutions up to 150°C," NUREG/CR-3513 (ORNL/TM-8664), 1984.
5. A. M. Deane, "Iodine Hydrolysis Reactions: Preliminary Species Identification and Rate Measurements," in the *Proceedings of the Specialists' Workshop on Iodine Chemistry in Reactor Safety*, AERE-R-11974, Sept. 11-12, 1985.
6. D. J. Wren and J. Paquette, "The Kinetics of Iodine Release from Aqueous Solutions," in the *Proceedings of the Specialists' Workshop on Iodine Chemistry in Reactor Safety*, AERE-R-11974, Sept. 11-12, 1985.
7. E. C. Beahm et al., "Chemistry and Transport of Iodine in Containment," NUREG/CR-4697 (ORNL/TM-10135), p. 18, 1986.
8. A. C. Hindmarsh, *ACM-SIGNOM Newsletter* **15**(4), 10-11 (1980).
9. R. Kalaba and K. Spingarn, *Control, Identification, and Input Optimization*, New York: Plenum Press (1982).
10. P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization*, Academic Press (1981).
11. R. E. Bellman and R. E. Kalaba, *Quasilinearization and Nonlinear Boundary Value Problems*, New York: American Elsevier (1965).
12. R. Bellman, J. Jacquez, R. Kalaba, and S. Schwimmer, *Mathematical Biosciences* **1**, 71-76 (1967).
13. P. Eykhoff, *System Identification*, New York: John Wiley & Sons (1974).
14. R. Fletcher, *Practical Methods of Optimization*, New York: John Wiley & Sons (1987).
15. M. J. D. Powell, *SIAM Review* **28**(4), 499, 1986.
16. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes*, Cambridge University Press (1987).
17. K. J. Laidler, *Chemical Kinetics*, New York: Harper & Row (1987).
18. D. A. Palmer, R. W. Ramette, and R. E. Mesmer, *J. Nucl. Materials* **130**, 280-286 (1985).
19. J. E. Dennis and J. J. More, *SIAM Review* **19**, 46-89 (1977).

20. Y. Bard, *Nonlinear Parameter Estimation*, New York: Academic Press (1974).

APPENDIXES

APPENDIX A

Rate Equations for the Ten-Reaction System

The derivation of rate equations for the set of reactions (4a)– (4j) follows the same procedure developed in Section 3 for the six-reaction system; concentration variables (g-atom/dm³) are defined as

$$C_1 = [\text{I}^-], \quad C_2 = 2[\text{I}_2], \quad C_3 = [\text{HOI}], \quad C_4 = [\text{OI}], \quad C_5 = [\text{HOI}_2],$$

$$C_6 = [\text{OI}_2], \quad C_7 = [\text{HIO}_3]$$

and identifiable rate process of the reaction substeps are

$$\dot{C}_1 = -\frac{1}{2}\dot{C}_2 = \dot{C}_3 = \frac{1}{2}k_1C_2$$

$$\dot{C}_{i-1} = -\frac{1}{2}\dot{C}_i = \dot{C}_{i+1} = k_{i-1}C_i^2, \quad i = 3, 4, 5, 6$$

$$-\dot{C}_1 = \dot{C}_2 = \dot{C}_{i-1} = -\dot{C}_i = k_{13-i}C_1C_HC_i, \quad i = 4, 5, 6, 7,$$

$$-\dot{C}_1 = \frac{1}{2}\dot{C}_2 = \dot{C}_3 = k_{10}C_1C_HC_3$$

These are combined into the system of ordinary differential equations

$$\dot{C}_1 = \frac{1}{2}k_1C_2 - C_1C_H(k_6C_7 + k_7C_6 + k_8C_5 + k_9C_4 + k_{10}C_3) \quad (\text{A.1a})$$

$$\dot{C}_2 = -k_1C_2 + k_2C_3^2 + C_1C_H(k_6C_7 + k_7C_6 + k_8C_5 + k_9C_4 + 2k_{10}C_3) \quad (\text{A.1b})$$

$$\dot{C}_3 = \frac{1}{2}k_1C_2 - 2k_2C_3^2 + k_3C_4^2 + C_1C_H(k_9C_4 - k_{10}C_3) \quad (\text{A.1c})$$

$$\dot{C}_4 = k_2C_3^2 - 2k_3C_4^2 + k_4C_5^2 + C_1C_H(k_8C_5 - k_9C_4) \quad (\text{A.1d})$$

$$\dot{C}_5 = k_3C_4^2 - 2k_4C_5^2 + k_5C_6^2 + C_1C_H(k_7C_6 - k_8C_5) \quad (\text{A.1e})$$

$$\dot{C}_6 = k_4C_5^2 - 2k_5C_6^2 + C_1C_H(k_6C_7 - k_7C_6) \quad (\text{A.1f})$$

$$\dot{C}_7 = k_5 C_6^2 + C_1 C_H (-k_6 C_7) \quad (\text{A.1g})$$

which is analogous to system (6). Scaling is accomplished by introducing new variables as in Eq. (8), and results in the dimensionless system

$$\dot{y}_1 = \frac{1}{2} u_1 y_2 - y_1 (u_6 y_7 + u_7 y_6 + u_8 y_5 + u_9 y_4 + u_{10} y_3)$$

$$\dot{y}_2 = -u_1 y_2 + u_2 y_3^2 + y_1 (u_6 y_7 + u_7 y_6 + u_8 y_5 + u_9 y_4 + 2u_{10} y_3)$$

$$\dot{y}_3 = \frac{1}{2} u_1 y_2 - 2u_2 y_3^2 + u_3 y_4^2 + y_1 (u_9 y_4 - u_{10} y_3)$$

$$\dot{y}_4 = u_2 y_3^2 - 2u_3 y_4^2 + u_4 y_5^2 + y_1 (u_8 y_5 - u_9 y_4)$$

$$\dot{y}_5 = u_3 y_4^2 - 2u_4 y_5^2 + u_5 y_6^2 + y_1 (u_7 y_6 - u_8 y_5)$$

$$\dot{y}_6 = u_4 y_5^2 - 2u_5 y_6^2 + y_1 (u_6 y_7 - u_7 y_6)$$

$$\dot{y}_7 = u_5 y_6^2 + y_1 (-u_6 y_7)$$

where the dimensionless rate coefficients are

$$u_1 = k_1 T,$$

$$u_i = k_i T C_T, \quad i = 2, \dots, 5$$

$$u_i = k_i T C_T C_H, \quad i = 6, \dots, 10.$$

APPENDIX B

Extensions to the Basic Systems

If the reaction (5e) is used in place of (4e), the treatment is similar to that of reaction (5a). The rate expression is

$$\dot{C}_5 = -\frac{1}{2}\dot{C}_6 = \dot{C}_7 = k_5 C_6^2 [\text{OH}^-],$$

where now $C_7 \equiv [\text{IO}_3^-]$. Assume now that the water dissociation equilibrium (5a) holds at all times. Substitution of Eq. (5a) into the above produces a slight change in system (12), where the term u_5 should be replaced by

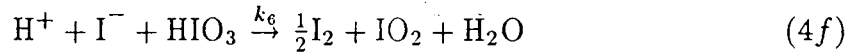
$$\tilde{u}_5 = u_5 [\text{OH}^-] = k_5 T C_T K_w / C_H.$$

If Reaction (5f) is used instead of (4f), then the reaction rate is,

$$-\dot{C}_1 = \dot{C}_2 = \dot{C}_6 = -\dot{C}_7 = k_6 C_1 C_7 C_H^2,$$

and all occurrences of u_6 in system (12) should be replaced by $\tilde{u}_6 = u_6 C_H$.

The alternative is to continue using Eq.(4f), together with the ionization relationship between $C'_7 \equiv [\text{HIO}_3]$ and $C_7 \equiv [\text{IO}_3^-]$:



Considering just these two equations, individual rate expressions of interest are

$$\dot{C}_1 = -k_6 C_1 C'_7 C_H$$

$$\dot{C}'_7 = -k_6 C_1 C'_7 C_H - k_+ C'_7 + k_- C_H C_7$$

$$\dot{C}_7 = k_+ C'_7 - k_- C_H C_7$$

The last expression can be rewritten in terms of the other two as

$$\dot{C}_7 = \dot{C}_1 - \dot{C}'_7. \quad (\text{B.2})$$

If the reaction (B.1) is very fast, then equilibrium can be assumed:

$$C'_7 = C_7 C_H / K_a \quad (\text{B.3})$$

Now K_a decreases with temperature, and its minimum value in this study (at 25°C) is 0.16.¹⁸ Furthermore, the minimum pH used is 2.8, which implies that $C_H \leq 0.0016$. Thus, if pH and temperature are fixed, then $K_a/C_H \geq 100$, and Eq. (B.3) can be differentiated to get

$$\dot{C}'_7 = \dot{C}_7 (C_H / K_a)$$

Substituting this into Eq. (B.2) and rearranging yields

$$|\dot{C}_1| \geq 101 |\dot{C}'_7|,$$

which allows Eq. (B.2) to be rewritten as

$$\dot{C}_7 \cong \dot{C}_1 = -k_6 C_1 C'_7 C_H = -k_6 C_1 C_7 C_H^2 / K_a.$$

To include the additional reaction (5g) requires the use of an additional iodine specie:

$$C_i = 3[\text{I}_3^-], \quad y_i = C_i / C_T,$$

where $i=6$ or $i=8$ for use with systems (3) and (4), respectively. The reaction rates are

$$-\dot{C}_1 = -\frac{1}{2}\dot{C}_2 = \frac{1}{3}\dot{C}_i = \frac{1}{2}k_j C_1 C_2$$

$$\dot{C}_1 = \frac{1}{2}\dot{C}_2 = -\frac{1}{3}\dot{C}_i = \frac{1}{3}k_{j+1} C_3$$

where $j=7$ with system (3) and $j=11$ for system (4). The right-hand sides of certain equations in systems (6) and (A.1) must then include additional terms as follows:

$$-\frac{1}{2}k_7C_1C_2 + \frac{1}{3}k_8C_6 \quad \text{in Eq. (6a)}$$

$$-k_7C_1C_2 + \frac{2}{3}k_8C_6 \quad \text{in Eq. (6b)}$$

$$-\frac{1}{2}k_{11}C_1C_2 + \frac{1}{3}k_{12}C_8 \quad \text{in Eq. (A.1a)}$$

$$-k_{11}C_1C_2 + \frac{2}{3}k_{12}C_8 \quad \text{in Eq. (A.1b)}$$

In addition, the following equations must be added to systems (6) and (A.1), respectively:

$$\dot{C}_6 = \frac{3}{2}k_7C_1C_2 - k_8C_6 \quad (6f)$$

$$\dot{C}_8 = \frac{3}{2}k_{11}C_1C_2 - k_{12}C_8 \quad (A.1h)$$

This reaction has been well characterized and is considered to reach equilibrium very quickly. Hence, an alternative implementation of Reaction (5g) is accomplished by using the equilibrium constant, whose value is fairly well known:¹⁸

$$K = \frac{[I_3^-]}{[I^-][I_2]} = \frac{2}{3} \cdot \frac{C_i}{C_1C_2}. \quad (B.4)$$

This allows the elimination of the reverse reaction rate coefficient through the substitution $k_8 = k_7/K$ in System (6) and $k_{12} = k_{11}/K$ system (A.1). The resulting formulation will have only one parameter, but will still move toward equilibrium.

APPENDIX C

Numerical Solution Procedures

C.1. BFGS Algorithm for Minimizing a Single Nonlinear Equation

The BFGS method is used to obtain an update vector $\mathbf{p}^k$ to a parameter vector $\mathbf{u}^k$ in the process of minimizing a nonlinear objective function such as Eq. (19a). To motivate the method, consider the Taylor series expansion for a nonlinear function of several variables

$$\begin{aligned} R(\mathbf{u} + \mathbf{p}) &= R(\mathbf{u}) + \sum_i \frac{\partial R}{\partial u_i} p_i + \sum_{i,j} \frac{\partial^2 R}{\partial u_i \partial u_j} p_i p_j + \cdots \\ &\cong R(\mathbf{u}) + \mathbf{p}^T \frac{dR}{d\mathbf{u}} + \frac{1}{2} \mathbf{p}^T \mathbf{A} \mathbf{p}. \end{aligned} \quad (\text{C.1})$$

For the moment, the superscripts denoting iteration step have been dropped; nevertheless, the vectors $\mathbf{u}$ and $\mathbf{p}$ can be regarded as current value and correction, respectively, to the parameter vector. The vector $dR/d\mathbf{u}$ is the gradient and $\mathbf{A}$ is the matrix containing all second derivatives (the Hessian matrix). From Eq. (C.1), $R(\mathbf{u} + \mathbf{p})$ is a minimum when $\mathbf{p}$ is chosen to minimize the quadratic function

$$\phi(\mathbf{p}) = \mathbf{p}^T \frac{dR}{d\mathbf{u}} + \frac{1}{2} \mathbf{p}^T \mathbf{A} \mathbf{p}.$$

This occurs when the gradient of ϕ is zero, which implies that $\mathbf{p}$ satisfies the linear equation

$$\mathbf{A} \mathbf{p} = -\frac{dR}{d\mathbf{u}}, \quad (\text{C.2})$$

Methods which solve Eq. (C.2) directly are called "Newton Methods." Often, the vector $dR/d\mathbf{u}$ can be determined readily, but the matrix $\mathbf{A}$ is very expensive or difficult to calculate exactly. In such situations (which include ours) it is necessary to approximate $\mathbf{A}$ using previous iterates and gradients. Such methods are termed

“quasi-Newton” or “variable metric.” It is also possible to rewrite Eq. (C.2) as

$$\mathbf{p} = -\mathbf{B} \frac{dR}{d\mathbf{u}}, \quad (\text{C.3})$$

where $\mathbf{B} = \mathbf{A}^{-1}$. In practice, most quasi-Newton methods undertake to update the inverse Hessian $\mathbf{B}$ and apply Eq. (C.3), rather than updating $\mathbf{A}$ and solving a complete set of linear equations. Due to certain computational advantages, this is the procedure used in the present work.

Letting $\mathbf{B}_k$ denote the estimate of the inverse Hessian matrix at iteration step k , the BFGS algorithm involves the following update:

$$\mathbf{B}_{k+1} = \mathbf{B}_k - \frac{\mathbf{B}_k \mathbf{s} \mathbf{s}^T \mathbf{B}_k}{\mathbf{s}^T \mathbf{B}_k \mathbf{s}} + \frac{\mathbf{q} \mathbf{q}^T}{\mathbf{q}^T \mathbf{s}}, \quad (\text{C.4})$$

where $\mathbf{s} \equiv \mathbf{u}_{k+1} - \mathbf{u}_k = \alpha \mathbf{p}_k$ and $\mathbf{q} \equiv \frac{dR}{d\mathbf{u}^{k+1}} - \frac{dR}{d\mathbf{u}^k}$ and superscript “T” denotes transpose. The algorithm requires an initial guess $\mathbf{B}_0$, which is usually the identity matrix.

As mentioned earlier, the BFGS procedure has superlinear convergence properties provided that certain conditions are met.¹⁹ If Eq. (C.1) is a good approximation in the vicinity of the optimal point, and the Hessian matrix $\mathbf{A}$ is positive definite, then the convexity of $R(\mathbf{u})$ is assured. The BFGS update was formulated to preserve positive definiteness and symmetry, provided these qualities are present in the initial guess $\mathbf{B}_0$. In addition, if the step lengths α (see next section) converge to 1, then local convergence of the algorithm is assured and superlinear.

Although the BFGS method is fairly successful at converging to local minima, occasionally a problem arises from the fact that the exact Hessian is not used in Eq. (C.2). Thus, it is possible to have a good approximation $\mathbf{B}_k \frac{dR}{d\mathbf{u}} \cong \mathbf{A}^{-1} \frac{dR}{d\mathbf{u}}$, and yet still have $\mathbf{B}_k$ and $\mathbf{p}^k$ quite different from $\mathbf{A}^{-1}$ and $\mathbf{p}$, respectively. In addition, the updates $\mathbf{B}_k$ may lose positive definiteness due to round-off error, again causing

the algorithm to fail. Thus, when the method appears to be nearly converged, it is necessary to check for true optimality by restarting the algorithm.

C.2. Step-Length Selection

The task of selecting a good step length is essential in facilitating rapid convergence to a solution. For an initial parameter vector $\mathbf{u}_0$ and correction $\mathbf{p}$, it is desired to find a value for α so that

$$R(\alpha) \equiv R(\mathbf{u}_0 + \alpha \mathbf{p}) \quad (\text{C.5})$$

is as low as possible. However, it is also important that the process of finding a good step length be efficient and not an overwhelming computational burden. Some updating schemes require that the actual minimum to Eq. (C.1) be found. One reason for using the BFGS updating procedure is that it tolerates well the use of nonoptimal step lengths.

As a function of α , the residual behaves as depicted in Fig. C.1. The residual value at $\alpha = 0$ is assumed known, since this should be available from initialization or the preceding step. Also known at $\alpha = 0$ is the derivative

$$R'(0) = \left. \frac{dR}{d\alpha} \right|_{\alpha=0} = \mathbf{p}^T \frac{dR}{d\mathbf{u}}(\mathbf{u}_0). \quad (\text{C.6})$$

This quantity is always negative when either the steepest descent or BFGS updates are used, since these specify the respective directions

$$\mathbf{p} = -\frac{dR}{d\mathbf{u}}(\mathbf{u}_0), \quad \mathbf{p} = -\mathbf{B} \frac{dR}{d\mathbf{u}}(\mathbf{u}_0).$$

Substituting either form for $\mathbf{p}$ in Eq. (C.2) and recognizing that $\mathbf{B}$ is positive definite implies that $R'(0) < 0$.

In Fig. C.1 a minimum exists at α^* , but any step length producing sufficient residual reduction should be acceptable. Thus, the value α_1 would produce the

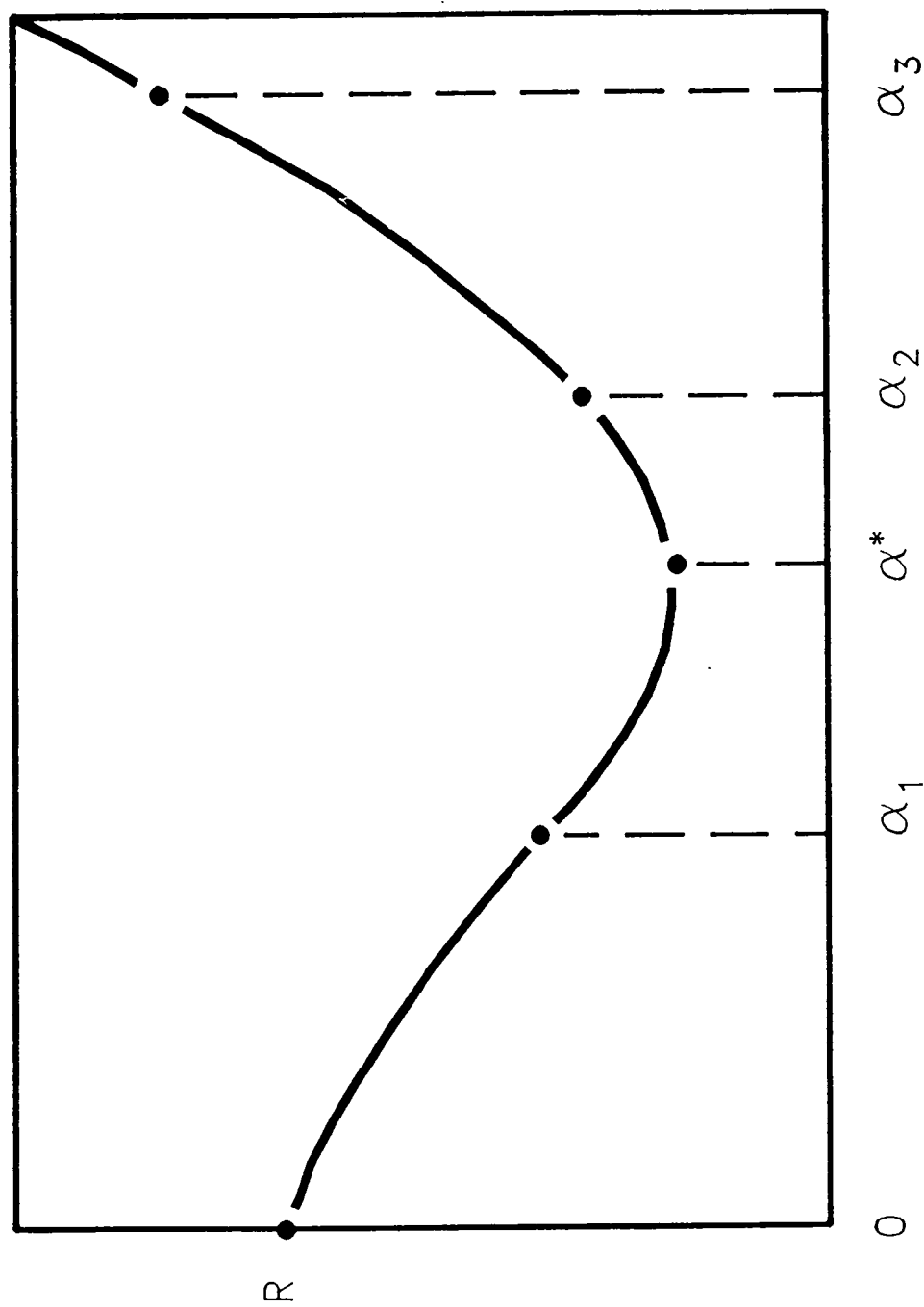


Figure C.1. Residual dependence on step length.

residual $R(\alpha_1)$, which may be acceptable. If not, a larger value should be used (α_1 should be increased if $R'(\alpha_1) < 0$). Likewise, if α_2 is not acceptable, a smaller value should be used (since $R'(\alpha_2) > 0$). The value α_3 would definitely *not* be acceptable, since

$$R(\alpha_3) > R(0).$$

In this case, a fairly good estimate of α^* can be obtained by minimizing the cubic polynomial interpolated through $R(0)$, $R'(0)$, $R(\alpha_3)$, and $R'(\alpha_3)$.

The very first correction step necessarily involves a step-length estimate which is user input. Subsequent steps can use previous iterations to estimate the step length. However, until $R(\alpha)$ and $R'(\alpha)$ are actually calculated, the goodness of this estimate is not known. Occasionally, the initial estimate is good enough, but often one or two corrections are required. The exact procedures used in step-length determination were taken from ref. 14, where a more rigorous and detailed discussion can be found.

APPENDIX D

Efficient Procedure for Determining the Objective Function Gradient

For a given parameter vector $\mathbf{u}$, the objective function $R(\mathbf{u})$ is calculated by solving the differential equations (19b, c) and calculating the error residual (19a). For most optimization routines, the gradient $dR/d\mathbf{u}$ is also necessary, and can be determined by numerical approximation, direct computation, or indirect computation using an adjoint function. The last two procedures are discussed below, where the latter is shown to provide a particularly useful approach. As a preliminary step, the least squares residual in Eq. (19a) can be rewritten as

$$R(\mathbf{u}) = \int_0^1 r[\mathbf{Y}(\mathbf{u}, s), s] ds \quad (\text{D.1})$$

where

$$r[\mathbf{Y}(\mathbf{u}, s), s] = \sum_n \mathbf{E}_n^T \mathbf{E}_n \delta(s - s_n), \quad (\text{D.2})$$

and $\delta(s - s_n)$ is the Dirac- δ function. Differentiating Eq. (D.1) with respect to the parameters u_j yields

$$\frac{\partial R}{\partial u_j} = \int_0^1 \phi_j^T \left(\frac{\partial r}{\partial \mathbf{Y}} \right) ds, \quad (\text{D.3})$$

where $\phi_j \equiv \partial \mathbf{Y} / \partial u_j$ and

$$\frac{\partial r}{\partial \mathbf{Y}} \equiv \left(\frac{\partial r}{\partial y_1^I}, \frac{\partial r}{\partial y_2^I}, \dots, \frac{\partial r}{\partial y_1^I}, \dots, \frac{\partial r}{\partial y_1^I}, \frac{\partial r}{\partial y_2^I}, \dots, \frac{\partial r}{\partial y_1^I} \right)^T.$$

Assuming sufficient smoothness, the unknown functions ϕ_j are determined by differentiating Eqs. (19b, c) to yield

$$\dot{\phi}_j = \mathbf{H} \phi_j + \frac{\partial \mathbf{F}}{\partial u_j}, \quad j = 1, 2, \dots, J, \quad (\text{D.4a})$$

$$\phi_j(0) = 0, \quad (\text{D.4b})$$

where $\mathbf{H} = \partial \mathbf{F} / \partial \mathbf{Y}$ is the Jacobian matrix, which is block diagonal:

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 & & & \\ & \mathbf{H}_2 & & \\ & & \ddots & \\ & & & \mathbf{H}_M \end{bmatrix}, \quad \mathbf{H}_m = \begin{bmatrix} \frac{\partial f_1}{\partial y_1^m} & \cdots & \frac{\partial f_I}{\partial y_I^m} \\ \vdots & & \vdots \\ \frac{\partial f_I}{\partial y_1^m} & \cdots & \frac{\partial f_I}{\partial y_I^m} \end{bmatrix}.$$

Note that $\mathbf{H}$ and $\partial \mathbf{F} / \partial u_j$ may depend on the solution $\mathbf{Y}$ of Eq. (19b), but are independent of ϕ_j ; hence, Eq. (D.4) is a linear differential system. For any estimate of the parameter vector $\mathbf{u}$, a solution to the primary differential system (19b,c) can be obtained and the residual (D.1) calculated. Furthermore, by then solving Eq. (D.4) for each parameter u_j and substituting into Eq. (D.3), the gradient $\partial R / \partial \mathbf{u}$ can also be computed. The principal drawback of this approach is that system (D.4) is a vector equation that must be solved J times. This process must be repeated for each update of the parameter vector $\mathbf{u}$. Thus, the computations involved are extremely laborious, and perhaps prohibitive; reduction of the necessary computations is quite important.

As mentioned above, the Jacobian $\mathbf{H}$ is independent of u_j ; hence, Eq. (D.4) is analogous to solving a single differential system J times, each with a different "source" $\partial \mathbf{F} / \partial u_j$. Using this observation, an adjoint function can be introduced which will allow Eqs. (D.3) and (D.4) to be replaced by a single differential system and a series of integrals. The saving in computation effort is considerable.

Let ϕ^* denote an arbitrary, differentiable vector function of dimension $I \cdot M$ (the same as ϕ_j and $\mathbf{Y}$). Premultiplying Eq. (D.4a) by ϕ^{*T} and integrating in time yields

$$\int_0^1 \phi^{*T} \dot{\phi}_j ds = \int_0^1 \phi^{*T} \left(\mathbf{H} \phi_j + \frac{\partial \mathbf{F}}{\partial u_j} \right) ds. \quad (\text{D.5})$$

Integrating the left-hand side by parts and utilizing Eq. (D.4b) yields

$$\int_0^1 \phi^{*T} \dot{\phi} ds = \phi^{*T} \phi_j \Big|_{s=1} - \int_0^1 \phi_j^T \dot{\phi}^* ds. \quad (\text{D.6})$$

Choosing now the terminal condition $\phi^*(1) = 0$, Eqs. (D.5) and (D.6) can be combined to get

$$\int_0^1 \phi_j^T (\dot{\phi}^* + \mathbf{H}^T \phi^*) ds = - \int_0^1 \phi^{*T} \frac{\partial \mathbf{F}}{\partial u_j} ds.$$

Comparing the above with Eq. (D.3), it is apparent that if ϕ^* is now required to satisfy the system

$$\dot{\phi}^* = -\mathbf{H}^T \phi^* + \frac{\partial r}{\partial \mathbf{Y}} \quad (\text{D.7a})$$

$$\phi^*(1) = 0 \quad (\text{D.7b})$$

then the residual derivatives can be expressed as

$$\begin{aligned} \frac{\partial R}{\partial u_j} &= \int_0^1 \phi_j^T \left(\frac{\partial r}{\partial \mathbf{Y}} \right) ds = \int_0^1 \phi_j^T (\dot{\phi}^* + \mathbf{H}^T \phi^*) ds = - \int_0^1 \phi^{*T} \frac{\partial \mathbf{F}}{\partial u_j} ds, \\ & \quad j = 1, 2, \dots, J. \quad (\text{D.8}) \end{aligned}$$

The adjoint ϕ^* is independent of ϕ_j for each j , and system (D.7) is linear. While the last integral in Eq. (D.8) is more difficult than Eq. (D.3), the need of solving only one differential system [Eq. (D.7) instead of Eq. (D.4)] more than compensates. Thus, for a given estimate of the parameter vector $\mathbf{u}$, both the residual R and its gradient $\partial R / \partial \mathbf{u}$ can be calculated by solving the primary system (19b,c), the adjoint system (D.7), and by performing the integrals

$$\frac{\partial R}{\partial u_j} = - \int_0^1 \phi^{*T} \left(\frac{\partial \mathbf{F}}{\partial u_j} \right) ds, \quad j = 1, 2, \dots, J. \quad (\text{D.9})$$

APPENDIX E

Optimization of Initial Inventories

The residual in system (19) is considered to be not only a function of the rate coefficients $\mathbf{u}$, but of the initial inventories $\mathbf{Y}_0$ as well. Optimization of these additional parameters follows closely the procedures of Appendix D.

Analogous to the development of Eqs. (D.3) and (D.4), the response (D.1) is differentiated with respect to the j^{th} element of $\mathbf{Y}_0$,

$$\frac{\partial R}{\partial Y_{0j}} = \int_0^1 \psi_j^T \left(\frac{\partial r}{\partial \mathbf{Y}} \right) ds, \quad (\text{E.1})$$

where now $\psi_j^T \equiv \partial \mathbf{Y} / \partial Y_{0j}$. This vector function must satisfy the system

$$\dot{\psi}_j = \mathbf{H} \psi_j \quad (\text{E.2a})$$

$$\psi_j(0) = \frac{\partial \mathbf{Y}_0}{\partial Y_{0j}}, \quad (\text{E.2b})$$

where $\mathbf{H}$ is the Jacobian matrix defined in Appendix D. The right-hand side of Eq. (E.2b) is a unit vector with zeroes everywhere except the j^{th} position. The differential equation (E.2a) is the homogeneous form of Eq. (D.4a). Assuming that ϕ^* satisfies Eqs. (D.7) as before, then the analogue of Eq. (D.8) is

$$\begin{aligned} \frac{\partial R}{\partial Y_{0j}} &= \int_0^1 \psi_j^T \left(\frac{\partial r}{\partial \mathbf{Y}} \right) ds = \int_0^1 \psi_j^T (\dot{\phi}^* + \mathbf{H}^T \phi^*) ds \\ &= \psi_j^T \phi^* \Big|_0^1 + \int_0^1 (-\dot{\psi}_j^T + \psi_j^T \mathbf{H}^T) \phi^* ds \end{aligned}$$

The last integral is zero, due to Eq. (E.2a). However, unlike Eq. (D.6), the boundary term is non-zero; from Eq. (E.2b), it follows that

$$\frac{\partial R}{\partial Y_{0j}} = \phi_j^*(0). \quad (\text{E.3})$$

Implementation of Eq. (E.3) is remarkably simple, since the adjoint ϕ^* must already be calculated, and no further integration is required. However, the inclusion of Y_{0j} within the optimization does increase the dimension of the vector-matrix operations, and is therefore not used in preliminary calculations.

APPENDIX F

Description of Optimization Procedures for Six-Reaction Model

Optimization of rate coefficients for system (10) involves judicious selection of initial guesses as well as operation of the mathematical procedures described in Section 5. Numerical trials indicate that use of Eq. (5a) is superior to the forward step of reaction (3a) and is therefore used for the remainder of this work. For completeness, the formation of tri-iodide in reaction (5g) is also included, although it had very little impact on the results. The final formulation is a system of six ordinary differential equations involving seven free parameters at a constant temperature.

In Section 6.1 the Arrhenius equation is used to estimate rate coefficients k ($\text{M}^{-1}\text{s}^{-1}$) at 25°C , resulting in the estimated range shown in Eq. (27). Converting units to $\text{M}^{-1}\text{hr}^{-1}$, and defining $\xi_i \equiv \ln k_i$, this yields

$$17.4 \leq \xi_i \leq 32.4 \quad (\text{F.1})$$

This presents a range in which the search for coefficients begins. As a first step, a specific value was assigned to all rate coefficients, the system (19b,c) was solved once, and the residual (19a) determined. This was done for several values in or near the range shown in Eq. (F.1), and the results are shown in Table F.1. For several cases the optimization routine was run until it stopped, yielding the final values in the last column of Table F.1. The optimized residual for $\xi = 16$ is markedly different, indicating a level of global convergence not present for other initial guesses. Additional runs were made for $\xi = 14$ and 15 , which also appeared to have superior global optimization.

Table F.1. Initial Scoping Calculation for Six-Reaction Model

Initial ξ	Residual	
	Before Optimization	After Optimization
16	7.960	1.945
17	7.581	
18	6.326	4.567
19	5.095	
20	4.605	4.558
21	5.201	
22	6.540	4.558

Results for these three promising cases are shown in Table F.2, with the following observations:

1. ξ_1 and ξ_7 converge to a standard value regardless of initial estimate.
2. ξ_2 , ξ_4 and ξ_6 rise slightly from initial value, but are highly dependent on initial guess.
3. ξ_3 and ξ_5 exhibit erratic behavior.

The first observation indicates that the choice of ξ_1 has a far greater impact on the residual (and hence, on the model) than any of the other parameters, indicating that the first reaction is dominant. Recognizing that ξ_2 represents the natural reverse reaction, the insignificance of this parameter indicates that the value of ξ_1 resembles irreversible reaction rate constant. If ξ_2 is increased then it may reach a level of significance, and the residual may decrease. Furthermore, ξ_2 , ξ_4 , and ξ_6 all represent reverse reactions and are third order. Thus, it is reasonable to group these three together in searching for improved initial guesses.

Table F.2. Optimization Results for Initial Scoping Calculations

Initial ξ	ξ_1	ξ_2	ξ_3	ξ_4	ξ_5	ξ_6	ξ_7	Residual
16	14.83	16.36	15.33	16.18	15.88	16.38	8.215	1.945
15	14.85	15.08	14.54	15.04	15.13	15.10	8.153	2.011
14	14.83	14.04	17.42	14.02	15.91	14.08	8.147	1.991

Next, a number of single iteration trials were performed using different combinations of $\xi_A = \xi_2 = \xi_4 = \xi_6$ and $\xi_B = \xi_3 = \xi_5$, with ξ_1 and ξ_7 fixed. The most pertinent choices are shown in Table F.3 along with the error residuals obtained from the single iteration. For each value of ξ_B a particular value of ξ_A produces a minimum residual. Furthermore, the table indicates that values near $\xi_A = 19$ or 20 appear the most promising for further study.

Table F.3. Advanced Scoping Calculations for Six-Reaction Model

$\xi_1 = 14.83, \xi_7 = 8.210, \xi_A = \xi_2 = \xi_4 = \xi_6, \xi_B = \xi_3 = \xi_5$					
ξ_A	ξ_B	Residual	ξ_A	ξ_B	Residual
18	14	1.572	19	18	1.404
19		1.340	20		1.252
20		1.426	21		1.689
18	15	1.565	19	19	1.525
19		1.306	20		1.227
20		1.380	21		1.544
18	16	1.604	19	20	1.663
19		1.302	20		1.271
20		1.345	21		1.378
18	17	1.678	19	21	1.783
19		1.330	20		1.391
20		1.304	21		1.255
21		1.776	22		1.863

In order to ascertain more exactly the optimal choices at 25°C, several parameter combinations in Table F.3 were allowed to optimize through many iterations. These results are shown in Table F.4, where it is seen that the best results were for initial values of $\xi_A = 19$ and $\xi_B = 15$ or 16.

Table F.4. Optimization of Advanced Scoping Calculations

Initial Values ^a		Final Values							Residual
ξ_A	ξ_B	ξ_1	ξ_2	ξ_3	ξ_4	ξ_5	ξ_6	ξ_7	
19	14	14.77	18.96	19.06	18.97	16.34	20.89	7.600	1.178
19	15	14.76	18.91	20.58	18.53	17.56	24.20	7.575	1.125
19	16	14.77	18.49	19.59	18.75	16.80	23.46	7.574	1.131
20	17	14.75	19.31	17.18	19.62	17.25	19.94	7.588	1.213
20	19	14.73	20.03	19.01	20.00	19.00	20.27	7.590	1.198
21	21	14.73	20.68	21.06	20.88	21.09	20.82	7.591	1.193

^a $\xi_1 = 14.83$, $\xi_7 = 8.210$, $\xi_A = \xi_2 = \xi_4 = \xi_6$, $\xi_B = \xi_3 = \xi_5$

Although additional trials did not lower the residual, they did provide an estimate for the range of uncertainty (sometimes called an “indifference region,” see ref. 20) for each parameter due to calculational error. Based largely on experience and observation of many optimization runs, it is estimated that any residual within 1% of the optimum (in this case, row 2 of Table F.4) is just as likely to reflect parameter values which accurately model the data. (Thus, the final values in row 3 could also be considered as a reasonable choice for optimal parameters. The choice of row 2 as the real optimum is somewhat arbitrary but is aesthetically appealing since its residual is numerically lower.) A number of runs were made in which each parameter was varied individually, while the others were held at the final optimal values in row 2. Table F.5 gives results of such runs, where for each parameter are shown a value higher (ξ_i^+) and lower (ξ_i^-) than the optimum (ξ_i), and the resid-

ual that resulted from replacing ξ_i with ξ_i^+ or ξ_i^- . Letting $R^* = 1.125$ denote the optimal residual, a quadratic polynomial was interpolated through the three points $(\xi_i^-, R(\xi_i^-))$, $(\xi_i^+, R(\xi_i^+))$, (ξ_i, R^*) , as illustrated in Fig. F.1. The minimum of this polynomial ξ_i^{min} is given, and in every case is quite close to the corresponding optimal value. The uncertainty is here defined as the change $\pm\Delta\xi_i$ for which the interpolating quadratic would be equal to $1.01 R^*$. The actual rate coefficients k_i can be determined by the relation $k_i = \exp(\xi_i)$; these are also shown in Table F.6. The ninth column results from the basic parameter minimization only (i.e., using the ξ_i values in the sixth column of Table F.5). These values change slightly when initial inventories are included in the optimization, which lowers the residual from 1.125 to 0.9958. The resulting rate coefficients are shown in the last column of Table F.5. The absolute uncertainties $\Delta\xi_i$ are good estimates for the relative uncertainties of the k_i from the relationship

$$\Delta\xi_i \cong \frac{d\xi_i}{dk_i} \Delta k_i = \frac{\Delta k_i}{k_i}. \quad (\text{F.2})$$

Table F.5 Parameter Changes and Uncertainty in Six-Reaction Model

i	ξ_i^-	$R(\xi_i^-)$	ξ_i^+	$R(\xi_i^+)$	ξ_i	ξ_i^{min}	Uncertainty $\Delta\xi_i$	$k_i(\text{M}^{-1} \text{ hr}^{-1})$	
								Basic ^a	Final ^b
1	14.65	1.144	14.90	1.151	14.76	14.77	0.088	2.572×10^6	2.546×10^6
2	18.56	1.139	19.11	1.135	18.91	18.86	0.264	1.631×10^8	1.582×10^8
3	19.91	1.135	21.11	1.133	20.58	20.54	0.664	8.665×10^8	8.668×10^8
4	17.85	1.136	19.21	1.136	18.53	18.53	0.680	1.116×10^8	1.111×10^8
5	16.86	1.142	18.31	1.137	17.56	17.65	0.635	4.229×10^7	4.247×10^7
6	23.50	1.136	25.02	1.138	24.20	24.23	0.728	3.235×10^{10}	3.236×10^{10}
7	7.08	1.145	8.00	1.136	7.58	7.60	0.392	1.949×10^3	1.938×10^3

^aCalculated using ξ_i from column 6.

^bCalculated by including optimization of initial inventories.

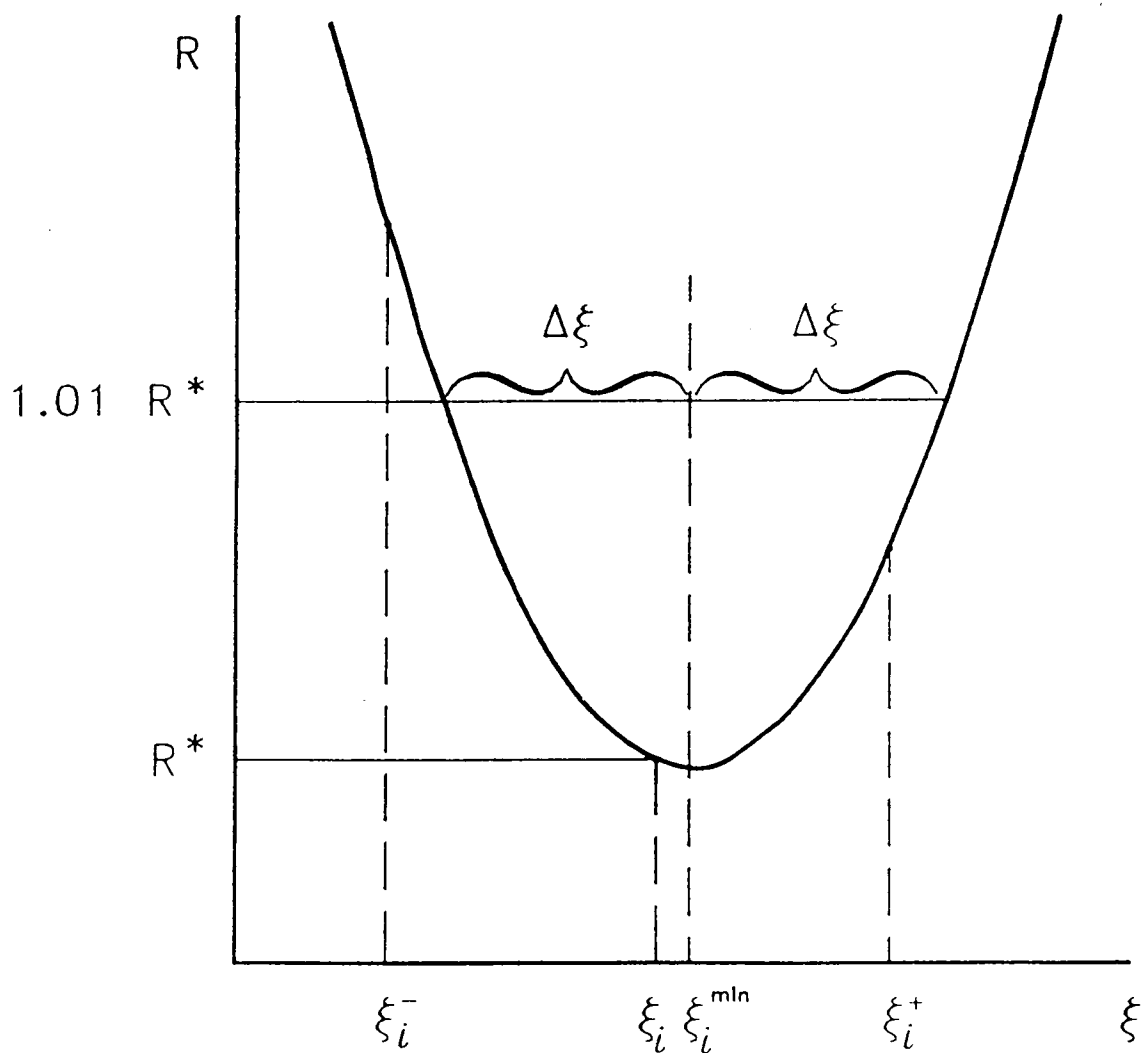


Figure F.1. Interpolating quadratic polynomial for determining parameter uncertainties.

APPENDIX G

Optimization Procedures for Ten-Reaction Model at 25°C

The process of obtaining an optimal set of rate coefficients for system (12) is similar in many ways to the procedure for the six-reaction model, described in Appendix F. The use of reactions (5a) and (5e) in place of Eqs. (4a) and (4e), respectively, produced demonstrably superior results and hence were used for all calculations described in this study. Also included was the formation of tri-iodide (5g), although its effect was not appreciable. Using the equilibrium constant (B.4), the resulting formulation has 11 rate coefficients, implying 11 free parameters at constant temperature.

Defining $\xi_i \equiv \ln k_i$ as in Appendix F, the suggested range for rate coefficients is given by Eq. (F.1). A single value within or near this range was assigned to all rate constants as an initial guess. The optimization procedure of Section 5 was implemented and allowed to operate until no further improvement occurred. The results for several different initial guesses are shown in Table G.1. As seen in the table for initial guesses 14, 15, and 16, most of the final values do not deviate much from the initial guesses. This was even more the case for initial guesses 17, 20, and 24, which terminated without approaching any sort of global optimum. Following are observations for the two cases resulting in the lowest final residual (resulting from initial values $\xi=15$ and 16):

1. The values of ξ_1, ξ_2, ξ_6 and ξ_{11} are fairly independent of initial guess.
2. Most values are somewhat below the range suggested in Eq. (F.1).
3. Considerable reduction in the residual is possible and the procedure is capable of altering parameter values by several orders of magnitude.

Table G.1. Primary Optimization for Ten-Reaction Model at 25°C

Initial Guess ^a	Final Values											Residual
	ξ_1	ξ_2	ξ_3	ξ_4	ξ_5	ξ_6	ξ_7	ξ_8	ξ_9	ξ_{10}	ξ_{11}	
14	15.25	14.21	14.12	14.04	14.01	14.02	14.02	14.01	14.01	14.01	7.72	2.136
15	15.20	13.22	14.88	15.00	15.01	20.37	15.56	15.18	15.13	15.08	7.31	1.337
16	15.20	13.07	15.75	15.97	16.03	20.61	14.87	17.39	16.33	16.10	7.26	1.327
17												4.833
20												4.599
24												5.565

^aValues of ξ_i for all $i=1, \dots, 11$.

A number of scoping calculations were performed in a similar manner to those described in Appendix F for the six-constant model, in which different combinations of initial values were used. A range was determined for each parameter which ensured it would make some contribution to the model, and the optimization procedure was then run until it converged.

The resulting values for ξ_i are given in column 6 of Table G.2. A sensitivity analysis was then performed, again in a manner identical to that described in Appendix F for the six-constant model. Each parameter was varied in turn from its optimal value; two choices for each parameter and the corresponding residuals are shown in columns 2–3 and 4–5 of Table G.2. For each parameter, an interpolating quadratic polynomial (see Fig. F.2) was then determined through the optimal point $(\xi_i, R(\xi_i))$ and the two variation points $(\xi_i^-, R(\xi_i^-))$ and $(\xi_i^+, R(\xi_i^+))$. The minimum of this quadratic is $\xi_i^{\min}$, given in column 7 of Table G.2. As seen in the table, these values are generally quite close to the original optima. Again, the uncertainty $\Delta\xi_i$ is the change from $\xi_i^{\min}$ which produces a polynomial value that exceeds the optimal residual by 1% (see Fig. F.1). These values are given in the eighth column of Table G.2. The last two rows of this table consider identical changes in two parameters si-

multaneously. For the two pairs shown, there is a close relationship in both optimal value and uncertainty.

Actual rate coefficients are determined from the relation $k_i = \exp(\xi_i)$. Values corresponding to column 6 of Table G.2 are given in the final column. When initial inventories were included in optimization, the final values of rate coefficients were unchanged except for $k_{11} = 777$; however, the error residual was reduced from .5195 to .3586. As in Table F.5, the absolute uncertainties $\Delta\xi_i$ are related to the relative uncertainties $\Delta k_i/k_i$ by Eq. (F.2).

Table G.2. Parameter Changes and Uncertainty in the Ten-Reaction Model

i	ξ_i^-	$R(\xi_i^-)$	ξ_i^+	$R(\xi_i^+)$	ξ_i	ξ_i^{min}	$\Delta\xi_i$	$k_i(\text{M}^{-1} \text{ hr}^{-1})$
1	15.65	.5270	15.80	.5241	15.73	15.73	0.069	6.7×10^6
2	12.95	.5272	13.45	.5225	13.25	13.26	0.250	5.68×10^5
3	26.2	.5229	36.2	.5201	32.2	33.0	8.3	5.85×10^{13}
4	26.2	.5267	35.2	.5216	32.2	32.0	4.9	5.85×10^{13}
5	28.7	.5281	37.7	.5248	32.7	33.7	3.9	1.59×10^{14}
6	19.21	.5270	19.56	.5244	19.43	19.40	0.16	2.74×10^8
7	15.83	.5240	17.53	.5242	16.93	16.68	0.89	2.25×10^7
8	22.3	.5229	28.3	.5266	25.3	24.8	3.0	6.52×10^{10}
9	21.3	.5205	28.3	.5229	25.3	23.9	5.2	6.52×10^{10}
10	28.21	.5242	28.53	.5258	28.36	28.36	0.15	2.07×10^{12}
11	6.359	.5257	6.909	.5242	6.659	6.653	0.267	780
3,4	26.2	.5267	35.2	.5216	32.2	31.7	4.1	
8,9	22.3	.5260	27.3	.5249	25.3	24.9	2.32	

APPENDIX H

Optimization Theory Extension to Arrhenius Constants

Using other quantities defined in Section 5.2, the standard optimization problem (19) can be expressed as

$$\min_{\mathbf{u}(A,E)>0} R = \sum_n \sum_m (\mathbf{e}_n^m)^T \mathbf{e}_n^m \quad (\text{H.1a})$$

$$\text{Subject to :} \quad \dot{\mathbf{y}}^m = \mathbf{f}[\mathbf{y}^m, \mathbf{u}(A, E), s] \quad m = 1, 2, \dots, M, \quad (\text{H.1b})$$

$$\mathbf{y}^m(0) = \mathbf{y}_0^m \quad (\text{H.1c})$$

where the index m indicates that optimization over M different transients is occurring. As discussed in Section 5.4, the gradient that reflects a useful scaling of rate coefficients is

$$\frac{\partial R}{\partial v_j} = \frac{\partial R}{\partial(\ln u_j)} \quad (\text{H.2})$$

Dropping the subscript, each dimensionless rate coefficient is the product of some scaling factor and the actual rate coefficient:

$$u \equiv z k, \quad (\text{H.3})$$

where z can readily be determined for each rate coefficient from Eqs. (11) and (13). From the Arrhenius form (26) it can be seen that the gradients involving A and E which are consistent with Eq. (H.2) are

$$\frac{\partial R}{\partial(\ln A)} = \frac{\partial R}{\partial(\ln u)} \quad \text{and} \quad \frac{\partial R}{\partial E} = \frac{\partial R}{\partial(\ln u)} \frac{\partial(\ln u)}{\partial E} = -\frac{1}{R_g T} \frac{\partial R}{\partial(\ln u)}$$

where the subscript ' g ' is used to distinguish the gas constant R_g from the residual R .

Letting T_m and $R_m \equiv \sum_n (e_n^m)^T e_n^m$ denote the temperature and error residual for the m^{th} transient, the overall gradient terms are

$$\frac{\partial R}{\partial(\ln A)} = \sum_m \frac{\partial R_m}{\partial(\ln u^m)}, \quad \frac{\partial R}{\partial E} = \sum_m \left(\frac{1}{R_g T_m} \right) \frac{\partial R}{\partial(\ln u^m)}. \quad (\text{H.4})$$

where $u^m \equiv z A \exp(-E/R_g T_m)$ is the rate coefficient for transient m . Thus, if the gradients $\partial R_m / \partial(\ln u^m)$ are calculated for each transient just as they are in the constant temperature case, then optimization with respect to the Arrhenius parameters A and E can be accomplished by applying Eq. (H.4).

APPENDIX I

Computer Coding for Optimization Calculations

A FORTRAN code was written to perform the optimization calculations described in Sections 5 and 6. Block diagrams for the principal code organization and two important subsections are shown in Figures I.1–I.3. Composed of approximately 600 statements, the code is arranged in four principal routines:

MAIN Program. This routine performs all input, initialization, and most output. It contains the main iteration loop, which involves:

1. Calculation of the error residual R and gradient dR/du . This is done by calling the routine SOLVIT, described below.
2. Calculation of a correction p to the parameter vector u . This is done initially by the steepest descent method of Eq. (21) and switches to the BFGS method (described in Appendix C) when improvements in the error residual are too small. In the latter case, an update is first made to the inverse Hessian matrix B via Eq. (C.4); this is followed by determination of a correction p from Eq. (C.3). The switching logic of the method selection is illustrated in Fig. I.3.
3. Calculation of step-length α . An estimate α_0 is input initially or approximated using the step length of the previous iteration. This value is refined using an algorithm in ref. 14, as described in Appendix C. A simplification of the decision process is illustrated in Fig. I.2.
4. Updating the parameter vector by replacing u with $u + \alpha p$.

SUBROUTINE SOLVIT – Given a parameter vector u and a vector of initial conditions y_0 for all transients to be simulated, this routine calculates the solution to Eqs.

(19b,c) and the error residual R from Eq. (19a). It then calculates the solution to the adjoint system (D.7a,b) and uses this to calculate the residual gradient dR/du from Eq. (D.8). The integrals are approximated numerically using Simpson's rule.

SUBROUTINE CUBMIN – This routine determines a cubic polynomial, given function values and derivatives at two points. It then calculates the minimum value in a given interval.

SUBROUTINE RCALC – This is an interface and data manipulation routine between the Main Program and SOLVIT.

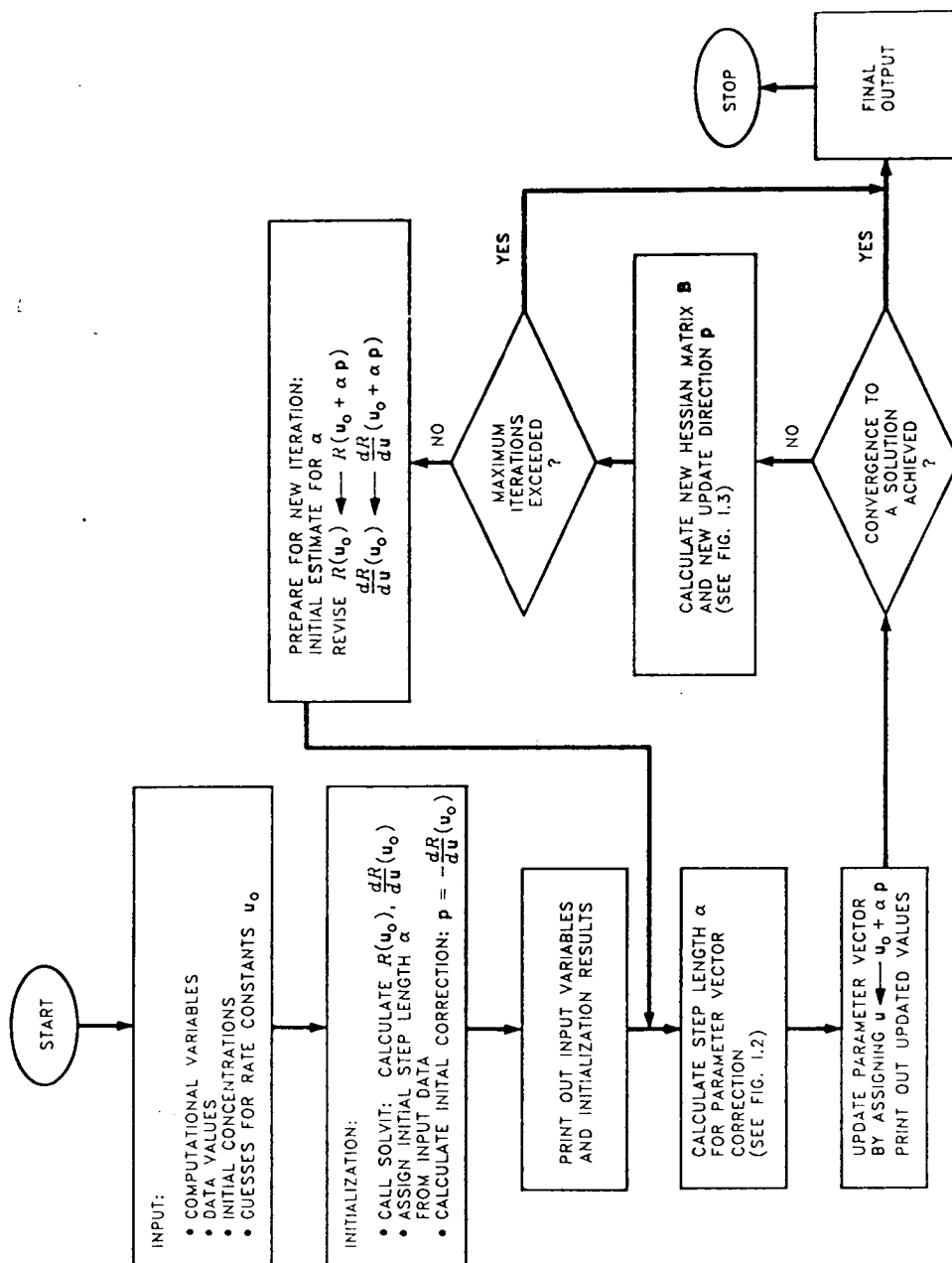


Figure I.1. Main routine block diagram.

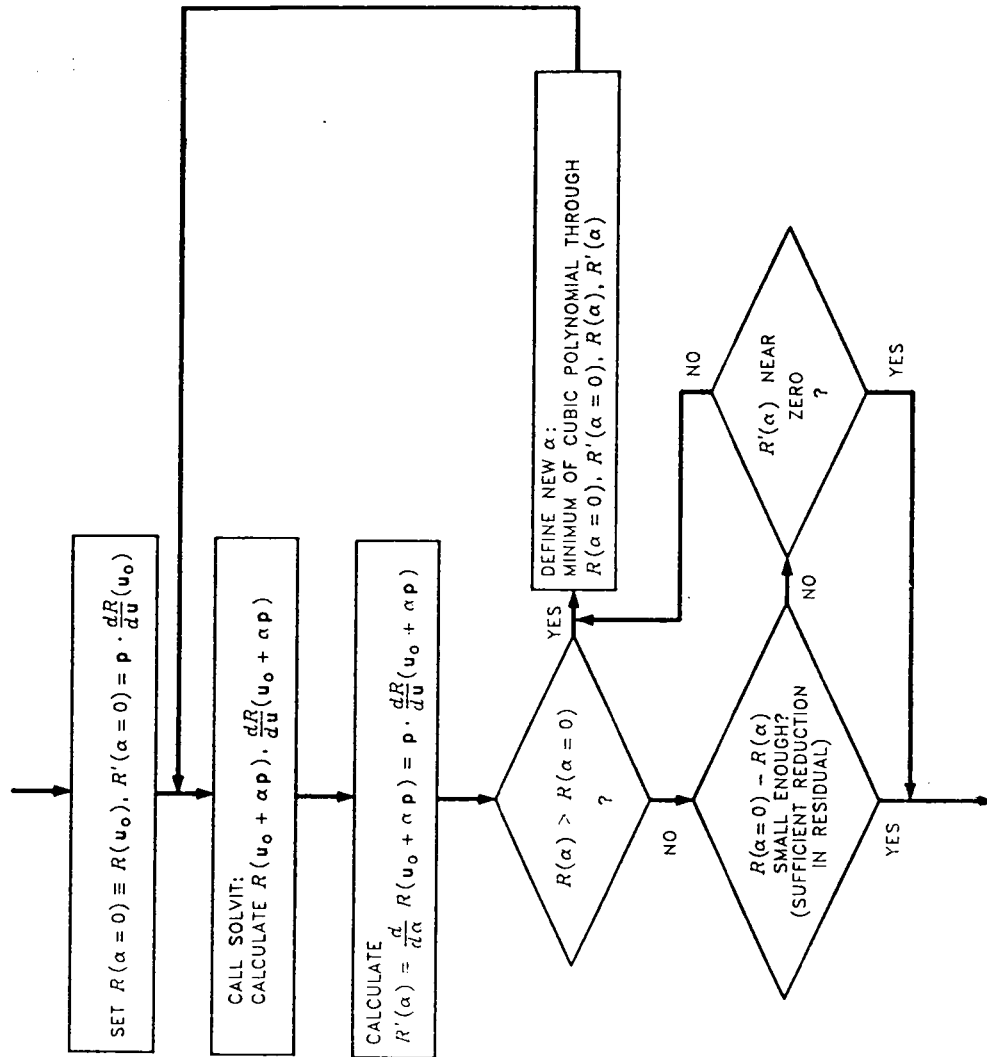


Figure I.2. Block diagram for step-length calculation.

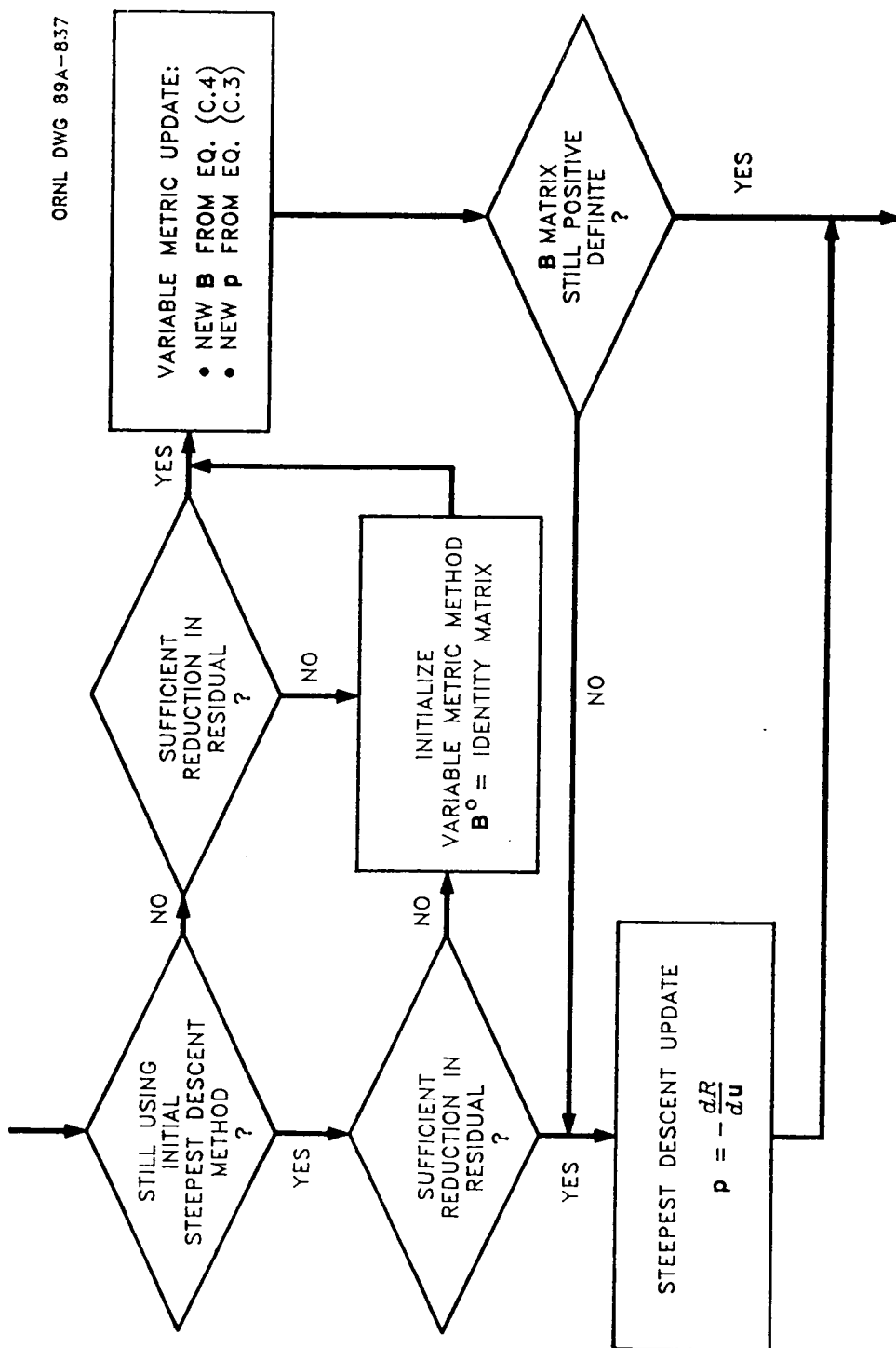


Figure I.3. Block diagram for parameter update calculation.

VITA

Charles F. Weber was born in Tucson, Arizona, on August 13, 1953. He graduated in 1971 from Oak Ridge High School in Oak Ridge, Tennessee. He then attended The University of Tennessee, where in 1975 he received a Bachelor of Arts Degree, majoring in mathematics and minoring in computer science. Continuing his studies there, he received two Master of Science degrees, one in management science (1976), and the other in mathematics (1979). In September 1976, Mr. Weber accepted a teaching assistantship in the Mathematics Department of The University of Tennessee. In January 1978, he began a research assistantship with the Nuclear Engineering Department, in conjunction with the Oak Ridge National Laboratory (ORNL). He later was employed by ORNL, joining the Computer Sciences Division in October 1978. At ORNL, his research has touched several areas of applied mathematics and nuclear reactor safety. Most recently, he has been involved in fission product transport and chemistry in postulated severe reactor accidents.

In addition to a previous thesis, Mr. Weber has authored several reports and open literature publications. He is a member of the American Nuclear Society and the Society for Industrial and Applied Mathematics.

Mr. Weber is married to the former Renee Rios and has a three-year old daughter, Cristina.