DETERMINATION OF OPTIMAL AIR POLLUTION CONTROL STRATEGIES

Thesis by

Chwan Pein Kyan

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1973

(Submitted May 14, 1973)

ACKNOWLEDGMENTS

I wish to express my sincere appreciation to my research advisor, Professor J. H. Seinfeld for his inspiration, guidance and encouragement throughout the course of this work.

Financial support from the California Institute of Technology in the form of a teaching assistantship, from the National Science Foundation, the John A. McCarthy Foundation, the Earle C. Anthony Foundation and from the Li Ming scholarship is gratefully acknowledged.

I wish to express my gratitude to my father, Chwan Wint Hway for his inspiration and my brother Chwan Pein Fan for his selfless devotion, encouragement and understanding throughout my whole educational career. The endurance of my family during some of the most difficult times in my life is also acknowledged.

I wish to express my appreciation to all my fellow graduate students for all the valuable interaction and discussion. I owe a special gratitude to Steve Reynolds for his help with part of the source and meteorological data.

ABSTRACT

One of the important environmental problems facing urban officials today is the selection and enforcement of air pollutant emission control measures. These measures take two forms: long-term controls (multi-year legislation, such as the Federal new car emission standards through 1976) and short-term controls (action taken over a period of hours to days to avoid an air pollution episode). What is required for each form of control is a methodology for the systematic determination of the "best" strategy from among all those possible. In this thesis, a general theoretical framework for the determination of optimal air pollution control strategies is presented for both long-term and real-time controls.

For the long-term control problem, it is assumed that emission control procedures are changed on a year-to-year basis. The problem considered is to determine the set of control measures that minimizes the total cost of control while maintaining specified levels of air quality each year. It is assumed that an airshed model exists which is capable of predicting pollutant concentrations as a function of source emissions in the airshed. Both single-year and multi-year problems are treated. Computational methods are developed based on mathematical programming techniques. The theory and computational methods developed are applied to the evaluation of long-term air pollution control strategies for the Los Angeles basin. Optimal strategies for the control of carbon monoxide, nitrogen dioxide and ozone for 1973 to 1975 in the Los Angeles basin have been obtained.

The problem of determining real-time (short-term) air pollution control strategies for an urban airshed is posed as selecting those control measures from among all possible such that air quality is maintained at a certain level over a given time period and the total control imposed is a minimum. The real-time control is based on meteorological predictions made over a several hour to several day period. A computational algorithm is developed for solving the class of control problems that result.

Typical control measures include restrictions on the number of motor vehicles allowed on a freeway, reduced operation of power plants, and substitution of low emission fuel (e.g. natural gas) for high emission fuel (e.g. coal) in power plants. The control strategy is assumed to be enforced over a certain period, say, one hour, based on meteorological predictions made at the beginning of the period. The strategy for each time period could be determined by an air pollution control agency by means of a computer implementing the algorithm presented. The theory is applied to a hypothetical study of implementation of the optimal control on September 29, 1969 in the Los Angeles basin.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. INTRODUCTION

The establishment of strategies for air pollution control is one of the key environmental problems facing urban officials and legislators today. Because of the complexity of the problem, it will be necessary to establish rational, systematic techniques for evaluating and comparing the multitude of possible air pollution control measures for a particular air quality control region. It will then be possible to elucidate the effect of control measures on air quality, to coordinate and utilize the resources of air pollution control in an efficient manner, and to develop an appropriate time-table for long-term control.

A literature survey is given below:

Kohn (1969, 1970) determined the least cost way of achieving given set of reductions in mass emissions of CO, SO_2, hydrocarbons, NO_X and particulate matter for St. Louis in 1975. Farmer et al.(1970) also discussed the formulation of emission control strategies for SO_2 and particulate matter for St. Louis. Burton and Sanjour (1970) employed a computer-assisted system simulation to determine the cost and measure of effectiveness for a given abatement of SO_2 and particulate matter in Kansas City and Washington D.C. This kind of system simulation approach where given a set of control measures, the performance of the system is then evaluated, has been used by Bounds (1971) and Morgenstern et al. (1973). The latter evaluated alternative SO_2 control strategies for Boston intrastate air quality control region. A cost benefit approach for the comparison of different emission

control strategies has been used by Wilson and Minnotte (1969), Hamburg and Cross (1971), Kohn (1972) and Babcock and Nagda (1973). Trijonis (1972) has performed an analysis of air pollution control in the Los Angeles basin in 1975, and obtained optimal control strategies based on a statistical airshed model. All these studies were concerned with long-term control determination.

Using a systems approach, Bibbero (1971) discussed at length the concept of air pollution management and control on a nationwide basis. Malone (1972) also used a systems approach to model air pollution control as a large scale, complex system.

Herzog (1969) has outlined how an urban air diffusion model can be incorporated into zoning decisions. Reiquam (1971) treated the optimal allocation of source emission in an airshed to minimize the likelihood of violating air quality standards. More general concepts of air pollution control in the context of urban and economic planning, and management programs were treated by Fensterstock et al.(1971), Sporn (1971), Croke et al.(1971), Smith et al.(1972), Kleiman (1971) and Muller (1973).

For real-time control Savas (1969) discussed a series of conventional feedback control diagrams from which an integrated system of control may be developed. Parson and Croke (1969) evaluated the economics of SO_2 incident control for Chicago. Friedlander (1969) discussed the computer control of vehicular traffic to ease emissions in episode conditions. Croke and Booras (1970) treated the real-time control of SO_2 in Chicago. Shepard (1970) treated the real-time

load shifting among the various power plants in an airshed in case of an episode. Leavitt et al. (1971) discussed limiting power plant emissions according to meteorology. Rossin et al. (1972) analysed CO episode control, introducing a health criterion based on the dissolved CO concentration in the blood and outlined a number of possible real-time control measures.

The main objective of this thesis is to develop a general theoretical framework for the determination of optimal air pollution control strategies for both the long-term and the real-time control problems. Application to the Los Angeles basin will then be attempted.

CHAPTER 2. GENERAL CONSIDERATIONS IN AIR POLLUTION CONTROL PLANNING

In this chapter, a general description of the air pollution control problem for an airshed is given. Those components of the airshed system, important in control studies are delineated. Some conceptual ideas of air pollution control are then introduced.

2.1 Description of an air pollution control system

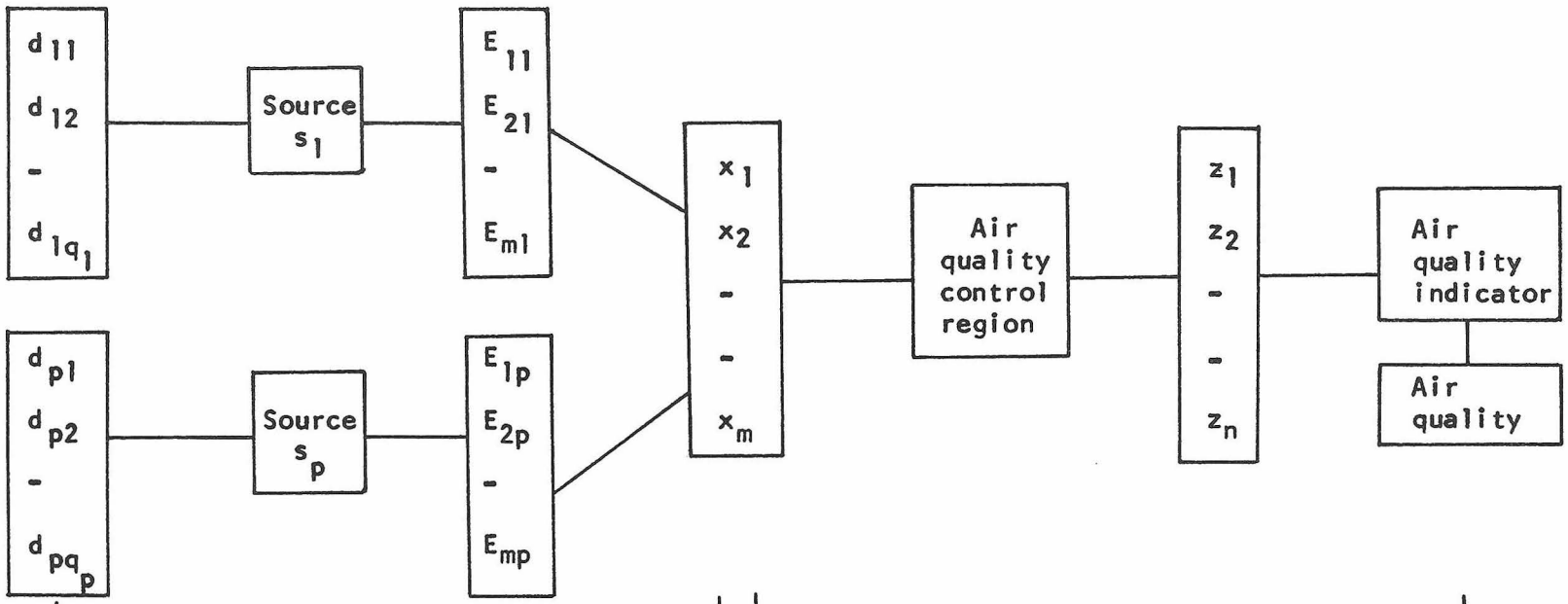
The airshed system consists of the following components:

(1) Various pollutant-emitting sources such as motor vehicles, power plants, industries, petroleum marketing and solvent users, aircrafts etc.

(2) Various chemical species, i.e. pollutants. The primary pollutants which are emitted directly from sources consist of predominantly CO, hydrocarbons, NO_x, SO_2 and particulates. The secondary pollutants formed from the primary ones by atmospheric chemical reactions, consist of O_3, NO_2, H_2SO_4 and organic compounds.

(3) A multitude of control methods for controlling the pollutant emissions of the various sources. For example, emissions from motor vehicles can be controlled by evaporative control devices, crankcase control devices etc. Emissions from power plants can be controlled by burner modifications, substitution of natural gas for fuel oil etc. These control methods are the variables that we can manipulate to achieve a certain objective (such as cleaner air).

(4) Meteorological and topographical parameters. These consist



Emission control methods or devices	Sources of emission	Emission level of various sources	Emissions of primary pollutants	Atmospheric simulation model	Concentrations of pollutants in atmosphere	Air quality for the airshed
d_{ij}	s_i	E_{ik}	X_k		z_i	g_i
$i=1,2,..,p$ $j=1,2,..,q_i$	$i=1,2,..,p$	$i=1,2,..,p$ $k=1,2,..,m$	$k=1,2,..,m$		$i=1,2,..,n$	$i=1,2,..,L$

p = Number of sources in the airshed. q_i = Number of control methods available for source i
m = Number of primary pollutants. n = Number of total airborne pollutants (primary+secondary)
L = Number of air quality indicators desired.

Figure 2.1 Schematic representation of an airshed system

of wind flow, temperature inversion, turbulent mixing, atmospheric irradiation and the geographical location of the airshed. These are measurable but not manipulable variables.

(5) Atmospheric chemical reactions. These constitute the means of transforming primary pollutants to the secondary ones.

Schematically, an air pollution control system is depicted in Figure 2.1.

The sources are spatially distributed inside the airshed and their activities are dictated by certain temporal distributions. Thus, there are definite spatial and temporal patterns by which the pollutants are emitted into the atmosphere. The distribution of airborne pollutant concentrations as a function of time and position inside the airshed depends on the influence of turbulent mixing and chemical reactions, processes presumably describable by a suitable airshed simulation model. The air quality can then be evaluated from the airborne pollutant concentrations according to the definition of an air quality indicator. Thus, a set of control measures applied to the sources will produce a given level of pollutant emissions having a certain spatial and temporal distribution, resulting in a certain level of air quality. Conversely, for a given air quality, there may exist many possible sets of control measures that can achieve the same air quality. Therefore, there arises naturally the question of which set of the possible control measures is the best in some sense.

The air pollution control system can be decomposed into the following two subsystems:

(1) The control-method-emission subsystem which relates the control methods to the source emissions in the airshed and is independent of the meteorological conditions of the airshed.

(2) The emission-air-quality subsystem which relates source emissions to air quality. This subsystem depends directly only on the spatial and temporal distributions of the source emissions.

In later chapters, we shall see that this decomposition can simplify the treatment of the whole air pollution control system greatly.

2.2 Analysis of an air pollution control system

An analysis of the air pollution control system will involve the following steps:

(1) Establishing the desired air quality goals, in terms of atmospheric concentrations of pollutants.

(2) Elucidating the entire spectrum of control measures.

(3) Establishing the criteria by which the alternative control strategies are to be evaluated.

(4) Determining the set of control actions (from among all those possible) which in some sense provide an "optimal" solution.

Air quality goals will ultimately be established on the basis of medical, aesthetic, and economic effects of air pollution. At the present time, however, we do not have enough information to provide a firm, quantitative link between these factors and airborne pollutant concentrations. Thus, we cannot say, for example, that an exposure over a one-year period to a certain level of concentration will lead to an increase in lung disorders of a given percent. The measures of

air quality that we use will vary, in general, for each pollutant species, depending on the factors mentioned above.

The next step is the elucidation of the alternative control strategies. The broadest classification of air pollution control measures could be made on the basis of where in the total system the control is exercised. Three points in the air pollution system are amenable, at least in principle, to control action. First, control can be exercised at the sources of emission, resulting in lower quantities or a different distribution of primary effluents reaching the atmosphere. For the internal combustion engine, for example, emission control actions are fuel modifications, engine modifications, and catalytic and thermal afterburners. Rapid transit and traffic control are also emission controls, since they affect the spatial and temporal emissions of the sources and do not change the atmospheric transport and mixing capacity. Second, control could be levied on the atmosphere, for example, in the form of diverting wind flows or discharging huge quantities of heat to break a temperature inversion. Finally, air pollution control could be reserved for receptors, for example, by extensive use of filtered air conditioning systems, or, in the limit, use of gas masks. Of the three, control at the emission source is not only the most feasible but also the most practical. In short, the best way to control air pollution is to prevent contaminants from getting into the atmosphere in the first place. Thus, we will consider here only those control techniques which are exercised directly on the sources, that is, those which affect the quantity or the spatial

and temporal distribution of emissions.

Control of contaminant emissions can assume several forms. The most obvious is the control of the quantity of material emitted over a certain time period. Also important is control of emission timing, namely, the rescheduling of certain activities so that those pollutants which must be discharged are done so at as advantageous a time as possible during the day in terms of atmospheric accumulation. The spatial distribution of the source emissions can also be varied. Finally, the location of emissions can be controlled by proper zoning for freeways and industrial development or requiring the use of high stacks for dispersal. While each of these forms plays an important role in air pollution control, the most prevalent and in many ways the most feasible, at least for existing sources, is the control of the quantity of material emitted.

Emission control programs can be divided into two categories:

(1) Short-term control.

(2) Long-term control.

Short-term control involves measures such as shutdown and slowdown procedures which are adopted over periods of several hours to several days under impending adverse meteorological conditions. Long-term control strategies involve a legislated set of measures to be adopted over a multi-year period.

An example of a short-term strategy are the emergency procedures for fuel substitution by coal-burning power plants in Chicago when SO_2 concentrations reach certain levels. An example of a long-term

control policy is Los Angeles County Rule 68, which provides for a two-step reduction in allowable emissions of oxides of nitrogen (NO_x) from fuel burning equipment producing more than 1775 million Btu per hour. The rule specifies that the maximum parts-per-million (ppm) by volume of NO_x in the effluent gases from gas-fired equipment must be 225 ppm after December 31, 1971, and 125 ppm after December 31, 1974.

The next step in the analysis is the establishment of criteria by which the alternative strategies are to be evaluated. This is by no means a simple task, although we can state that in general we should consider the economic feasibility, the social desirability, and the political acceptability of each alternative. In our analysis, we will concentrate on the criterion of economic feasibility only. What we will determine are the optimal strategies in an economic sense. These policies must then, of course, be screened for social and political desirability.

Our objective is to develop a systematic way of comparing alternative control strategies on an economic basis so that the "best", or at least a suitable one, can be chosen. The selection of an appropriate objective function by which to evaluate alternative strategies is a key part of the problem. In principle, an economic objective function should represent the total cost of air pollution to the community. The total cost of air pollution can be roughly divided into a sum of two costs:

(1) The control costs — both the direct (cost of equipment to be installed, cost of new raw materials needed, etc.) and indirect

(costs due to resulting unemployment, costs of enforcement, etc) costs resulting from emission reduction procedures and devices adopted by sources.

(2) The damage costs —both the tangible (damaged materials and crops, hospital bills for respiratory illnesses, etc.) and intangible (unpleasantness of smoggy air, decreased life expectancy in urban climates, etc) costs incurred by the public from living in polluted air.

If both of these cost functions could be determined accurately as functions of air quality, then a solution to the control problem would be to adopt those control measures leading to air quality yielding the minimum in the total cost curve. Unfortunately, it is difficult even to estimate, much less determine accurately, the damage cost of air pollution. In fact, we probably cannot even catalog all the adverse effects of air pollution. Thus, at this time there is little chance of basing air pollution control on a minimum of total cost. The alternative, which actually makes more sense than dealing with total costs, is to determine the minimum cost of control of reaching a given level of air quality. This can be done for a variety of air quality levels, resulting in the minimum cost as a function of level of air quality. The ultimate choice of which level of air quality should be demanded could presumably be made on the basis of the costs involved and other information, for example, on the basis of levels believed to cause adverse health effects, as we had indicated earlier.

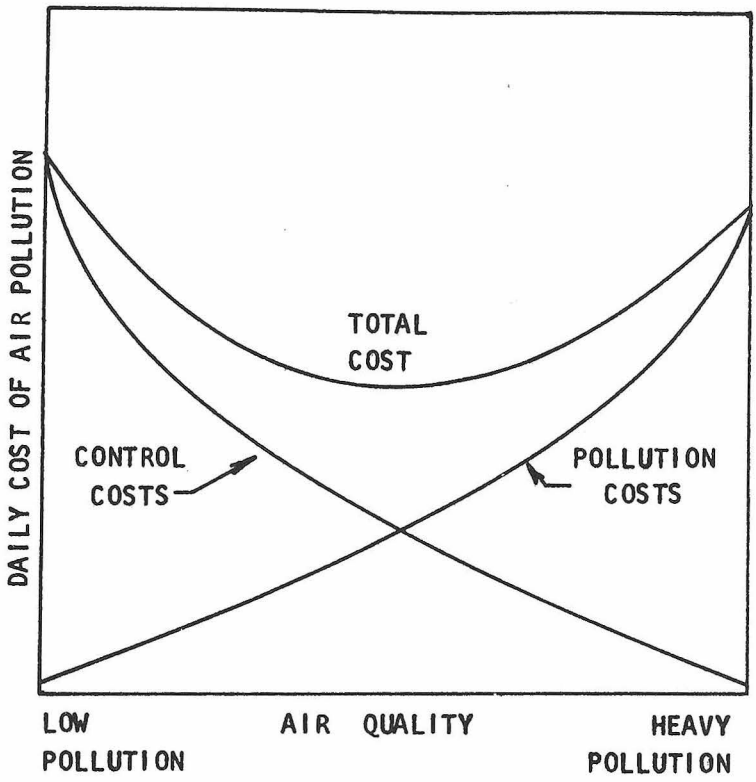


Figure 2.2 Costs of air pollution

The final step in the analysis is the actual evaluation of the alternative strategies to determine the optimal strategy. Since the control will be exercised on the sources and the air quality standard to be met is related to the airborne concentrations, it is clear that it will be necessary to have a simulation model of the airshed which predicts pollutant concentrations as a function of emission levels and meteorology.

An urban airshed is a dynamic system, the state of which can be considered to be the airborne pollutant concentrations as a function of time and location. The source emissions of contaminants as a function of time and location constitute the controllable inputs to the dynamic airshed system. In order to determine the effect of changes in the source inputs on atmospheric concentrations, it is necessary to have a mathematical model of the airshed. There are two basic types of models we can use:

(1) A dynamic model which describes changes occurring over time spans the order of a day, including wind patterns, solar radiation, atmospheric chemistry, and diffusion. The result is concentration values on spatial and temporal scales of the order of 1-2 miles and every 15 minutes, respectively.

(2) A static model which yields long-term (say yearly) average concentrations in the airshed as a function of yearly emissions.

A dynamic model will yield information on actual concentrations given all the required meteorological and emission inputs for the day. Such a model is clearly necessary for short-term air pollution control,

but can also be employed in determining long-term controls. Dynamic airshed models proposed have generally been deterministic in nature relying on solution of various forms of mass conservation equations (Seinfeld 1970-1972).

A static model, on the other hand, will yield average concentrations over a long period of time. These models often incorporate wind roses, or other meteorological frequency distributions and much more simplified treatments of diffusion and reaction than dynamic models (Trijonis 1972, Slade 1968, Pasquill 1962). Static models are by the nature stochastic models (often statistical regression models) since probability distributions of meteorology are prime inputs. The choice of whether to employ a dynamic or a static model in control studies depends on the air quality measure to be met.

2.3 Further discussion of long-term and real-time controls

There are two basic strategies for controlling a dynamic system: open-loop and closed-loop control. In open-loop control, the control policy to achieve a desired objective is determined on the basis of the initial state of the system and any expected inputs during the evolution of the system, i.e. open-loop control is predetermined and not altered during the evolution of the system. In feedback closed-loop control, the control policy is determined at each time during the evolution of the system by comparing the actual output of the system and the desired output and manipulating system inputs to make the actual output match the desired output.

The long-term air pollution control problem is an open-loop

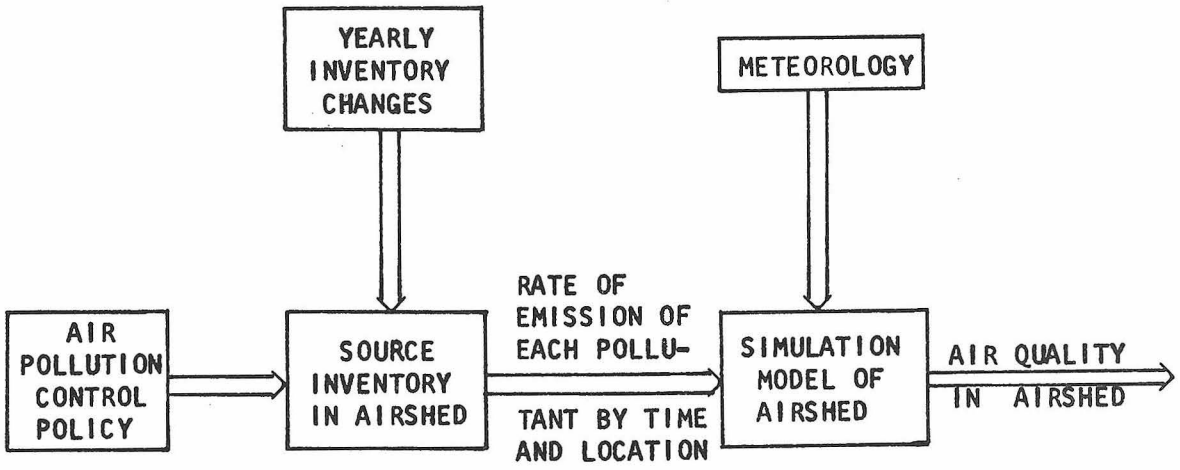


Figure 2.3 Open-loop control of air pollution in an airshed

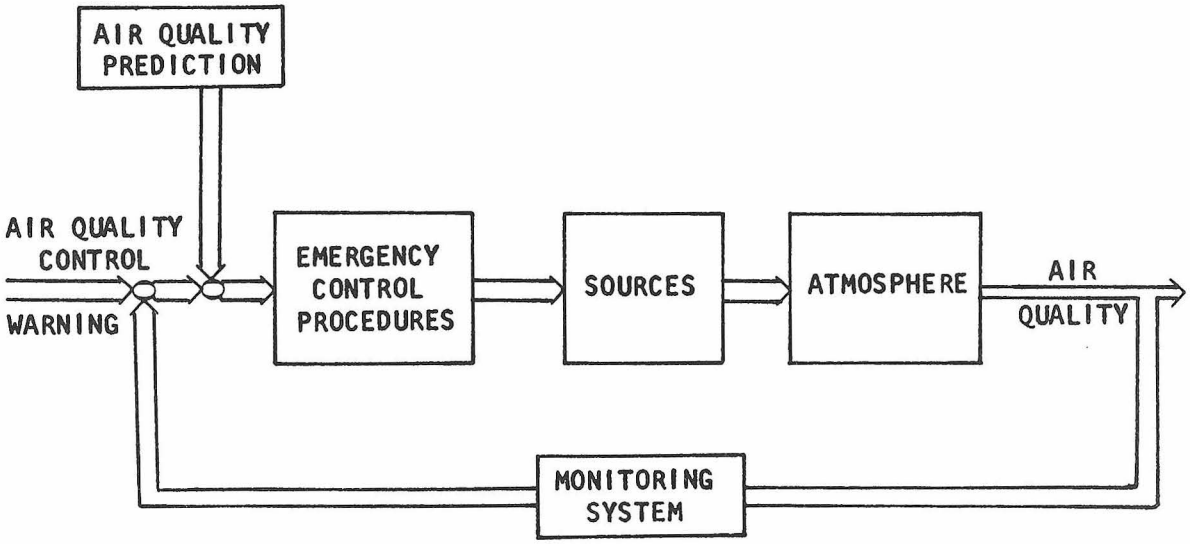


Figure 2.4 Closed-loop control of air pollution in an airshed

control as depicted in Figure 2.3. The determination of emission control strategies for a particular year or for a number of years is an exercise of open-loop control. The real-time air pollution control problem can also be an open-loop control, important in an hour-to-hour or day-to-day capacity. If weather predictions indicated the possibility of forthcoming adverse meteorological conditions, e.g. low inversion and light winds, control measures could be announced that would have to be instituted by various sources during the affected period.

Closed-loop control is most important when combined with an air quality monitoring system, from which measurements of air quality made during the day can be used to put into operation rapid control actions when pollutant concentrations begin to exceed specific warning levels. These warning levels would normally be somewhat lower than those considered to be injurious to health because of the inherently sluggish response of the entire airshed or portions thereof to source emission changes. The system of smog alerts existing in Los Angeles is an example of such control. Savas (1967) and Croke et al. (1969) have discussed the role of feedback control in an urban air monitoring system. Figure 2.4 illustrates the closed-loop control of an airshed.

2.4 Realization of the general concepts

The above general concepts will be made concise in later chapters by detailed mathematical formulations.

In Chapter 3, a systematic mathematical theory for the

determination of optimal air pollution control strategies for the long-term problem (see also Seinfeld and Kyan, 1971,1972) is given, incorporating the causality relationships of control methods, source emissions, simulation model, airborne pollutant concentrations and air quality. The treatment of the long-term air pollution control problem is given for both single-year problem and multi-year problem.

In chapter 4, a general theoretical framework for the determination of real-time air pollution control strategies in the context of an optimal control problem is given (see also Kyan and Seinfeld, 1973). Application of the theory to the real-time control of Los Angles basin is illustrated.

In chapter 5, the evaluation of the long-term air pollution control strategies for the Los Angeles basin is attempted.

All the theoretical formulation and computational methods developed in the later chapters can be used with any airshed model either a dynamic one or a static one. However, for most of the illustrative purposes in the later chapters, a well-mixed cell model, as outlined below, is used.

2.5 A simple airshed simulation model.

Our primary purpose is not to consider atmospheric simulation, thus, we will only consider this subject in enough detail to make clear its relationship to the control problem. The necessary components of an urban airshed model are:

(1) The transport and diffusion model.

This is really the overall model, the major descriptive aspect

of which is the atmospheric transport and dispersive processes. This model will include:

(2) The reaction kinetics model.

This describes the rates of reactions occuring in the atmosphere as a function of concentration, intensity of radiation, temperature, etc.

(3) The emissions model.

This includes a complete source inventory of the airshed describing mass emissions of pollutants as a function of time and location.

A rigorous approach to urban diffusion modeling is direct integration of the three-dimensional, time-dependent partial differential equations of continuity for each species (Seinfeld et al.1972). However a somewhat simpler approach can be adopted, based on the concept of well-mixed cells.

Assume the airshed has been divided into an array of L cells, each of which is considered as a well-mixed reactor. The volumes of the cells, which need not be equal, are $v_1,\ldots,v_L$. The concentration of species i in cell j is z_{ij}. In each cell there is a time-varying source of each pollutant, the rate of emission of species i into cell j being s'_{ij}. Also, there exists the possibility that pollutants can be formed by chemical reaction at a rate r'_{ij}, or removed by deposition, the rate of deposition being d'_{ij}. Finally, the volumetric rate of air flow from cell j to cell k is q_{jk}.

Thus, a dynamic material balance for species i in cell k, when

the volume v_k can vary with time, is

$$v_k \frac{dz_{ik}}{d\tau} = - z_{ik}\frac{dv_k}{d\tau} + \sum_{j=0}^{L} q_{jk}\, z_{ij} - z_{ik}\sum_{j=0}^{L} q_{kj} + s'_{ik} - d'_{ik} + r'_{ik} \qquad (2.1)$$

$$z_{ik}\,(o) = z^{o}_{ik} \qquad (2.2)$$

Normally, $dv_k/d\tau$ is set equal to $A'_k(dh'_k/d\tau)$, where A'_k is the area of the base of a cell having vertical sides and h'_k is the height of the base or an inversion of a convenient mixing height. In effect, the cell is a box with permeable walls and a movable lid. The subscript zero on q_{kj} and z_{ik} relates to flows into and out of the airshed. If we divide the airshed into L cells and consider M components, LM ordinary differential equations of the form of (2.1) will be required to describe the system. Such a model has been introduced for airshed modeling by Ulbrich (1968).

The advantages of this approach are as follows:

(1) Aspects of complicated topographical variations can be easily handled.

(2) Changing inversion levels can be easily handled.

(3) The model is conceptually easy to understand and implement.

However, this approach has several drawbacks:

(1) In the absence of an inversion, the concept of a mixing cell is somewhat artificial.

(2) The assumption that pollutants are instantaneously mixed throughout the entire cell may be a poor one. If vertical mixing is

slow, as under stable meteorological conditions, strong vertical concentration gradients can develop and the well-mixed assumption will not hold.

In spite of its potential drawbacks, the well-mixed cell model represents a reasonable compromise between the complexity of a rigorous partial differential equation diffusion model and the statistical Gaussian plume formulation, inapplicable when chemical reactions are occurring.

The second component of the airshed model is the reaction kinetics model, which, in the cell model, appears in r'_{ik}. A discussion of atmospheric chemistry is beyond our scope and intentions (Altshuller & Bufalini, 1971; Johnston et al. 1970). A generalized kinetic model for photochemical smog that has been successful in simulating both laboratory and atmospheric data is that of Hecht and Seinfeld (1972).

The third component of an airshed model is the sources. Emission magnitudes must be specified as a function of time and location. Sources can be conveniently divided into mobile sources (motor vehicle, aircraft etc.) and fixed sources (power plants, refineries, factories, etc.). An extensive treatment of the source modeling for the Los Angeles basin can be found in Roberts et al. (1971).

CHAPTER 3. LONG-TERM AIR POLLUTION CONTROL

The long-term air pollution control problem involves the evaluation of control actions by sources to be implemented over one year or over a number of years.

As objectives of the long-term control problem, we want to evaluate the following in a systematic manner.

(1) Preferential selection of sources to be controlled

(2) Preferential selection of control methods

(3) Preferential selection of primary pollutants to be controlled

(4) Optimal allocation of control resources in a multi-year period

(5) Minimum cost of meeting a given air quality standard

It will become clear later that the above objectives are complementary and can be achieved simultaneously by solving the problem stated in section 3.1. The long-term problem is formally stated in section 3.1. Its mathematical formulation is given in section 3.2. Section 3.3 treats extensively the computational methods necessary for solving the class of optimal control problems that result.

Applications of the theory developed in this section are delayed until chapter 5.

3.1 Statement of the problem

Given the following for each of a successive number of years:

(a) Various polluting sources and their associated distribution and emission levels of pollutants in an airshed,

(b) Emission control methods and their associated costs and emission reduction characteristics for each of the sources,

(c) A given set of air quality criteria and

(d) An airshed simulation model relating the emission levels and emission spatial and temporal distrubution to the air quality,

determine the set of control measures over a specified period of years , such that a multi-year control cost criterion is minimized and a set of air quality criteria are satisfied.

The above problem assumes the form of a multi-stage optimal control problem.

3.2 Mathematical formulation

For the mathematical formulation of the problem, the following definitions are used.

$s(t)$ = p - source vector in year t , with $s_i(t)$ being the units of source i in year t (e.g. $s_1(1)$ may be the total number of pre-1966 vehicles in the L.A. basin in year 1 (1973) .

p = total number of sources in the airshed

$E(t)$ = m x p emission level matrix with E_{ij} being the emission of pollutant i from unit source j (e.g. $E_{12}(2)$ may be the grams of RHC emitted per vehicle mile of a 1970-model vehicle).

$E^{o}(t)$= m x p emission level matrix without control.

m = number of primary pollutants

$e(\xi, \tau, t)$ = m - emission rate vector with $e_i(\xi, \tau, t)$ being

the emission rate of pollutant i from all the sources at (ξ,τ). (e.g. e $(\xi^*, \tau^*, 2)$ may be the tons/hour of NO_x emitted from all the sources at 8 a.m. in downtown Los Angeles for year 2 (1974)).

E accounts for the emission by each type of sources everywhere in the airshed. e accounts for the emission rate of pollutants from all types of sources in a certain location of the airshed at a certain time. Thus, E has "source resolution" while e has " space resolution". E is for later use in control method constraints and e in the airshed simulation model.

ξ = spatial variable.

τ = time (hourly) variable.

d_{ijt} = Number of units of control method j per unit of source i in year t. (e.g. d_{123} may be the number of cubic feet of natural gas substituted for fuel oil for every mega-watt-hour of power generated in year 3 (1975)).

$w(t)$ = K-dimensional emission control vector instituted in year t. (e.g. $w_1(1)$ may be the number of million cubic feet of natural gas substituted for fuel oil in power plants in year 1 (1973)). For later convenience, we define $w(0) = 0$.

q_j = number of control methods available for source j,

$K = \sum_{i=1}^{p} q_i$ = total number of control methods available for all the sources in the airshed.

$\overline{w}(t)$ = K-dimensional emission control vector which can be instituted or taken-off on a yearly basis. (e.g. substitution of natural gas in power plants).

$\overline{\overline{w}}(t)$ = K-dimensional emission control vector which once instituted will remain on the sources for the life of the control device. (e.g. evaporative control device for motor vehicles).

The components of w(t) can be ordered such that

$$w(t) = \begin{bmatrix} \overline{w}(t) \\ \\ \overline{\overline{w}}(t) \end{bmatrix} \quad \text{and} \quad K = \overline{K} + \overline{\overline{K}}$$

d_{ijt} and w(t) are equivalent notations of control methods for the sources. The index notation d_{ijt} shows explicitly the kinds of sources on which the control methods can be instituted. The vector notation w(t) is convenient for compact mathematical formulations, although it does not show the explicit dependence on sources. w(t) is a single-index variable. It can be obtained by a proper ordering of the double indices i and j of d_{ijt}.#

R(t) = mxK reduction matrix with R_{ij} being the reduction in the emission of pollutant i per unit control $w_j(t)$. (e.g. R_{11} may be the reduction in grams of NO_x emission from pre-1966 model motor vehicles per

For example, let s_1 and s_2 each have two control methods: d_{11}, d_{12} and d_{21}, d_{22} respectively. Let d_{11} and d_{21} be of the temporary type. Then,

$$w(t) = (\overline{w}_1(t), \overline{w}_2(t); \overline{\overline{w}}_1(t), \overline{\overline{w}}_2(t))^T$$

$$= (d_{11}s_1, d_{21}s_2; d_{12}s_1, d_{22}s_2)^T$$

unit of control method w_1 which is the installation of one capacitor-discharge-ignition-optimization system to one pre-1966 model motor vehicle).

$A(t)$ = pxK limited source matrix with A_{ij} being the units of source i controlled by one unit of w_j.

Similar to $w(t)$, $\bar{w}(t)$, $\bar{\bar{w}}(t)$, we define $\bar{R}(t)$, $\bar{\bar{R}}(t)$, $\bar{A}(t)$, $\bar{\bar{A}}(t)$ and we have

$R(t) = (\bar{R}(t), \bar{\bar{R}}(t))$ and $A(t) = (\bar{A}(t), \bar{\bar{A}}(t))$

c_{ijt} = Cost of one unit of control method j for one unit of source i, i= 1,...,p; j= 1,...,q_i, for year t.

$c(t)$ = cost vector of control methods in year t, c_j being the cost of one unit of w_j.

c_{ijt} and $c(t)$ are related to each other as d_{ijt} and $w(t)$ are related to each other. The cost of each control method is to be properly annualized.

$\mu(t)$ = Scaling factor for $c(t)$, so that proper weights can be attached to the cost incurred in each year.

Similarly, we define $\bar{c}(t)$, $\bar{\bar{c}}(t)$, $\bar{\mu}(t)$ and $\bar{\bar{\mu}}(t)$.

$\ell(t)$ = M-dimensional limited supply vector in year t.

$D(t)$ = MxK limited supply coefficient matrix with D_{ij} being the amount of i-th limited supply consumed by one unit of control method w_j.

$x(t)$ = m-dimensional emission vector for year t with x_i being the emission of pollutant i from all the sources in the airshed after institution of controls.

(e.g. $x_1(1)$ may be the tons of NO_x emitted by all the sources during a certain reference time period, say one day, in year 1 (1973)).

$x^o(t)$ = The input x to year t as well as the output x from year (t-1). It is the x(t-1) with only $\bar{\bar{w}}$ (t-1) applied. i.e.

$$x^o(t) = x(t-1) + \bar{R}(t-1)\bar{w}(t-1) \qquad (3.1)$$

$y^o(t)$ = Original emission vector for year t without any control w(i), i = 1,...,t. (e.g. $y^o(3)$ may be tons/day of reactive hydrocarbons (RHC) emitted in the Los Angeles basin in year 3 with no controls, w(1), w(2) and w(3)).

$z(\xi,\tau,t)$ = n-dimensional concentration vector of airborne pollutants. For dynamic airshed model, z may be in units of ppm of pollutants as a function of time and location of the airshed in a typical day of the year. For statistical (static) airshed model, z may be the frequency of violation of an air quality standard (e.g. number of days per year that CO standard is violated in downtown Los Angeles from 6 a.m. to 9 a.m.) during a certain time period of the day and at a certain location of the airshed in any year.

n, m = n is the total number of primary and secondary pollutants and m is the number of primary pollutants.

$g(z(\xi',\tau',t))$ = Air quality vector defining air quality at given reference locations and time (ξ',τ') in year t.

$g^*(t)$ = Air quality criteria vector which is the maximum allowable g(z). g(z) is L-dimensional.

$\alpha(\xi')$ = p-dimensional spatial distribution vector of the sources. $\alpha_i(\xi')$ is the fraction of source i at location ξ' of the airshed.

$\beta(\tau')$ = p-dimensional temporal distribution vector of the sources. $\beta_i(\tau')$ is the fraction of daily activity of source i at time τ'.

Therefore, at (ξ', τ', t) the activity of source i, defined as a p-dimensional vector has component

$$a_i(\xi', \tau', t) = \alpha_i(\xi')\beta_i(\tau')s_i(t)$$

In a simulation model such as the well-mixed cell model, a instead of s is of direct use.

Any airshed simulation model can be represented, in general, by

$$F(z(\xi, \tau, t),\ e(\xi, \tau, t)) = 0 \qquad (3.2)$$

where the parameters a, meteorological data, etc., are assumed known and are not shown explicitly in (3.2). The arguments z and e in (3.2) emphasize the key requirement of an airshed simulation model for control studies, namely given e , z is determinable.

Equation (3.2) may represent a static model, consisting of a set of algebraic equations. It may be a dynamic model, consisting of a set of differential equations or even their solutions. In any event, once

a particular model is decided upon, the explicit form of the model is to be used in place of F.

The mathematical formulation of the long-term problem is given below as a multi-stage optimal control problem:

$$\underset{\substack{w(t),t=1,\\ \ldots,T}}{\text{Minimize } J} = \sum_{t=1}^{T} \bar{\mu}(t)\, \bar{c}^{T}(t)\bar{w}(t) + \bar{\bar{\mu}}(t)\, \bar{\bar{c}}^{T}(t)\bar{\bar{w}}(t) \tag{3.3}$$

subject to the state transition equations

$$x(t) = x(t-1) + \bar{R}(t-1)\bar{w}(t-1) - R(t)w(t)$$

$$= x^{o}(t) - R(t)w(t) \tag{3.4a}$$

$$x^{o}(1) = y^{o}(1) \tag{3.4b}$$

the control technology constraints

$$A(t)w(t) + \sum_{i=1}^{t-1} \bar{\bar{A}}(i)\bar{\bar{w}}(i) \leq s(t) \tag{3.5}$$

$$D(t)w(t) + \sum_{i=1}^{t-1} \bar{\bar{D}}(i)\bar{\bar{w}}(i) \leq \ell(t) \tag{3.6}$$

$$w(t) \geq o, \qquad x(t) \geq o \tag{3.7}$$

the air quality constraints

$$g(z(\xi, r, t)) \leq g^{*}(t) \tag{3.8}$$

and constraints imposed by the airshed simulation model

$$F(z(\xi, r, t), e(\xi, r, t)) = 0 \tag{3.9}$$

where airborne pollutant concentrations z, predicted by the airshed simulation model, are affected by the control measures w(t) through the following equations:

$$E(t)a(\xi, r, t) = e(\xi, r, t) \tag{3.10}$$

where the emission matrix under control action, E is given by

$$E_{ij}(t) = E_{ij}^{o}(t) - \sum_{k=1}^{q_j} r_{ikj}\ (t)\ d_{jkt} \qquad (3.11)$$

$$i = 1,\ldots,m \text{ (pollutants)}$$

$$j = 1,\ldots,p \text{ (source)}$$

where r_{ikj} is defined as the reduction in the emission of pollutant i per unit of control method k for source j. r_{ikj} is similarly related to R in the same way as d_{jk} is related to w. The subscript j shows that these are the control methods and the corresponding reduction in emissions relevant to source j. $E_{ij}^{o}(t)$ is the mass emission of pollutant i from unit source j without control w(t).

(3.3) to (3.11) apply for t = 1,...,T.

The solution of the state transition equations (3.4a) and (3.4b) is

$$x(t) + R(t)w(t) + \sum_{i=1}^{t-1} \bar{\bar{R}}(i)\bar{\bar{w}}(i) = y^{o}(t) \qquad (3.4)$$

Therefore the final mathematical formulation is (3.3), (3.4) and (3.5) through (3.11).

Some brief remarks on the system equations are given below:

In (3.3), the multi-year cost criterion is expressed as a sum of two costs, namely, the cost of temporary control device $\bar{w}$ and the cost of permanent control device $\bar{\bar{w}}$, properly weighted by scaling factors μ according to how the costs of control are to be accounted for in the various years. As an example, consider a two-year problem with only permanent controls (therefore $\bar{\mu} = o$). With c being the annualized

cost, the actual total cost of control incurred during the two-year period can be accounted for by setting $\bar{\bar{\mu}}(1) = 2$ and $\bar{\bar{\mu}}(2) = 1$. On the other hand, to avoid unbias choicesof control methods, we may set $\bar{\bar{\mu}}(1) = 2$ and $\bar{\bar{\mu}}(2) = 2$.

(3.4) says that the controlled emission level in year t ($x(t)$) plus the reduction in the emission levels by all the control methods on the sources up to year t must be equal to the uncontrolled emission level $y^{o}(t)$. In other words, (3.4) specifies the emission reduction requirement of the problem.

(3.5) is a source magnitude constraint, namely, the sources under controlled up to year t can not exceed the available sources. (3.6) is a control method constraint, namely, the amount of control to be used, must not exceed the available resources of control. These constraints are necessary to make the choice of control meaningful, because some control methods may be very favorable and their maximum amount that can be used has to be limited by these constraints.

(3.8) and (3.9) state that the choice of any set of control methods must give an emission level (and distribution) such that the air quality as predicted by the airshed simulation model (3.9) must comply with the air quality standard (3.8).

(3.11) gives the emission level of each of the sources in the airshed due to the institution of control measures. Then, the emission rate vector e to be used in the simulation model, is given by (3.10)

In summary, the mathematical formulations (3.3) to (3.11)

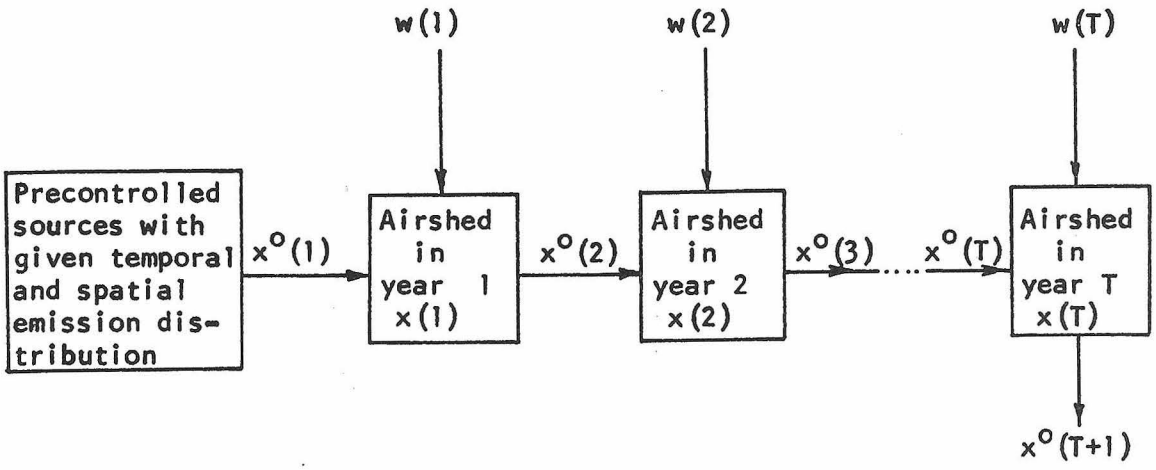
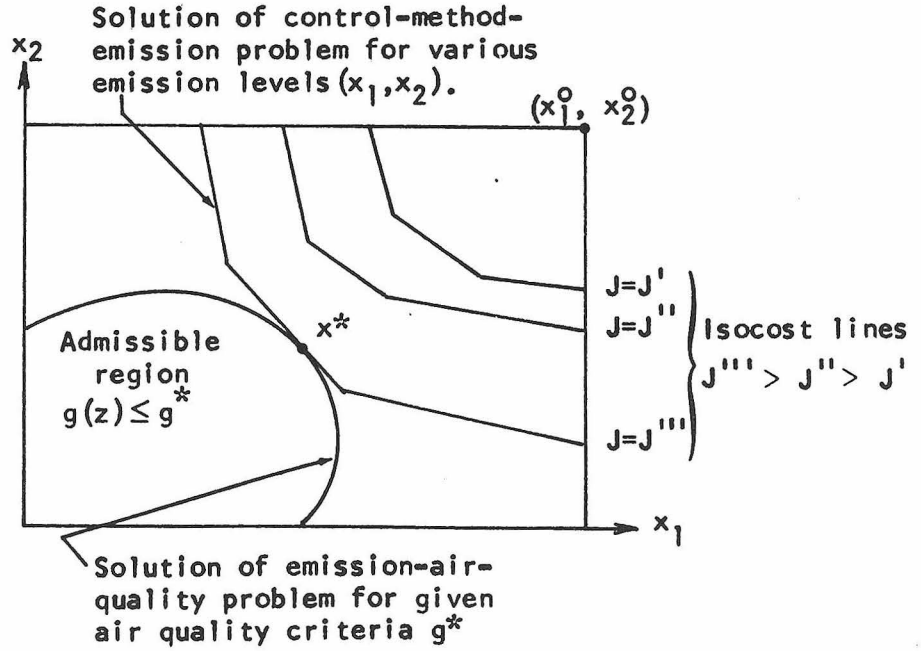


Figure 3.1 Schematic diagram of multi-year problem



x^*= reduced emission level for optimal solution.

Figure 3.2 Single-year minimum cost strategy for two pollutants using linear programming (Trijonis, 1972)

state that the multi-year cost criterion J is to be minimized by choosing control methods w(t), t = 1,...,T, out of all those physically possible such that the air quality standard is satisfied.

Knowing the optimal w(t), t = 1,..T, the controlled status of sources, pollutant emissions and cost etc. are easily evaluated. In short, the objectives as mentioned at the outset of this chapter can be achieved.

For T =1, (3.3) to (3.11) become a single-year problem and for $T \geq 2$, a multi-year problem.

3.3 Methods of solution

Solution methods for the single-year problem and the multi-year problem are different because the system strutures are different. We shall develop them separately.

3.3.1 Single-year problem:

For T = 1, (3.3) to (3.11) reduce to the following single-year problem. The index t for year will be omitted.

$$\underset{w}{\text{Minimize}} \quad J = c^T w \tag{3.12}$$

$$x + Rw = y^o \tag{3.13}$$

$$Aw \leq s \tag{3.14}$$

$$Dw \leq \ell \tag{3.15}$$

$$w \geq 0, \quad x \geq 0 \tag{3.16}$$

$$g(z(\xi, \tau)) \leq g^* \tag{3.17}$$

$$F(z(\xi, \tau), e(\xi, \tau)) = 0 \tag{3.18}$$

$$Ea(\xi, \tau) = e(\xi, \tau) \tag{3.19}$$

$$E_{ij} = E^{o}_{ij} - \sum_{k=1}^{q_j} r_{ikj}\, d_{jk} \qquad (3.20)$$

$i = 1,\ldots,m$ (pollutants)

$j = 1,\ldots,p$ (sources)

For a static airshed simulation model (3.18) is an algebraic relationship and (3.12) to (3.20) becomes a standard mathematical programming problem. For a dynamic airshed model, (3.18) will be a set of differential equations and (3.12) to (3.20) becomes a typical optimal control problem (if $z(\xi, r)$ is considered as a state variable) with inequality state and control variable constraints. In principle, computational methods exist for either case. However, due to the large dimensionality of the control vector w and the source vector s, usually associated with an airshed and the inherent nonlinearity of the (dynamic) simulation model, adaptation of the existing computational methods with special regard to the structure of the system is necessary.

System (3.12) to (3.20) consists of three distinct parts, each of which can be considered as a subproblem. Part (1) is (3.12) to (3.16). This part is linear and describes the relationship of control methods w, control cost J and emission level x. This relationship depends only on the characteristics of the sources and the control methods. This is a linear programming problem, the structure of which results implicitly from several assumptions inherent in the definitions of the various quantities. We have assumed, for example, that the cost per unit of control method c is independent of the number of units of

control w. Also, the reduction in the emissions of the species R is also independent of the level of control w. Similarly, the amount of limited supply inputs consumed is independent of the level of control. Part (2) is the system (3.17) to (3.18). This part is nonlinear in general and describes the relationship of emission distribution e (and not just the emission level x) and air quality g. Part (3) is the system (3.19) to (3.20), which describes the relationship between the emission level x and the emission distribution e and thus bridges part (1) and part (2).

For two-dimensional x, a convenient (and illustrative) computational method is the following graphical solution:

Graphical single-year algorithm

Step 1. Solve (3.12) to (3.16) for various values of x by linear programming. For each value of x, record the optimal cost J and the optimal control methods w corresponding to the reduction in emission as specified by x. We thus generate the emission level x and control cost relationship.

Step 2a. For each value of x in step 1, using the optimal w, compute E and then e by equations (3.19) and (3.20). We thus generate the relationship of x and e.

Step 2b. For each value of e in step 2a, compute air quality g by (3.17) and (3.18).

Steps 2a and 2b, in effect, generate the relationship between the emission level x and air quality g.

Step 3. Superimposing the results obtained in steps 1 and 2,

the optimal solution is given by that x which satisfies the air quality criteria at least cost.

The above computational method involves essentially the decomposition of the single-year problem into two problems. One is a linear programming problem relating optimal control cost and reduction of pollutant emissions. The other is a dynamic optimization (or nonlinear programming, if static airshed model is used) problem relating reduction of pollutant emissions and air quality. The decomposition of the single-year problem into two steps is advantageous from a computational point of view. In general, there will be many control methods, so that dimensionality problems can be expected. In addition, the airshed simulation model is generally nonlinear (e.g. for dynamic airshed model, it involves n differential equations, nonlinear if chemical reactions are occurring). Since linear programming can handle a large number of variables easily, it makes sense to separate the high dimensional, linear part of the problem from the nonlinear airshed model which is generally lower dimensional. The two-step solution in the foregoing represents such a separation.

The above computational method is applicable for x being two dimensional or less. Figure 3.2 illustrates the form of the results of the foregoing calculations for a problem involving two kinds of primary pollutants.

The above method has also been used by Trijonis (1972). For many practical air pollution control problems for a particular year, this method is sufficient and very useful.

If the dimension of x is greater than two, the above graphical method is no longer convenient. We shall develop a gradient method for this case based on the following observations:

Observation 1: Supposing the precontrolled source emission is such that the air quality constraint is violated. Then, the optimal policy $w = w^*$ has the property that at least one component of the air quality vector g is equal to its constraint g^*.

Observation 2: For any given set of controls, w, the corresponding components of g are not mutually independent. That is, g_i can not be arbitrarily varied while keeping other g_j fixed.

By the above observations, the key to a computational method is to find the w that minimizes J and such that the air quality constraint is just satisfied. A simple first order gradient method is proposed below:

Gradient (programming) single-year algorithm

Step 1. Assume an initial $x = x^{old}$ (e.g. at the first iteration, we may set it to be y^o). Solve for w from (3.12) to (3.16) by linear programming. Using w so obtained, determine the corresponding air quality $g = g^{old}$.

Step 2. Perturb each component of x^{old} one at a time and repeat step 1. Then, evaluate the Jacobian matrix $\partial g/\partial x$ numerically by

$$\partial g_i/\partial x_j = (g_i^{old} - g_i)/(x_j^{old} - x_j) \qquad i = 1,\ldots,L.$$

where $x_j = x_j^{old} - \epsilon$ and g corresponds to $x = x^{old} - \epsilon(\delta_{1j},..,\delta_{kj},..,\delta_{Mj})^T$ with δ being the Kronecker δ. It is to be emphasized that perturbation

of the component x_j gives the j-th column of the Jacobian matrix by the foregoing formula.

Step 3. Evaluate new x by solving with linear programming, (3.12) to (3.16) along with

$$(\partial g/\partial x)(x - x^{old}) = -\kappa(g^{old} - g^*) \qquad 0 < \kappa \leq 1$$

Step 4. Repeat steps 1 to 3 with the new x until $g \leq g^*$ and with at least one component of g satisfying the equality.

The question of convergence is irrelevant here, because we are merely iteratively decreasing g until $g = g^*$ by reducing x. We already know where to start the iteration and in what direction. By decreasing a vector in the above context, we mean decreasing each component of the vector. Another point to note in the above algorithm is that the effect of the airshed simulation model is embodied in step 2 and thus the algorithm is applicable to any airshed simulation model.

3.3.2 The multi-year problem

We have presented the general theory of determining the minimum cost set of controls to achieve a specified level of air quality for a single year (subject to the assumption inherent in the linear programming approach). In general, air pollution legislation will prescribe control actions for a number of years. The problem is to determine the combination of controls over a T year period that minimizes the total cost of control over the T-years while maintaining a specified level of air quality each year. One way of approaching the problem is to consider each year as independent of the others and solve a single

year problem for each of the T years. Since many control methods involve installation of equipment, however, the decision to install such equipment in an early year without considering its impact in later years may unnecessarily constrain our freedom to act in later years. Thus, control decisions made on the basis of what is optimal for this year may not be optimal over a long period. We would like to consider the optimal allocation of controls over all T years.

In section 3.2, we have distinguished two types of control measures. The first type of control measures $\bar{w}$ are those which can be undertaken on a yearly basis independently of the control measures used in any other year. This type of control arises normally for those sources the emission control of which depends on the grade or nature of raw materials used. For example, SO_2 emissions from power plants can be controlled by burning low sulfur fuel oil or natural gas in place of coal and high sulfur fuel oil. The amount of low sulfur fuel oil or natural gas burned in any year, while limited by the total amount available, should not necessarily depend on the amount burned in previous years. Therefore, for this type of control measure the decision on the level of the measure is made on a year-to-year basis, and the cost of the control is borne completely in the year in which the control action is taken.

The second type of control measures $\bar{\bar{w}}$ are those which, once instituted, remain for a fairly long period of time. Such measures include, for example, improving or changing the operating conditions of a process or adding a new piece of equipment for cleaning effluents.

An example of such a control measure is the installation of a catalytic muffler on a car which will be expected to remain on the car for its life. The decision on whether to institute this type of control measure must take into account future years when the control action is still in effect. In addition, the cost of such a control measure is not normally totally borne in the year of purchase, but rather amortized over the life of the device.

If all the control measures were of the first type, then the problem of choosing air pollution control strategies for a T-year period would become one of choosing controls for T individual and independent years, since a control measure used in year t would not necessarily depend on control measures used in year t-1.

In the most general case the control methods are of both types. Therefore, the choice of what controls to employ in a particular year will be affected by prior choices and the consideration of their future impact. Unfortunately, this makes the optimal control problem much more difficult since the various years cannot be treated independently.

The mathematical formulation for the multi-year problem is (3.3) to (3.11). In the sequel, we shall develop four computational methods for its solution.

Backward dynamic programming algorithm

When the number of sources involved is not large (such as two or three), an efficient computational method can be developed using (backward) dynamic programming. In this section, we shall employ

slightly different notation for convenience and explicitness.

As defined in section 3.2, the control methods are represented by the variables d_{ijt}, $i = 1,2,\ldots,p$; $j = 1,2,\ldots, q_i$; $t = 1,2,\ldots, T$. Of the q_i control methods for source i we assume that q_i' are of type one (corresponding to temporary control methods $\bar{w}$ whose use is independent of prior years) and q_i'' are of type two (corresponding to the permanent control methods $\bar{\bar{w}}$ which once installed remain for more than one year). Then, $q_i = q_i' + q_i''$ since all the control methods must be one of the two types. We will order the controls such that the first q_i' are those of type one and the last q_i'' are of type two.

The multi-year cost functional J stated in equation (3.3) in section 3.2 can be alternatively expressed as

$$J = \sum_{t=1}^{T} J_t = \sum_{t=1}^{T} C_t + B_t \qquad (3.21)$$

where the total cost of control paid for in year t is a sum of two costs:

C_t= the cost of controls instituted for the first time in year t. These controls may be of either type: in the first case the entire cost of the control is used; in the second case only that amount of the cost attributable to the first year is included in C_t

B_t= the cost of controls instituted in prior years which are still being paid for in year t.

In the case in which all controls are paid for completely in the year they are installed, $B_t = 0$, $t = 1,2,\ldots, T$.

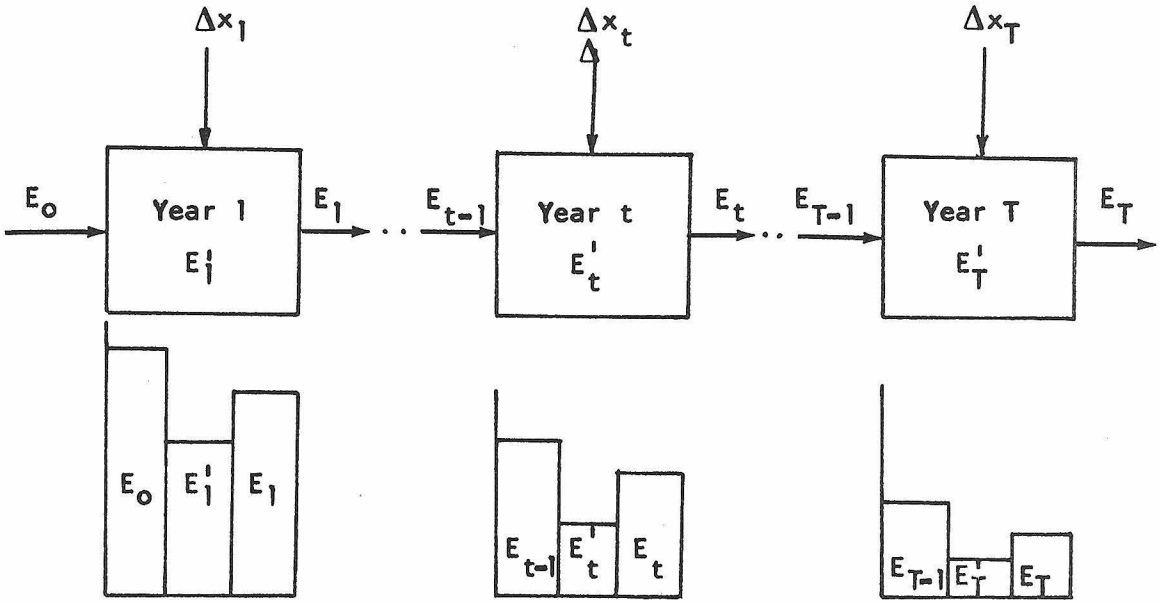
The multi-year control problem is essentially a sequence of interrelated single-year problems, that is, a discrete multi-staged optimal control problem, with each year representing a stage. Instead of taking the d_{ijt} as the control variables, let us select the total daily reduction of pollutant k in year t, Δx_{kt}, as the control variables. This choice is made because of the much lower dimension of m, which is at most, 3 or 4, as compared to p and q_i. Moreover, the determination of the least cost value of Δx_{kt} will automatically generate the d_{ijt}.

In the backward version of dynamic programming, we shall need a state variable which can fully describe the airshed system. For this purpose, we shall choose E_t as the state variables for the multi-year sequence, instead of x_t as given in the general formulation. Although E_t has a larger dimension than x_t, once E_t is specified, x_t is simply given by

$$x_t = E_t s_t \tag{3.22}$$

and x_t, in turn , completely specifies w(t) and g(t) by the single-year problem. For later convenience, we have written E(t) as E_t, x(t) as x_t, etc.

Consider Figure 3.3, in which we have depicted a T-year situation with one pollutant (m = 1) and one source (p = 1). The input to year 1 is denoted by E_o and represents the uncontrolled level of emission of the source. The controlled level of emission in year 1 is denoted by E_1'. This controlled level represents the effect of all



E_{t-1} = Daily mass of emission in year t before control (input to year 1)
E'_t = Daily mass of emission in year t after control (state of year t)
E_t = Daily mass of emission in year t+1 before control (output of year t).

Figure 3.3 The T-year air pollution control problem for a single source and one pollutant. The stagewise representation of the discrete years is shown above. The year-by-year reduction in the daily mass emissions is shown below.

control measures, those of types one and two. The input to year 2, i.e., the output from year 1, is, however, the controlled level of emission resulting only from those controls of type two.

For the general case,

$$E'_{kit} = E_{ki,t-1} - \sum_{j=1}^{q_i} r_{kji}\, d_{ijt} \tag{3.23}$$

$$E_{kit} = E_{ki,t-1} - \sum_{j=q'_i+1}^{q_i} r_{kji}\, d_{ijt} \tag{3.24}$$

$i = 1,\ldots, p$ (sources)

$k = 1,\ldots, m$ (pollutants)

$t = 1,\ldots, T$ (years)

Since the summation in (3.23) is to q_i, all the control actions used in year t are used to compute E'_{kit}. However, in order to determine the input to year t + 1, that is, the output from year t, only those controls of the second type (those which involve capital equipment, etc.) will carry over.

The problem is to choose mT reductions Δx_{kt}, such that (3.21) is minimized subject to the air quality and other constraints as stated in the general mathematical formulation in (3.4) to (3.11).

As is customary in (backward) dynamic programming, let us begin with the final stage, year T. Given any input E_{T-1}, we desire to find Δx_T s.t. $J_T = C_T + B_T$ is minimized subject to the air quality and other constraints as stated in the single-year formulation (3.13) to (3.20). Let

$$f_1(E_{T-1}) = \min_{\Delta x_T} \left\{ C_T + B_T \right\} \tag{3.25}$$

where $C_T = C_T\ (E_{T-1}\ , \Delta x_T)$ and $B_T = B_T\ (E_o, \ldots\ E_{T-2}\ ; \Delta x_1, \ldots \Delta x_{T-1})$.

Note that B_T, the cost of controls of type 2 instituted prior to year T, depends only on the prior values of E and Δx. Thus, (3.25) becomes

$$f_1(E_{T-1}) = \min_{\Delta x_T} \left\{ C_T\ (E_{T-1}\ , \Delta x_T) \right\} + B_T\ (E_o\ , \ldots,\ E_{T-2}\ ; \Delta x_1, \ldots, \Delta x_{T-1}) \tag{3.26}$$

Proceeding backward, the general recurrence relation for any stage t is

$$f_{T-t+1}\ (E_{t-1}) = \min_{\Delta x_t} \left\{ C_t\ (E_{t-1}, \Delta x_t\ ) + f_{T-t}\ (E_t) \right\} + B_t\ (E_o\ , \ldots,\ E_{t-2};\ \Delta x_1, \ldots, \Delta x_{t-1}\ ) \tag{3.27}$$

with state transition equation given by (3.24)

To perform the minimization in (3.27), the single-year optimization procedure must be used, and only those Δx_t are used in the search that satisfy the air quality constraints in year t.

The backward dynamic programming algorithm is then as follows:

Step 1. Choose discrete values of E as $E^1, E^2, \ldots, E^K$. Corresponding to each of the E^K, for $t = T$, determine the optimal Δx_T and the minimum C_T in (3.26), using the linear programming solution of the single-year problem. Let this minimum C_T be denoted by $C_T^*(E_{T-1})$. Clearly, C_T^* only includes the cost of controls instituted in year T. The control and cost values are stored in year-T table.

Step 2. Proceeding backward to year T-1, we want to

determine

$$f_2(E_{T-2}) = \min_{\Delta x_{T-1}} \left\{ C_{T-1}(E_{T-2}, \Delta x_{T-1}) + f_1(E_{T-1}) \right\} + B_{T-1}(E_0, \ldots, E_{T-3}; \Delta x_1, \ldots, \Delta x_{T-2}) \tag{3.28}$$

Substituting (3.26) into (3.28) gives

$$f_2(E_{T-2}) = \min_{\Delta x} \left\{ C_{T-1}(E_{T-2}, \Delta x_{T-1}) + C_T^*(E_{T-1}) + B_T(E_0, \ldots, E_{T-2}; \Delta x_1, \ldots, \Delta x_{T-1}) \right\} + B_{T-1}(E_0, \ldots, E_{T-3}; \Delta x_1, \ldots, \Delta x_{T-2}) \tag{3.29}$$

We note that B_T is that portion of the cost of controls for years 1,2,..., T-1, that is to be paid for in year T. It can be decomposed into

$$B_T(E_0, \ldots, E_{T-2}; \Delta x_1, \ldots, \Delta x_{T-1}) = B_{T-1}(E_0, \ldots E_{T-3}; \Delta x_1, \ldots, \Delta x_{T-2}) + H_T(E_0, \ldots, E_{T-2}; \Delta x_1, \ldots \Delta x_{T-1}) \tag{3.30}$$

where $B_{T-1}(E_0, \ldots, E_{T-3}; \Delta x_1, \ldots, \Delta x_{T-2})$ is the contribution of the cost of controls instituted in years 1,2,.., T-2 to the total cost in year T and $H_T(E_0, \ldots, E_{T-2}; \Delta x_1, \ldots, \Delta x_{T-1})$ is the contribution of the cost of controls instituted in year T-1 to the total cost in year T. Depending on the scheme of amortization of the control cost, H_T

is a known function of C_{T-1} i.e.

$$H_T = H_T\,(C_{T-1}\,(E_{T-2}\,, \Delta x_{T-1})) \tag{3.31}$$

Using (3.31), (3.29) reduces to

$$f_2\,(E_{T-2}) = \min_{\Delta x_{T-1}} \left\{ C_{T-1}\,(E_{T-2}, \Delta x_{T-1}) + H_T\,(C_{T-1}\,(E_{T-2}, \Delta x_{T-1})) \right.$$

$$\left. + C_T^{*}\,(E_{T-1}) \right\} + 2B_{T-1}\,(E_o\,, \ldots, E_{T-3}\,; \Delta x_1, \ldots, \Delta x_{T-2}) \tag{3.32}$$

where quantities independent of Δx_{T-1} have been moved outside the minimization parathesis. E_{T-1} is related to E_{T-2} through the state transition equation (3.24) where $d_{ij,T-1}$ are the optimal control measures corresponding to any Δx_{T-1}.

According to each of the discrete values of E as chosen in step 1, we perform a minimization of (3.32) over Δx_{T-1}. We then store the control and cost values in the table for year T-1.

Step 3. The procedure in step 2 is repeated until year 1 is reached.

Step 4. The optimal solution for a given E_o of the multi-year problem is then obtained by a forward sweep of the tables constructed for years 1 to T.

The minimization over Δx_t in the above algorithm may be carried out by a suitable search algorithm (Wilde, 1964). e.g. if Δx_t is one-dimensional, a Fibonacci search may be used.

Due to the usual dimensionality problem associated with dynamic programming, the usefulness of this algorithm is limited by the

number of pollutants and the number of sources involved. The principal drawback is its dependence on the number of sources which can be large for many air pollution control systems. To remove this dependence, we propose the next algorithm given below.

Forward dynamic programming algorithm

Choosing x(t) as the state variable, we can develop a forward version of the dynamic programming solution the usefulness of which will not be limited by the number of sources present, since x(t) is a vector of the emission levels of pollutants from all the sources. The motivation for using the forward version is that x is but a quasi-multi-stage variable (in the sense that knowing x does not completely specify the system) and in transiting from year to year, we need also to know the kinds and amount of controls already instituted in the prior years. However, if we use E_{kj} (emission of pollutant k from source j) as state variables, then the backward version is feasible, since E completely defines our airshed system. However, using E as state variable causes dimensionality problems as we had noted before. The choice of x as the state variable makes the algorithm applicable for most of the practical cases. e.g. for photochemical smog, we need only consider the emissions of NO_x and RHC with x being 2-dimemsional. For an inert pollutant, the dimension of x is only one. On the other hand, since coupling of the inert and active pollutants can occur through the control vector, there may be cases where we want to consider inert and active pollutants together and therefore the dimension of x may be large. In the latter case, other computational methods will be given.

To begin with, let us define

$f(x(t))$ = Optimal cost of controls from year 1 to year t, when the controlled state variable of year t is $x(t)$.

$h(x(t),w(t))$ = Cost of control incurred in year t when the state of previous year is $x(t-1)$.

Then by the principle of optimality of forward dynamic programming, we have the following functional equation:

$$f(x(t)) = \min_{w(t)} \left[h(x(t-1), w(t)) + f(x(t-1))\right] \qquad (3.32)$$

$$t = 2,\ldots, T.$$

with initial condition

$$f(x(1)) = \text{known for any } x(1) \qquad (3.33)$$

and the state transition equation:

$$x(t) = x(t) - R(t)w(t) \qquad (3.34)$$

or

$$x(t) = x(t-1) + \bar{R}(t-1)\bar{w}(t-1) - R(t)w(t)$$

The forward dynamic programming algorithm is then as follows:

Step 1. Select a discrete set of values for x as $x^1,\ldots, x^k, \ldots, x^K$, with x^K at least as big as $\max_t y^o(t)$. For year 1, compute $f(x^k(1))$, $k = 1,2,\ldots, K$ as the minimum control cost of the single-year problem (3.12) to (3.16) corresponding to $x(1) = x^1, x^2,\ldots, x^K$ respectively. Also compute air quality g from (3.17) to (3.20) for each of the x^k, $k = 1,\ldots, K$.

Store all the results of control cost,control measures and air quality corresponding to each of the x^k in year-1 table. Set $f(x^j(1))$ to be arbitrarily large for those $x^j(1)$ which correspond to $g(1) > g^*$.

Step 2. For year 2, we want to solve

$$f(x(2)) = \min_{w(2)} \left[h(x(1), w(2)) + f(x(1)) \right] \tag{3.35}$$

where $f(x(1))$ are known in step 1 for each of the $x^k(1)$, $k = 1,..,K$. The minimization over $w(2)$ is extremely difficult since $w(t)$ is usually of large dimension. However, we are not interested in all possible sets of control measures $w(2)$, but are only interested in those sets of control measures $w(2)$ which can reduce the input state $x(1)$ to $x(2)$ in an optimal fashion. The minimization scheme for (3.35) is then as follows:

For $x(2) = x^k$ and $x(1) = x^j > x^k$, determine $h(x(1),w(2))$ as the optimal control cost J^{jk} of reducing the emission level x^j in year 1 to the emission level x^k in year 2, by solving the following single-year problem by linear programming:

$$J^{jk} = \min_{w(2)} \left[\bar{\mu}(2)\ \bar{c}^T(2)\ \bar{w}(2) + \bar{\bar{\mu}}(2)\ \bar{\bar{c}}^T(2)\ \bar{\bar{w}}(2) \right] \tag{3.35}$$

$$R(2)w(2) = \bar{R}(1)\ \bar{w}(1) + x^j(1) - x^k(2) \tag{3.37}$$

$$A(2)w(2) \leq s(2) - \bar{\bar{A}}(1)\bar{\bar{w}}(1) \tag{3.38}$$

$$D(2)w(2) \leq \ell(2) - \bar{\bar{D}}(1)\bar{\bar{w}}(1) \tag{3.39}$$

$$w(2) \geq o \tag{3.40}$$

where the only unknown is w(2). $f(x(1) = x^j)$ is known in step 1; therefore the scalar function defined as

$$f^j(x(2) = x^k) = h(x(1) = x^j, w(2)) + f(x(1) = x^j) \tag{3.41}$$

as well as w(2) and g(2) are known. (3.41) can be evaluated for j=1, ..., K. Then $f(x(2)=x^k)$ as defined by (3.35) is given by

$$f(x(2) = x^k) = \min_{j \in (1,..,K)} \left[f^j(x(2) = x^k) \right] \tag{3.42}$$

Carrying out (3.42) for $x(2) = x^k$, k = 1,...,K completes the evaluation of f(x(2)). Store control cost, control measures and air quality for each $x^k(2)$ in year-2 table.

Step 3. Repeat step 2 for t = 3,4,...,T.

Step 4. The minimum J(T) is then the smallest f(x(T)) which satisfies $g(T) \leq g^*(T)$. The optimal control strategies w(t), t = 1, ...,T can be obtained by one backward sweep of the tables constructed for years T, T-1,...,1.

Comment on the algorithm: Major computing effort is in evaluating (3.42). The main drawback is that minimization of (3.42) can only be conveniently done over the fixed set (1,...,K) corresponding to $x^1, x^2, \ldots, x^K$. Interpolation between two x^k, x^{k+1}, though possible in principle, is not computationally feasible, since the control method w(1) needed for (3.36) to (3.40) can not be obtained by interpolation.

If the dimension of x is too large, the following algorithm

is suggested.

Gradient (programming) multi-year algorithm

This algorithm closely follows the concept of the single-year gradient algorithm.

Step 1. Guess an intial set of $x(t)$, $t=1,..,T$. Solve (3.3) to (3.7) by linear programming. Using the $w(t)$, $t=1,...,T$, obtained from the linear programming solution, determine the corresponding $g(t)$, $t=1,...,T$ by (3.8) to (3.11)

Step 2. As was done in the single-year gradient algorithm, evaluate numerically

$$M(t) = \partial g(t)/\partial x(t) \qquad t = 1,...,T$$

Step 3. Evaluate new $x(t)$ by solving (using linear programming) (3.3) to (3.7) along with

$$M(t)(x(t)^{new} - x(t)^{old}) = \kappa(g(t) - g^*(t)),\ t=1,...,T$$

$$0<\kappa\leq 1$$

Step 4. Repeat steps 1 to 3 with $x(t)^{new}$, until $g(t) \leq g^*(t)$ and at least one component of $g(t)$ satisfying the equality, for $t = 1, ...,T$.

A simplified algorithm for multi-year problem

If (3.8) and (3.9) are linear or can be approximated by linear relationships in the vicinity of x where $g = g^*$, then (3.3) to (3.11) reduce to a big linear programming problem and optimal $w(t)$, $t = 1,..,T$ can be generated simultaneously by the Simplex method.

CHAPTER 4. REAL-TIME AIR POLLUTION CONTROL

The real-time control of air pollution is of importance to develop emergency alert programs of source emission control procedures for polluting sources in an airshed, to maintain a desired air quality and to provide preventive or remedial actions to counteract any adverse meteorological conditions. The development of a general framework for the determination of optimal real-time air pollution control strategies is the objective of this chapter.

This chapter consists of three parts: (1) the formulation of a general real-time air pollution control problem, (2) the development of a computational algorithm for solving the class of control problems which result, and (3) an application of the theory to a hypothetical study of the effect of implementation of the optimal control on September 29, 1969 in the Los Angeles basin. It is assumed that a mathematical model of pollutant behavior which includes provisions for dynamic meteorology and atmospheric chemical reactions exists for the airshed. The particular type of model utilized for this study consists of an array of well-mixed cells, as detailed in chapter 2 although the theory is applicable to other types of mathematical air pollution models. Based on the airshed model the real-time control problem is formulated as choosing the types and levels of control actions as a function of time and location in the airshed based on predicted meteorology such that a certain level of air quality is maintained over a given time period and with minimum necessary control action. In the hypothetical study of real-time control for Los

Angeles, the pollutant species considered are carbon monoxide, nitric oxide, nitrogen dioxide, reactive hydrocarbons and ozone. The two control measures assumed to be available were reductions in the number of cars permitted to use freeways and in the amount of fuel burned in the basin's power plants. Various reductions in ozone levels that would have been reached during the day are seen to result from implementation of the optimal strategy. The significance of this chapter lies in the framework it provides for the subsequent use of airshed simulation models in control strategy evaluation as such models become available.

4.1 General consideration of real-time control

Several points can be noted about the real-time problem:

(1) This is a dynamic problem in which we are concerned with emissions and pollutant concentrations over time scales of a few hours to a few days;

(2) The emergency control measures are in general more severe than emission controls normally in effect and would be only of short duration; and

(3) The measures must be capable of being instituted rapidly and effectively.

In principle, a strategy could be designed on the basis of feedback or feedforward control.

In this context, feedback control would imply that we institute control action on the basis of measured atmospheric pollutant concentrations. Thus, we would essentially have to wait until concentrations begin to get serious before taking action, at which time it is usually

too late to forestall high concentrations. In general, the tremendous sluggishness of an urban airshed precludes feedback control from being effective.

Feedforward control would imply that we institute control action on the basis of measured meteorological conditions such as wind speed and inversion depth. The distinction between feedback and feedforward control lies in the definition of the system. The airshed is the system, the state of which is the set of pollutant concentrations, the controllable inputs to which are the source emissions, and the uncontrollable inputs to which are the meteorological variables. Feedback control would be based on state (concentration measurements) whereas feedforward control would be based on uncontrollable input (weather factors) measurements. Feedforward control is favorable to feedback control for this problem because we can act before concentrations actually build up. In feedforward control we might, for example, measure wind speeds and inversion depth every few hours as a basis for setting control actions over the ensuing few hours until the next measurements. It is this type of control we will consider here.

4.2 DYNAMIC AIRSHED MODELS-SYSTEM EQUATIONS-

In order to determine the relationship between emission levels and air quality, a mathematical representation of pollutant behavior in the atmosphere is required. There is currently much interest in the mathematical modeling of urban air pollution. A general survey of the subject has been presented by Seinfeld et al. (1972), and studies (of varying approaches and degrees of success) on modeling

specific urban areas are found in Lamb and Neiburger (1971), Randerson (1970), Eschenroeder and Martinez (1971), Roth et al. (1971), and Reynolds, et al. (1973). Most of these studies are based on the numerical solution of some form of the partial differential equations of continuity for the mean concentrations of pollutant species in a turbulent fluid. Because this approach to air pollution modeling is still in a state of development, we have chosen to employ the simpler well-mixed-cell model given in chapter 2. We do this primarily to facilitate our real objective in the present work—the study of real-time control. Therefore, the airshed model to be used may not ultimately be the most desirable but, nevertheless, is a conceptually simple one which includes provisions for sources, meteorology and chemistry.

The system equations for the real-time control problem based on the well-mixed cell model are then equations (2.1) and (2.2) as given in chapter 2. The deposition term in equation (2.1) will be neglected, as we will consider only gaseous pollutants. For conciseness we express them in vector notation as

$$\dot{z}(t) = A(t)z(t) + B(t)b(t) + r(z) + Eu \qquad (4.1)$$

$$z(t_o) = z^o \qquad (4.2)$$

Where $z(t)$ is the nK-dimensional column vector, $(z_{11}, z_{12}, \ldots, z_{1K}, z_{21}, \ldots, z_{nK})^T$, n being the total number of pollutants involved and K the number of cells into which the airshed is conceptually divided. The nK x nK matrix A is defined by

$$A = \begin{bmatrix} D^T & & 0 \\ & D^T & \\ & & \ddots \\ 0 & & D^T \end{bmatrix} \tag{4.3}$$

where the KxK matrix

$$D = \begin{bmatrix} -\frac{1}{v_1}(\dot{v}_1 + q_{1j}) & \frac{q_{12}}{v_2} & \cdots & \frac{q_{1K}}{v_K} \\ \frac{q_{21}}{v_1} & & & \\ & & \ddots & \\ \frac{q_{K1}}{v_1} & \cdots & -\frac{1}{v_K}(\dot{v}_K + q_{Kj}) & \end{bmatrix} \tag{4.4}$$

Also, the nKxnK matrix

$$B = \begin{bmatrix} P & & \\ & P & \\ & & \ddots \\ & & P \end{bmatrix} \tag{4.5}$$

where the KxK matrix

$$P = \begin{bmatrix} \frac{q_{01}}{v_1} & & 0 \\ & \frac{q_{02}}{v_2} & \\ & & \ddots \\ 0 & & \frac{q_{0K}}{v_K} \end{bmatrix} \tag{4.6}$$

As defined in Chapter 2, z_{ik} is the mean concentration of pollutant i in cell k, v_k is the volume of cell k and q_{jk} is the volumetric wind flow from cell j to cell k. Subscript "0" refers to location outside the airshed. b is an NK-dimensional vector of pollutant concentrations outside the airshed, $(z_{10}, z_{10}, \ldots, z_{10}, z_{20}, \ldots, z_{20}, \ldots, z_{n0})^T$, $r'(z)$ is an nK-dimensional vector of reaction rates, u is an nK-dimensional vector of the mass emissions of the n pollutants in each of the K cells (i.e. if we let e_{ik} be the mass emissions of pollutant i per hour in cell k, then $u = (e_{11}, e_{12}, \ldots, e_{1K}, e_{21}, e_{22}, \ldots, e_{2K}, \ldots, e_{nK})^T = (u_1, u_2, \ldots, u_{nK})^T$, where $u_{(i-1)K+k} = e_{ik}$), and E is an nKxnK matrix required to convert the emissions (mass/time) into concentrations (patrs-per-million/time).

4.3 Statement of the problem

The problem we wish to solve is the following: Given a meteorological forecast from time t_0 to time t_f, determine the set of control measures applied to polluting sources over (t_0, t_f) such that a given set of air quality criteria are not violated and the amount of required emission reduction is a minimum.

The meteorological parameters in the airshed model are A(t), B(t), b and any that may occur in $r'(z)$, such as the temperature or intensity of radiation. Thus, at time t_0 measurements of wind speeds and directions, inversion heights, temperatures, etc. are assumed to be available. On the basis of these measurements, A(t), b, etc. are forecasted over (t_0, t_f), so that henceforth we will assume these quantities are given. A reasonable duration for the predictions would be

two hours.

A general measure of air quality at any time is some prescribed function g(z). For example, if air quality is given by the concentration of species 1 in each of the K cells, then g(z) is simply equal to $(z_{11}, z_{12},\ldots,z_{1K})^T$. The maximum allowable value of g can be called g*.

Pollutant emissions enter the airshed model through u(t), the mass emissions of each contaminant in each cell as a function of time. The uncontrolled level of emission can be denoted by $u_0(t)$. We could pose the real-time control problem as minimizing some measure of the deviation between the normal level of emissions $u_0(t)$ and that needed for control u(t) subject to (4.1) to (4.6) and $g(z) \leq g^*$.

As stated, this problem will yield the maximum allowable mass emission levels, $u_{nK}(t)$, over the time interval (t_0, t_f), needed to maintain a certain air quality index g*, given meteorological information over the interval, i.e. A(t),B(t) and b. The solution to this problem will not, however, tell us <u>what</u> controls to impose - it only tells us what maximum <u>mass emission levels</u> e_{nk} of each pollutant in each cell can be tolerated while still maintaining g(t) below g*. We have no guarantee that these mass emission levels can, in fact, be reached in the necessary proportions with existing control methods. This is a key point in what follows. The reason is that the emission reductions of various primary pollutants achieved with any control method are not <u>independent</u>. For example, an automobile emits carbon monoxide, hydrocarbons, and oxides of nitrogen in certain relative proportions dependent on the age of the car, its condition, etc.

These emissions cannot be altered independently by the types of strategies available for real-time control, such as reducing freeway traffic; a change in the driving patterns in an area will affect all emissions in a fixed manner. For this reason, it is necessary to consider as control variables not simply the total mass emissions u(t) but rather the level of employment of the actual control methods. For example, u_1 might represent the mass emissions of CO (species 1) in cell 1. This value is a result perhaps of a number of control methods acting on several sources of CO in cell 1. Merely determing u_1 will not tell us either how to achieve that value of u_1 or, in fact, even if that value is attainable given existing control methods. We must therefore enumerate the feasible control methods for each source, as well as all the important sources in the airshed.

In order to include information relating to the sources and their controls we introduce the following definitions:

p = number of source types in the airshed (e.g. 1970 motor vehicles, power plants, etc.)

q_i = number of control methods available for source i, $i = 1,2,\ldots, p$.

s_{ik} = magnitude of source i in cell k (e.g. the number of 1970 motor vehicles in cell k) $i = 1,2,\ldots, p$; $k = 1,2,\ldots,K$.

d_{ijk} = level of control activity j on source type i in cell k (e.g. the number of 1970 motor vehicles prohibited from freeway use in cell k) $j = 1,2,\ldots, q_i$; $i = 1,2,\ldots,p$; $k = 1,2,\ldots,K$.

$r_{ijn'}$ = the reduction in the mass emission of pollutant n by application of one unit of control method j for source i (e.g. the pounds/hr of CO reduced per 1970 motor vehicle prevented from freeway use in cell k) $j = 1,2,\ldots, q_i$; $i = 1,2,\ldots, p$; $n' = 1,2,\ldots, n$.

ω_{ij} = the number of units of source i controlled by one unit of control method j (e.g. $\omega_{11} = 1$ if one 1970 car is prevented from freeway use).

Therefore, the d_{ijk} are the variables which represent the level of source control by each method. We assume that the parameters above can be taken as constants, independent of the level of control.

We can now relate the overall mass emissions u_ℓ, $\ell = (n'-1)K+k$, to the individual sources and their controls by

$$u_{\ell_o} - u_\ell = \sum_{i=1}^{p}\sum_{j=1}^{q_i} r_{ijn'}\, d_{ijk}\, s_{ik}, \quad \ell = (n-1)K + k \qquad (4.7)$$

$$n' = 1,2,\ldots, n$$

$$k = 1,2,\ldots, K$$

where, as we noted previously, u_{ℓ_o} is the uncontrolled level of emission. (4.7) may be expressed somewhat more concisely as

$$u = u_o - Gw \qquad (4.8)$$

where G is a nK by $K.\Sigma q_i$ matrix

$$G = (\Gamma(1), \Gamma(2), \ldots, \Gamma(n))^T \qquad (4.9)$$

with the $K\sum q_i$ matrix $\Gamma(n')$ given by

$$\Gamma(n')=\begin{bmatrix} \vartheta_{n'}(1) & & 0 \\ & \vartheta_{n'}(2) & \\ & \cdot & \\ & \cdot & \\ 0 & & \vartheta_{n'}(K) \end{bmatrix} \qquad n' = 1,2,\ldots,n \qquad (4.10)$$

where the $\sum q$ -dimensional row vector $\vartheta_{n'}(k)$ is

$$\vartheta_{n'}(k)= (r_{11n'}s_{1n'}, r_{12n'}s_{1k},\ldots, r_{1q_1n'}s_{1k};\ldots; r_{p1n'}s_{pk},\ldots, r_{pq_pn'}s_{pk}) \;,\; k = 1,\ldots,K \qquad n' = 1,\ldots,n$$

The $K\sum q_i$ -dimensional vector w is defined by

$$w = (h^T(1), h^T(2),\ldots, h^T(K))^T$$

where the $\sum q_i$-dimensional vector h(k) is

$$h(k) = (d_{11k},d_{12k},\ldots,d_{1q_1K};d_{21k},\ldots,d_{2q_2k};\ldots;d_{p1k}\ldots d_{pq_pk})^T$$

We can now write (4.1) as

$$\dot{z}(t) = Az(t) + Bb + r'(z) + E(u_o - Gw) \qquad (4.11)$$

Thus, (4.11) is our new airshed model equation, and w is our new control vector. Our objective will be to determine w over the interval (t_o,t_f).

To simplify the problem somewhat, we make the following two

assumptions:

(1) The matrices A, B, b and E are constant over the interval (t_o, t_f). In other words, once a set of meteorological measurements are made at t_o, conditions are assumed to be constant until t_f, when, presumably, a new set of measurements are made. As noted before, reasonable value of $t_f - t_o$ might be two hours.

(2) The control actions w are constant over (t_o, t_f). (Since control strategies will involve actions such as reducing freeway traffic or power plant operations, it is impractical to update the strategy too frequently, so this is an entirely reasonable requirement.)

We choose as the explcit from of the air quality constraint $g(z) \leqq g^*$,

$$\psi(z(t_f)) + \int_{t_o}^{t_f} \phi(z(t))\, dt \leqq 0$$

The two terms account for the instantaneous concentrations at the end of the control period t_f and dosages during the entire control period, respectively. Although we do not include explicitly a constraint on concentrations over the whole control-period, appropriate choice of ψ and ϕ will serve to keep concentrations below a desired level.

The individual control variables d_{ijk} must satisfy two constraints, namely that

(1) the number of source units controlled not exceed the total source units

$$\sum_{j=1}^{q_i} \omega_{ij}\, d_{ijk} \leqq s_{ik} \qquad \begin{matrix} i = 1,2,\ldots, p \\ k = 1,2,\ldots,K \end{matrix} \tag{4.13}$$

(2) the number of control units be non-negative

$$d_{ijk} \geq 0 \qquad (4.14)$$

Finally, we must specify the objective function to be minimized. We have stated that we desire to minimize the deviation between normal emission levels and those required to meet the air quality criteria. Perhaps a better choice would seem to be to minimize the total cost of control rather than simply the amount of reduction required. However, costs associated with a certain measure are often not easy to estimate. This is particularly true in the case of real-time controls, such as rerouting of traffic or providing only limited access to freeways. Consequently, we will not consider control costs as our objective function, although control costs are almost always closely tied to the level of control required, so that omission of explicit costs is not a serious drawback in the problem formulation.

A reasonable choice for the objective function is the quadratic form, $J = w^T Q w$, where Q is a pre-specified weighting matrix. If no control is applied J = 0, since, by definition, w = 0. Thus, we want to keep J as close to zero as possible.

In summary, the general problem to be solved is the following: Minimize J with respect to w, subject to the constraints, (4.11) and (4.12) - (4.14). In the next section we develop a computational method to solve this problem.

4.4 GENERAL METHOD OF SOLUTION

Since w is a set of constant parameters, the general problem is a mathematical programming problem with both nonlinear and differential

equation constraints. We now present a computational method for determining w which minimizes J subject to the constraints above. The method is based on iterative improvement of an initial guess $w^{(o)}$. We begin by linearizing (4.11) about a nominal control $w^{(o)}$ and the corresponding nominal trajectory $z^{(o)}$. The perturbation $\delta z = z(t) - z^{(o)}(t)$ is governed by

$$\delta\dot{z}(t) = (A + r_z(z^{(o)}(t)))\delta z - EG\delta w \tag{4.15}$$

$$\delta z(t_o) = 0 \tag{4.16}$$

where $\delta w = w - w^{(o)}$. We wish to choose the increment δw such that j is minimized and (4.12) to (4.14) are satisfied and

$$\psi(z^{(o)}(t_f)+\delta z(t_f)) + \int_{t_o}^{t_f} \phi(z^{(o)}(t) + \delta z(t))dt = 0 \tag{4.17}$$

$$Z\delta w = y - Z\, w^{(o)} \tag{4.18}$$

$$-\delta w = w^{(o)} \tag{4.19}$$

where y is the pK-dimensional column vector $(s_{11}, s_{21},\ldots, s_{p1}, s_{12}, s_{22}\ldots, s_{p2},\ldots, s_{pK})^T$ and the block-diagonal matrix

$$Z = \begin{bmatrix} W & & 0 \\ & W & \\ & & \ddots \\ 0 & & W \end{bmatrix} \tag{4.20}$$

where W is the $p \times \Sigma q_i$ matrix

$$W = \begin{bmatrix} \sigma(1) & & 0 \\ & \sigma(2) & \\ & \ddots & \\ 0 & & \sigma(p) \end{bmatrix} \tag{4.21}$$

and $\sigma(i)$ is the q -dimensional row vector $(\omega_{i1}, \omega_{i2}, \ldots, \omega_{iq_i})$.

In order to obtain (4.17) in a form amenable to computation, we linearize the constraint about $z^{(o)}(t)$,

$$\psi(z^{(o)}(t_f)) + \psi_z(z^{(o)}(t_f))\, \delta z(t_f) + \int_{t_o}^{t_f} \Big[\phi(z^{(o)}(t)) + \phi_z(z^{(o)}(t))\, \delta z(t) \Big] dt = 0 \tag{4.22}$$

The solution of (4.15), a set of linear differential equations with time-varying coefficients (i.e. $r_z(z_o)(t)$) can be expressed as

$$\delta z(t) = - \int_{t_o}^{t} \Phi(t,\eta)\, EG\, \delta w\, d\eta \tag{4.23}$$

where the nKxnK transition matrix $\Phi(t,\eta)$ satisfies

$$\frac{\partial \Phi(t,\eta)}{\partial t} = (A + r_z(z^{(o)}(t)))\, \Phi(t,\eta) \tag{4.24}$$

$$(\eta,\eta) = I \tag{4.25}$$

Using (4.23), (4.22) becomes

$$(\psi_z(z^{(o)}(t_f))\, FEG + VEG)\, \delta w = \psi(z^{(o)}(t_f)) + \int_{t_o}^{t_f} \phi(z^{(o)}(t))dt \tag{4.26}$$

where

$$F = \int_{t_o}^{t_f} \Phi(t_f, \nu)\, d\nu \tag{4.27}$$

and

$$V = \int_{t_o}^{t_f} \phi_z(z^{(o)}(t')) \int_{t_o}^{t'} \Phi(t', \nu)\, dt'\, d\nu \tag{4.28}$$

Thus, the problem of determining w can be stated as follows:

$$\underset{\delta w}{\text{Min}}\ J = \underset{\delta w}{\text{Min}}\ (\delta w^T Q\, w^{(o)}) \tag{4.29}$$

subject to (4.18), (4.19) and (4.26). Since J (the R.H.S. of (4.29)) and the constraints (4.18), (4.19) and (4.26) are linear in δw, this problem can be solved by linear programming.

We note that $\Phi(t,\nu)$ is not explicitly required in this problem, rather only the integral (4.27) given by F. By inspection we see that

$$\int_{t_o}^{t} \Phi(t,\nu)\, d\nu \cdot E = \frac{\partial z}{\partial w} \tag{4.30}$$

where $\partial z/\partial w$ is the $nKxnK$ matrix of sensitivity coefficients of the state z to the control w. Thus, Φ need not be evaluated explicitly. The sensitivity matrix can be computed by perturbing $w^{(o)}$ by $w^{(o)} + \epsilon$ and computing by finite differences

$$\frac{\partial z_i}{\partial w_j} = \frac{z_i(t) - z_i^{(o)}(t)}{\epsilon_j} \tag{4.31}$$

The computational scheme is as follows:

(1) Choose a nominal control $w^{(o)}$ and solve (4.11) to obtain the nominal state $z^{(o)}(t)$. A convenient initial guess $w^{(o)}$ is the uncontrolled emission level $w = 0$.

(2) Evaluate F and V by the procedure described in conjunction with (4.30) and (4.31). The perturbations ϵ are arbitrary and are not necessarily related to δw.

(3) Determine δw by minimizing (4.29) subject to (4.18), (4.19) and (4.26) by linear programming. Compute the next iterate of w by $w^{(1)} = w^{(o)} + \delta w$. Return to step 1 with $w^{(1)}$ in place of $w^{(o)}$.

(3) When $[J(w^{(m+1)}) - J(w^{(m)})]/J(w^{(m)})$ is less than a certain criterion, stop.

4.5 Real-time control of photochemical smog in the Los Angeles basin

As an application of the general theory we have developed, we will consider real-time control of air pollution in the Los Angeles basin. The results to be presented should be viewed only as preliminary with respect to a final control scheme for Los Angeles. The primary reason is that the well-mixed cell airshed model that we will employ here is rather crude, compared to a model currently being developed based on the continuity equations for mean concentrations of the pollutant species (Roth et al., 1971). Thus, the evaluation of the real-time control method is the principal aim of this section as

opposed to the presentation of a validated mathematical model for photochemical smog in Los Angeles. The latter study is forthcoming (Reynolds et al. 1973).

Figure (4.1) presents a map of the Los Angeles basin with a 2 mi. x 2 mi. grid overlaying a 50 mi. x 50 mi. area. The locations of major sources as well as the Los Angeles County Air Pollution Control District (APCD) pollutant monitoring sites are shown. The primary pollutants of most importance in Los Angeles are CO, NO_x, and hydrocarbons, with SO_2 and particulate matter of somewhat less importance. The most significant secondary pollutants (those formed in the atmosphere by chemical reaction) are NO_2 and O_3. It is well-established that the major sources of primary pollutants in Los Angeles are motor vehicles and power plants, with smaller contributions from refineries, industrial operations and aircraft (Lemke, 1971). Prevailing wind patterns are essentially the same in summer and winter, that is, from the west to the east.

The behavior of the various species varies with summer and winter conditions. CO distributions (due entirely to motor vehicles) are approximately the same all year, except that yearly morning winter concentrations are higher than summer. NO concentrations are highest in early morning in the vicinity of freeways and power plants. NO_2, which is not emitted in significant quantities from sources, is formed in the atmosphere by oxidation of NO, and subsequently converted to nitric acid and organic nitrates in the photochemical smog reactions. NO_2 concentrations are higher in the winter, when, because of shorter days and less intense sunlight, the photochemical reaction sequence does

not proceed to completion. In summer, on the other hand, the primary NO and hydrocarbons react to completion to yield large quantities of O_3, as the air parcels traverse the basin from west to east.

Our real-time control study will center on summer time conditions, since it is in the summer that the typical Los Angeles smog is felt to be most damaging, primarily because of the high ozone concentrations achieved. We will consider the following species: CO, NO, hydrocarbons (HC), NO_2 and O_3. The first three are primary pollutants, while the latter two are secondary pollutants. We neglect SO_2 and particulate matter because control measures in effect in Los Angeles have reduced these two pollutants to considerably lesser importance than CO, NO and hydrocarbons. For our exercise we have chosen a typical day in 1969, namely September 29, on which pollutants concentrations were reasonably high. What we will examine, therefore, is the effect that real-time controls would have had if they had been imposed on that day.

As the sources amenable to control we have selected for this study freeway motor vehicle traffic and power plant operations. We exclude surface street motor vehicle traffic because of the difficulty of its control. In each case, a partial reduction in the source activity is chosen as the control measure, i.e. reducing the number of vehicles allowed on freeways and reducing the amount of fuel burned (power delivered) in certain power plants. The control strategies will depend on the location of the sources as well as their uncontrolled hourly emission rates.

The question of the practicality of these control measures is a

central one. In view of the nature of the Los Angeles air pollution problem, real-time control actions must certainly focus on motor vehicle traffic and perhaps to a somewhat lesser extent on power plant operations. With respect to motor vehicle traffic, the question then is -- what is an effective means of reducing traffic and still providing the means for people to get to work? We will not attempt to answer this question here, although a system currently under study, involving appreciably expanded use of buses on special freeway lanes, is a realistic approach. Because of the size of and the freeway patterns in Los-Angeles, it is unlikely that, in the face of restricted freeway traffic, a large number of people would elect surface street routes as opposed to available mass transit.

We will require several items to be able to carry out the control exercise, namely:

(1) An emissions inventory for CO, NO and HC in the Los Angeles basin for 1969. This inventory will provide information on the location and hourly emissions from all major sources of these contaminants.

(2) A kinetic mechanism for the atmospheric chemical reactions involving CO, NO_2, HC and O_3. This will provide the functional form of $r'(z)$.

(3) Meteorological data, including hourly averaged wind speeds and directions and inversion depths, for the area to be modeled on September 29, 1969. Clearly, these elements are required for the validation of any mathematical model of urban air pollution. As we noted, our primary intent here is to test and examine the theory developed

for real-time control, rather than to formulate and validate an air pollution model. Consequently, we will not dwell too extensively on the comparison of the predictions of the well-mixed cell model with actual monitoring data.

4.5.1 Emissions inventory for the Los Angeles basin

The major sources of pollutant emissions in an airshed may be classified as moving and fixed. The predominant moving source in all urban airsheds is vehicular traffic, primarily automobiles and trucks, with smaller contributions coming from aircraft. Power plants, refineries and industrial operations are the principal fixed sources of pollutants in the Los Angeles basin.

There is a multiplicity of models for pollutant emissions that may be applied to individual sources and source types. The model that is used, and the degree of detail that is incorporated, is dependent upon the spatial and temporal resolution of the overall airshed model, the type and amount of data available, and the accuracy of those data. For example, in attempting to estimate contours of pollution concentrations over the Los Angeles basin during the course of a day and under particular meteorological conditions, it is necessary to compute the distribution over space of pollutant emissions from automobiles with a resolution of the order of one mile, and over time with a resolution of the order of one hour.

A motor vehicle emissions inventory can be divided into two parts:

(1) estimation of spatial and temporal distribution of traffic;

and (2) estimation of average vehicle emission rates applicable to traffic in the area. The spatial and temporal distribution of traffic on the freeways and surface streets in an urban area can be estimated from traffic counts which are normally taken by state and local agencies. Vehicle exhaust emissions rates are estimated from data collected in tests that simulate the emissions of vehicles actually driven over typical routes in the urban area being studied.

Data for the motor vehicle emissions inventory for Los Angeles for 1969 have been compiled by Roberts et al. (1971, 1972). In this study the spatial distribution of motor vehicle traffic was obtained from the traffic counts of freeways and major and minor street intersections and compiled for each of the 625 2 mi. x 2 mi. grid squares shown in Figure 4.1. Figure 4.2 shows the geographical distribution of freeway traffic in the Los Angeles basin in 1969 in thousands of vehicle miles/day, as determined by Roberts et al. (1971). A similar distribution, not presented here, was compiled for surface street traffic. The temporal distribution of both freeway and surface street traffic was determined by traffic count information and is shown in Figure 4.3. The freeway distribution was derived from 15-minute traffic count data over a 24-hour period at 31 freeway locations, while the surface street distribution was compiled from traffic counts on 52 randomly selected city streets.

The emission rates of CO, NO and HC were based on the computation of emissions for an "average" vehicle in 1969 based on the distribution of vehicle ages, makes and sizes in the Los Angeles basin and

on the federal driving cycle as representative of an average trip. Details of the computation are presented by Roberts et al. (1971, 1972). Resulting emission rates are given in Table 4.1.

There are 11 power plants in the Los Angeles basin, the locations of which are indicated in Figure 4.1. Data relating to locations, capacities and emissions are published annually by the Los Angeles Country APCD (e.g. Lemke, 1971). It was assumed that total daily power plant emissions are distributed equally over the 24-hour period. Average emission rates applicable to 1969 appear in Table 4.1. (In the computations, account was taken of the fact that each particular plant may very in its emission characteristics.)

Subsequently we shall consider two control cases: (1) CO control only, and (2) CO, NO and HC control. Because of the large dimensionality in the latter case ($n = 5$) when all five species are considered, computing time requirements force us to employ only a four cell-model. The four regions bounded by the heavy lines in Figure 4.1 constitute the four cells. The spatial distribution of major sources in the four cells is summarized in Table 4.2. In the case of CO control only, however, since only a single species is involved, it is possible to use a model with considerably more cells. Thus, we let $K = 20$ in the CO control case, as used previously by Kyan and Seinfeld (1972). We do not illustrate the 20 cells here.

Table 4.1 Average Emission Rates from Motor Vehicles and Power Plants in 1969

	CO	NO	HC
Motor Vehicles	63.9 grams/mi.	2.726 grams/mi*	7.66 grams/mi.
Power Plants	0	544.42 grams/megawatt hour**	0

* NO_2 emissions assumed to be 0.13 grams/mile.

** NO_2 emissions assumed to be 35.86 grams/megawatt hour.

- - - - - - - - - - -

Table 4.2 Spatial Distribution of Major Sources in the Los Angeles Basin in 1969 in the Four Cells as shown in Figure 4.1

Sources	Cell 1	2	3	4
Freeway motor vehicles (10^6 miles/day)	9.765	4.245	19.19	8.595
Surface motor vehicles (10^6 miles/day)	16.5	7.368	27.31	21.105
Power plants (megawatt)	4410	0	1069	3217

4.5.2 Kinetic mechanism for photochemical smog

The reaction term $r'(z)$ accounts for the rate of production of each species by chemical reaction in the atmosphere, and depends in general on the concentrations of each of the n species. There will be instances in which the use of an airshed model will be limited to the prediction of concentrations of inert species. However, when chemical reaction processes are of importance, it is essential to include an adequate description of these phenomena in the model.

A discussion of the development of kinetic mechanisms for photochemical smog suitable for inclusion in an airshed model would take us too far afield. Reviews of smog chemistry can be found in Altshuller and Bufalini (1971) and Johnston et al. (1970). A generalized kinetic mechanisms for photochemical smog that has been successful in simulating both laboratory and atmospheric data is that of Hecht and Seinfeld (1972). The mechanism, together with values of the kinetic parameters used in this study, is given in Table 4.3. Differential equations are required for NO, NO_2, O_3, and HC (the generalized hydrocarbon species), while the other species are assumed to be in a pseudosteady state. To conserve space, we do not present the explicit form of $r'(z)$ here. These may be found in Seinfeld et al. (1971).

Table 4.3 Kinetic Mechanism for Photochemical Smog*

	Reaction	Rate constants employed
1.	$NO_2 + h \longrightarrow NO + O$	22.2 hr^{-1} #
2.	$O + O_2 + M \longrightarrow O_3 + M$	1.65×10^8 hr^{-1} (pseudo first order)
3.	$O_3 + NO \longrightarrow NO_2 + O_2$	1.308×10^3 ppm^{-1} hr^{-1}
4.	$O_3 + 2NO_2 \xrightarrow{H_2O} 2HNO_3$	0.36 ppm^{-1} hr^{-1} **
5.	$NO + NO_2 \xrightarrow{H_2O} 2HNO_2$	0.24 ppm^{-1} hr^{-1}
6.	$HO_2\cdot + NO_2 \longrightarrow HNO_2 + O_2$	600 ppm^{-1} hr^{-1}
7.	$HNO_2 + h \longrightarrow OH\cdot + NO$	0.3 hr^{-1} #
8.	$CO + OH\cdot \xrightarrow{O_2} HO_2\cdot + CO_2$	1.2×10^4 ppm^{-1} hr^{-1}
9.	$HC + O \longrightarrow \alpha RO_2\cdot$	1.2×10^5 ppm^{-1} hr^{-1}, $\alpha = 2.7$
10.	$HC + O_3 \longrightarrow \beta RO_2\cdot + RCHO$,	0.06 ppm^{-1} hr^{-1}, $\beta = 0.5$
11.	$HC + OH\cdot \longrightarrow \delta RO_2\cdot$	3.6×10^5 ppm^{-1} hr^{-1}, $\delta = 1.2$
12.	$RO_2\cdot + NO \longrightarrow NO_2 + \epsilon OH\cdot$,	1.08×10^5 ppm^{-1} hr^{-1}, $\epsilon = 0.6$
13.	$RO_2\cdot + NO_2 \longrightarrow PAN$	600 ppm^{-1} hr^{-1}
14.	$HO_2\cdot + NO \longrightarrow NO_2 + OH\cdot$	1.08×10^5 ppm^{-1} hr^{-1}

* reference - Hecht and Seinfeld (1972)

** reaction 4 is a composite of the three reactions:

$$O_3 + NO_2 \longrightarrow NO_3 + O_2, \quad NO_3 + NO_2 \longrightarrow N_2O_5, \quad N_2O_5 + H_2O \longrightarrow 2HNO_3$$

\# k_1 and k_7, the rate constants for reactions 1 and 7 respectively, depend on light intensity and are related to time in Los Angeles by

$$k_i(t)/k_{imax} = 1.017 - 0.06846((t-12)/6) - 1.0764((t-12)/6)^2, \quad i = 1,7$$

where t is the time in hours (t= 12 is noon) and k_{imax} is the value of rate contants given in the above table. (Reference-Reynolds 1972)

4.5.3 Meteorological data

The required meteorological data for implementation of the model are the intercell flow rates q_{jk} and the cell volumes v_k as a function of time. The intercell flow rates can be obtained from wind fields, whereas inversion heights are necessary to compute v_k. Wind speeds and directions as a function of time and location have been prepared for Roth et al. (1971b), based on hourly-averaged surface wind data at 34 stations inthe Los Angeles basin on September 29, 1969. Inversion heights were estimated based on measured vertical temperature profiles at three stations on the same day. Contours of constant inversion height were constructed on hourly maps by assuming the contours to be roughly parallel to the coastline. Inversion heights for the entire basin were then interpolated from the three contours.

The wind data in Roth et al. (1971b) were resolved into east-west and north-south components and then appropriately summed and averaged to produce the hourly q_{jk} required for the four cells. Inversion height data were used to give the cell volume v_k. Table 4.4 presents a typical set of this data for 11 a.m.

We have noted that our object here is not to present a validated mathematical model of Los Angeles air pollution and that such a study is forthcoming (Reynolds et al. 1973). Nevertheless, it is useful to have some idea of the validity of the cell model employed here. Since the cells are so large, particularly in the case of K = 4, it is not particularly revealing to compare the readings at one station to the average values in a 100 square mile region in which the station is located. However, as an indication of the concentration levels and temporal

trends of both the actual data and the model, we present Figures 4.4 and 4.5. Figure 4.4 shows the NO concentration simulated for cell 4 and the measured values at two stations in cell 4, Long Beach and Lennox stations. Figure 4.5 shows a similar comparison for O_3 in cell 3 and the measurements at the Reseda station in cell 3. Also shown in Figure 4.5 is the average O_3 concentration at all stations in cell 3.

Table 4.4 Intercell Flows q_{jk} and Cell Volumes v_k at 11 a.m. for the Four Cells Shown in Figure 4.1.

	$q_{jk} \times 10^{-10}$ meter^3/hour				
j	0	1	2	3	4
k					
0		1.193	0	1.024	3.922
1	2.199	0	0.684	0	0
2	1.459	0	0	0	0
3	2.052	0	0.775	0	0.465
4	0.429	1.69	0	2.268	0
v_k, $m^3 \times 10^{-10}$		4.053	2.573	7.013	4.736

4.5.4 Control parameters

The source types we consider amendable to control are freeway motor vehicle traffic and fuel consumption in power plants (p = 2). For each of the two source types the sole control method is a reduction in the source activity ($q_1 = q_2 = 1$). Thus, (4.7), (4.13) and (4.14) reduce to

$$\sum_{i=1}^{2} r_{iln'} d_{ilk} s_{ik} = e_{n'k_o} - e_{n'k}, \quad n' = 1,2,\ldots,5 \qquad k = 1,2,3,4 \tag{4.32}$$

and

$$0 \leq d_{ik} \leq 1 \qquad i = 1,2, \qquad k = 1,2,3,4 \tag{4.33}$$

where the control parameters are defined as follows:

$r_{11n'}$ = reduction in emissions of species n' (grams) per motor vehicle mile reduction below normal level. See Table 4.1.

$r_{21n'}$ = reduction in emissions of species n' (grams) per megawatt reduction in power plant output below normal level (equal to zero for all pollutants but NO_2). See Table 4.1.

d_{11k} = fractional reduction in freeway mileage in cell k during time period (t_o, t_f).

d_{21k} = fractional reduction in megawatt output in cell k during time period (t_o, t_f).

s_{1k} = normal vehicle miles travelled on freeways in cell k during (t_o, t_f). See Table 4.2.

s_{2k} = normal megawatts delivered in cell k during (t_o, t_f). See Table 4.2.

4.5.5 Control results

Figure 4.6 shows the results of implementation of the real-time control strategy for only CO in cells 13 and 20. The air quality constraint employed is $z_i(t_f) = 12$ ppm, $i = 1,2,\ldots,20$, that is, that at the end of the control period (one hour) that the CO concentration in any cell not exceed 12 ppm. Only values of concentrations at the end of each hour were assumed to be reported, and these values are connected by straight lines in Figure 4.6 and subsequent figures. Table 4.6 shows the temporal and spatial reductions in freeway traffic needed to achieve the results shown in Figure 4.7. Major reductions in freeway traffic are called for during the period 6-8 a.m. in central Los Angeles and the San Fernando Valley.

Figure 4.7 presents a comparison of ozone concentrations in cells 1 and 3 with and without real-time control. The air quality constraints employed were that the ozone concentrations at the end of the control period (5 hours) not exceed 0.29 and 0.13 ppm in cells 1 and 3, respectively. A control period of 5 hours (5-10 a.m.) was chosen in the NO_x, HC, O_3 case because the secondary pollutant O_3 resulting from early morning emissions attains its peak values several hours after the early morning rush hour. Therefore, it is necessary to choose a control period long enough to see the effect of the

control of early morning emissions. Different air quality constraints were chosen primarily to illustrate the flexibility of the theory.

Table 4.6 shows the spatial distribution of reductions in both freeway traffic and power plant operations during the period 5-10 a.m.

Table 4.5 Control Policy for the CO Case, expressed as Fractional Reduction in Freeway Motor Vehicle Traffic.

Cell	Hour 5-6 a.m.	6-7 a.m.	7-8 a.m.	8-9 a.m.
1				
2				
3		0.56		
4	0.11			
5	0.56		0.31	
6				
7				
8	0.06			
9	0.08	1.0	0.29	
10		0.75		
11				
12		0.75	0.78	
13		0.74	0.79	0.01
14		0.75		
15		0.71		
16		0.69		
17		0.47		
18				
19	0.54			
20	0.57	0.16		

Table 4.6 Control Policy for Photochemical Case expressed as Fractional Reduction of Freeway Traffic and Power Plant Output

Hour	Cell	1	2	3	4
5-10 a.m.	Freeway traffic	0.67	0	0.45	0.77
	Power plants	0.57	0	1.0	0

4.6 Discussion

The objective of this study has been the formulation and testing of a framework for considering real-time air pollution control strategies. The key aspects are the proper definition of the real-time control problem with respect to an airshed model and its solution rather than the advocacy of any particular model. Since, as we have noted, the well-mixed cell airshed model will probably not be the form employed in the future for large urban simulations, but rather one based on the partial differential equations for the mean concentrations of pollutants (the so-called semi-empirical equation of atmospheric diffusion - Monin and Yaglom, 1971), an important question to which we must address ourselves here is - can the theory we have presented be implemented with feasible computing requirements on the more complex models to come?

When dealing with the control of essentially inert pollutants, such as CO, particulate matter and SO_2 (which for control purposes can be considered inert), we feel the answer is yes. The computing storage and time requirements for the 20 cell CO control exercise for Los Angeles were quite modest (70,000 bytes of storage, 40 seconds for 8-hour control on an IBM 370/155). Both the storage and time requirements would increase proportionately with the number of grid squares used in the model for single pollutant control.

Although we have illustrated the theory in a case of chemically reacting air pollution in order to show its application under the most general circumstances, at this time it appears that extension

to a model consisting of coupled, three-dimensional partial differential equation, while theoretically feasible, is not practical given current computing capabilities. The theoretical feasibility is clear since the airshed simulation model comes into the picture through (4.15) and (4.31) only which are amenable to any simulation model. For example, the 4-cell Los Angeles exercise involving 5 species required storage of 220,000 bytes and 10 min. of computing time for the 8-hour control results in Table 5. Extension to 5 coupled partial differential equations on a 25 x 25 x 10 mesh with $t = 5$ minutes, as reported by Reynolds et al. (1973) would require more than one hour of computing for a comparable control exercise. As with similar problems, such as global weather simulation, mathematical modeling of chemically reacting urban air pollution will require considerable computing capacities, thereby making control exercises an expensive (but necessary) undertaking.

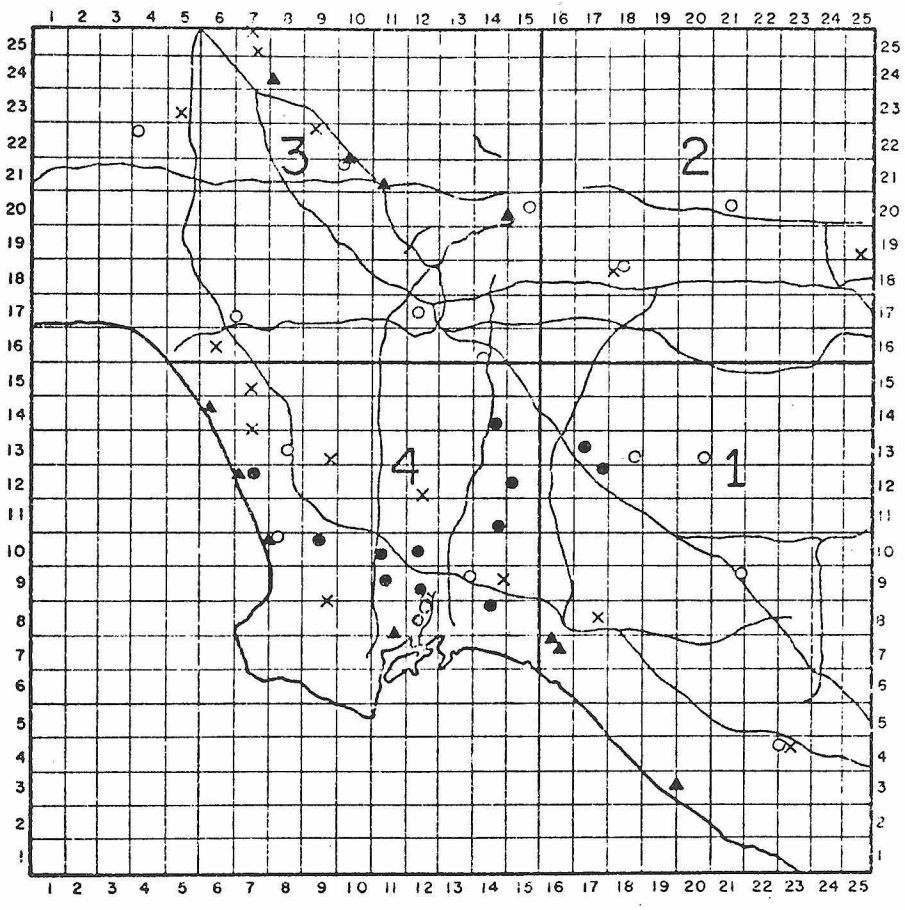


Figure 4.1 Los Angeles basin divided into 4 cells showing locations of major sources.

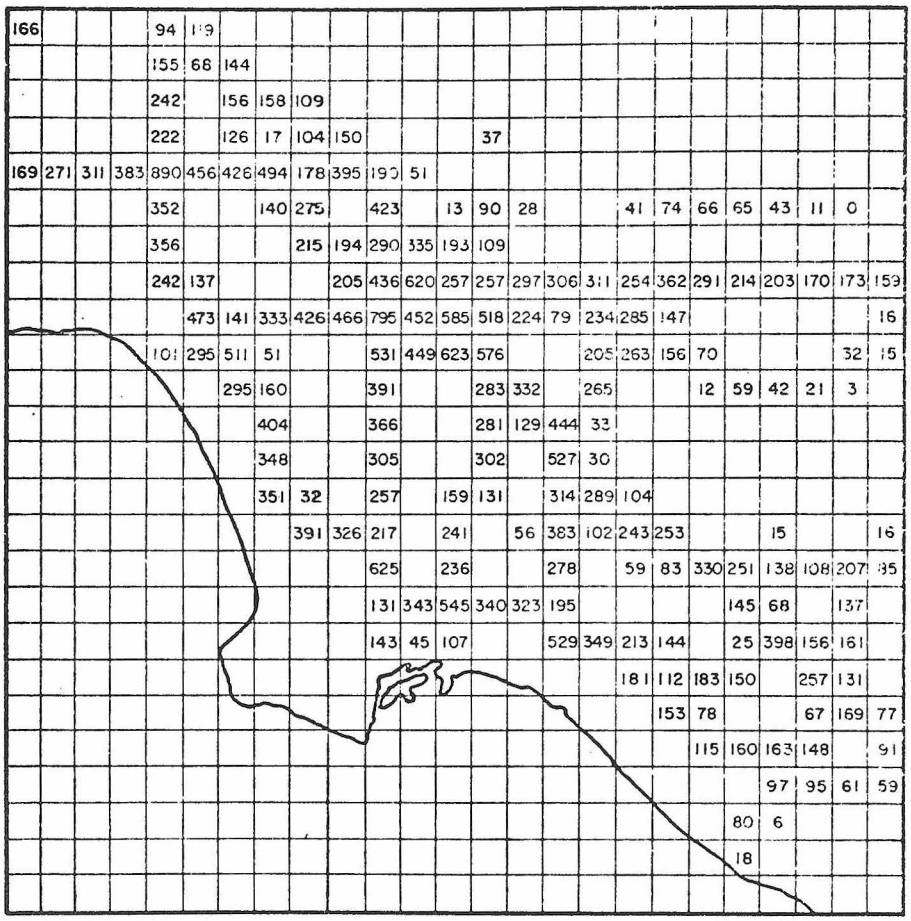


Figure 4.2 Spatial distribution of freeway traffic in the Los Angeles basin in 1000 vehicle miles/day.

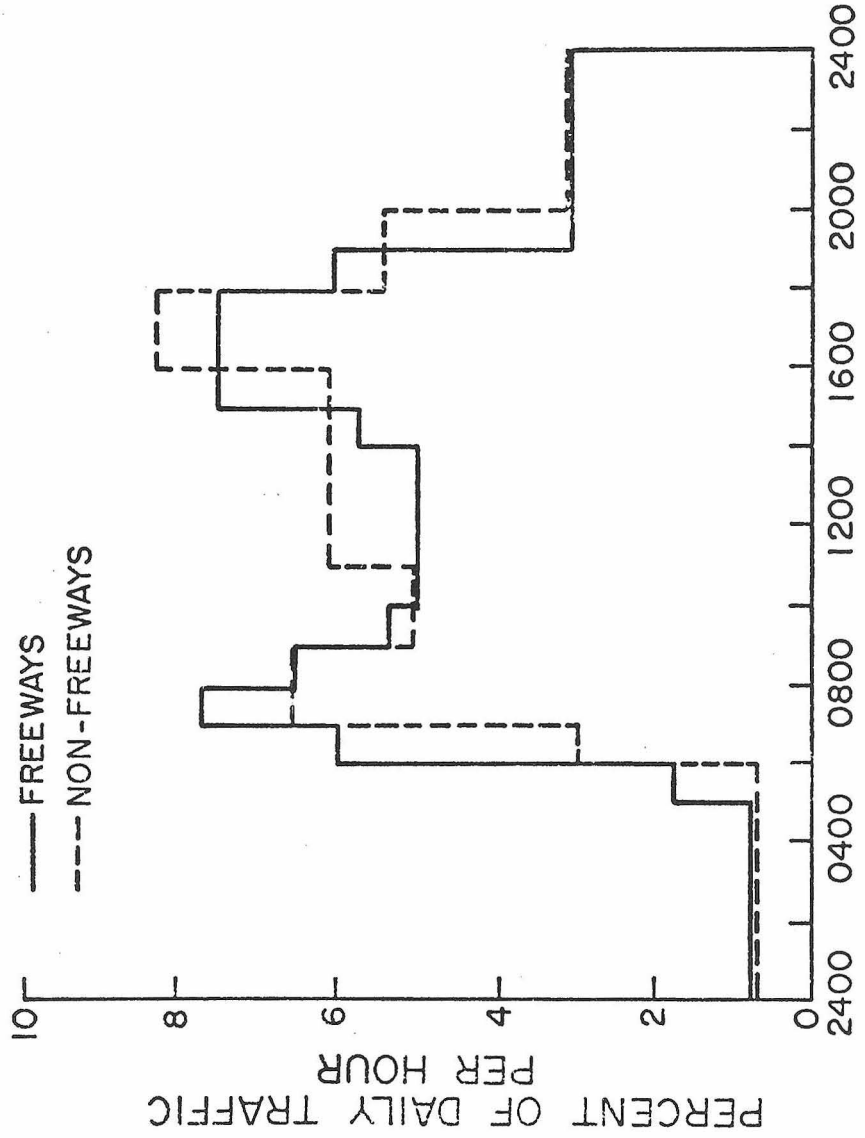


Figure 4.3 Temporal distribution of vehicular traffic in the Los Angeles basin.

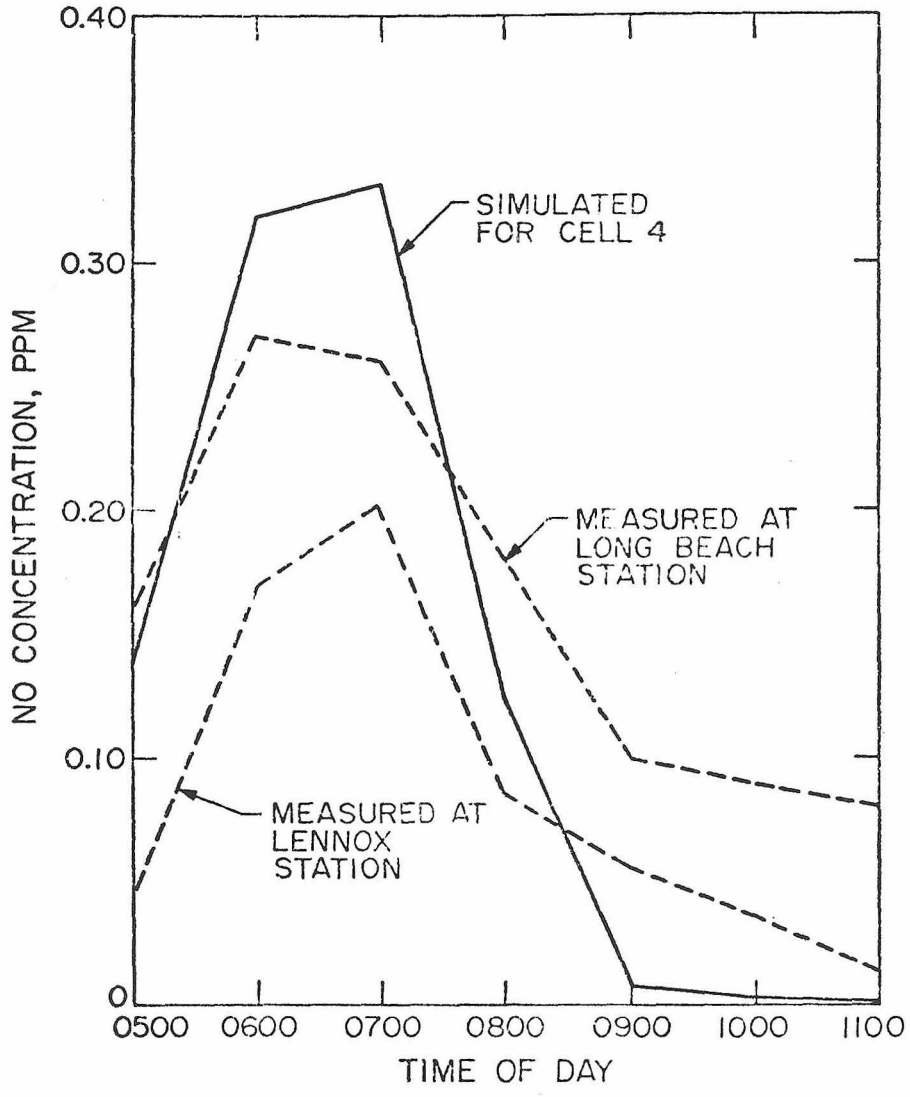


Figure 4.4 Comparison of measured and simulated NO concentrations for a 4-cell simulation of the Los Angeles basin.

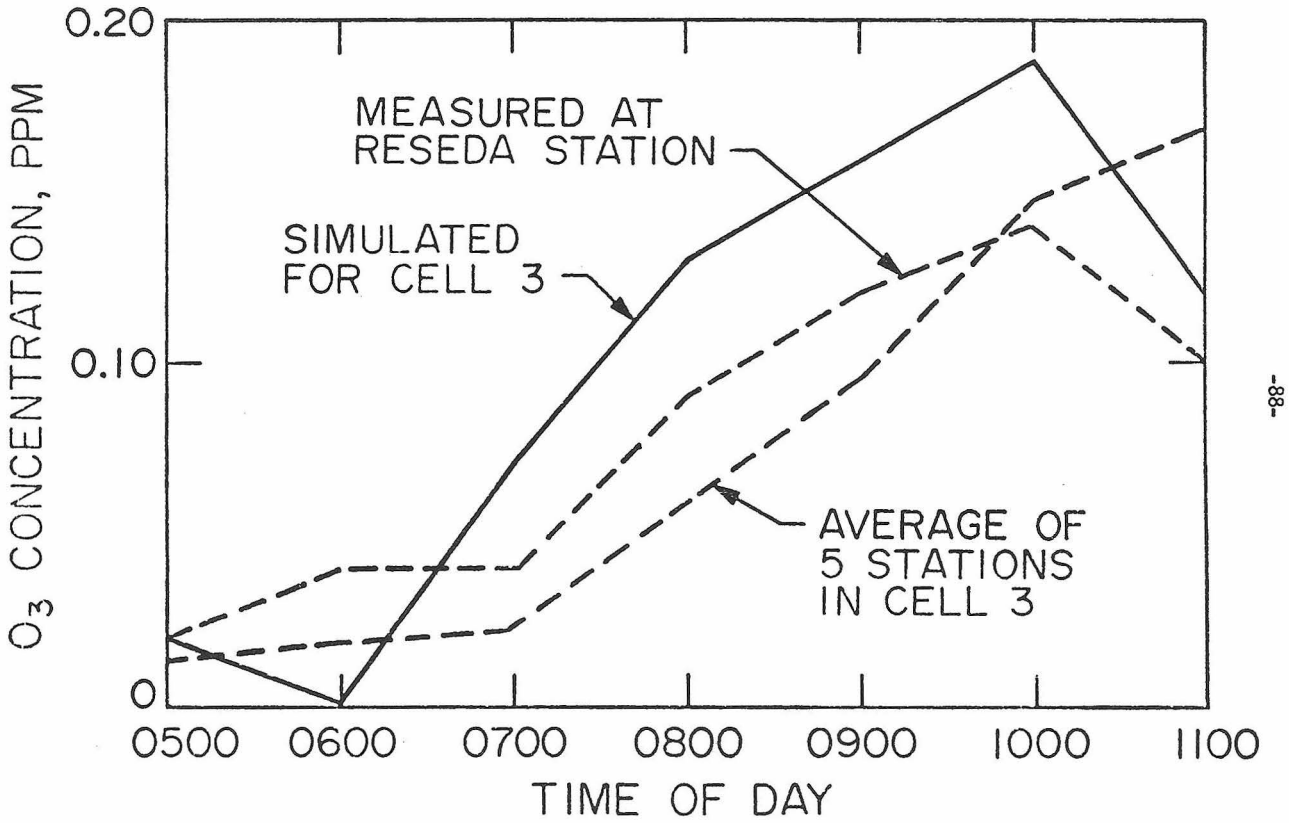


Figure 4.5 Comparison of measured and simulated O_3 concentrations for a 4-cell simulation of the Los Angeles basin.

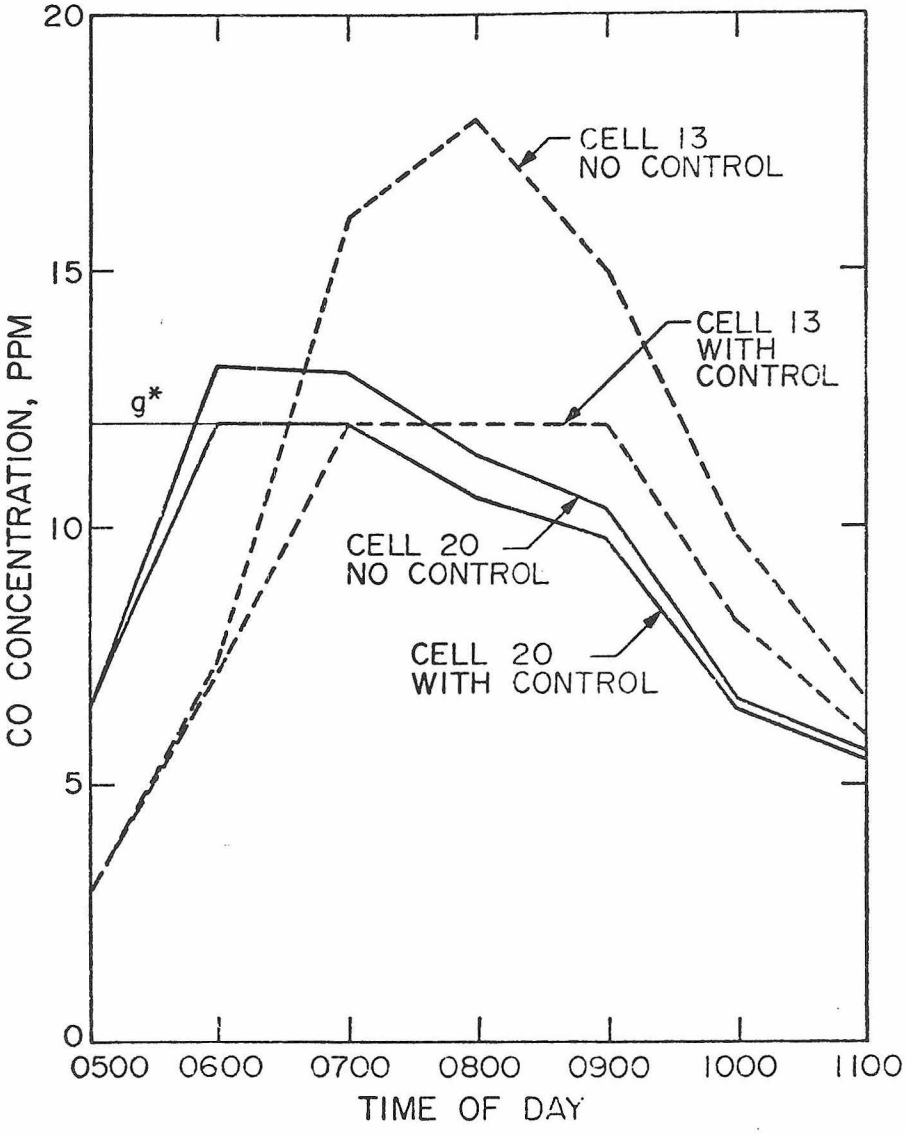


Figure 4.6 Results of CO control with Los Angeles basin divided into 4 cells.

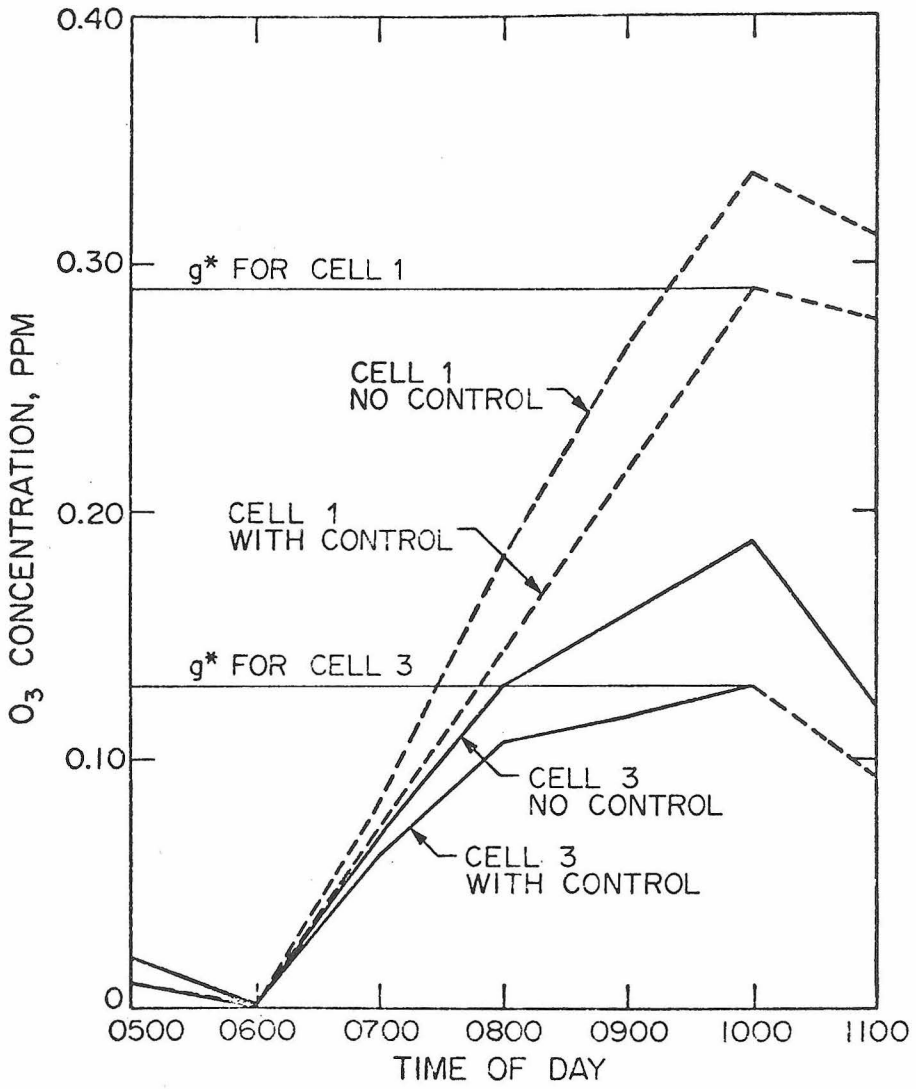


Figure 4.7 Results of O_3 control with the Los Angeles basin divided into four cells.

CHAPTER 5. EVALUATION OF LONG-TERM AIR POLLUTION CONTROL STRATEGIES FOR THE LOS ANGELES BASIN

To illustrate the theory of long-term control and to obtain some practical results, the evaluation of long-term air pollution control strategies for the Los Angeles basin is carried out for CO control and also for control of hydrocarbons and oxides of nitrogen.

5.1 Control of CO in Los Angeles

To illustrate the theory and explain in detail the mechanics of the discrete (backward) dynamic programming solution, we will consider the control of CO in the Los Angeles basin for 1972-1974. Carbon monoxide is essentially an inert pollutant the source of which is almost entirely motor vehicles. Thus we will consider the only source as motor vehicles.

The primary aim of this CO control exercise is to illustrate the use of the dynamic programming solution of the multi-year air pollution control problem. Although the example considered is CO control for the Los Angeles basin, and an effort has been made to employ realistic emissions and cost data, the exercise is artificial in the sense that CO control is not the critical problem for Los Angeles. Control of hydrocarbons and oxides of nitrogen emissions is a far more important issue. In addition, control of hydrocarbons and

NO_x cannot generally be accomplished without simultaneous control of carbon monoxide, so that the obvious way to approach long-term air pollution control in Los Angeles is to focus on hydrocarbons and NO_x, allowing carbon monoxide emissions to settle at that level which results from control of the two former classes. Section 5.2 is devoted to such an exercise.

We will employ the well-mixed cell model as detailed in chapter 2, i.e.

$$v_k(dz_{ik}/dt) = - z_{ik}(dv_k/dt) + \sum_{j=0}^{K} q_{jk} z_{ij} - z_{ik} \sum_{j=0}^{K} q_{kj}$$

$$+ s'_{ik} - d'_{ik} + r'_{ik} \qquad (5.1)$$

$$z^o_{ik}(o) = z^o_{ik} \qquad (5.2)$$

to describe the daily dynamic behavior of carbon monoxide in the Los Angeles basin.

We divide the Los Angeles basin into 20 cells (K = 20) as shown in Figure 5.1. Thus, (5.1) is a set of twenty coupled ordinary differential equations. The areas of the various cells are given in Table 5.1. For typical daily meteorological conditions we take those of September

Table 5.1 Cell areas, CO concentrations at 5 a.m. and fraction of total vehicle mileage for the 20 cells comprising the Los Angeles basin

Cell	Area, mile^2	CO conc. ppm	$m_k \times 10^2$
1	29.6	5	0.74
2	93.8	5	2.56
3	100	7	5.59
4	100	7	1.49
5	100	10	3.07
6	100	10	0.02
7	100	11	0.47
8	100	12	1.58
9	100	13	6.60
10	100	10	6.15
11	83.2	5	2.25
12	100	5	10.1
13	100	3	13.04
14	100	5	6.36
15	100	7	6.08
16	99	7	5.98
17	77.2	7	6.23
18	64	6	4.96
19	84	6	7.58
20	100	6.5	9.17

Table 5.2 Hourly source activity distribution for motor vehicles in the Los Angeles basin

Hour PST	Hourly source activity
5	0.0178
6	0.0591
7	0.0768
8	0.0648
9	0.0536
10	0.0484
11	0.0484
12	0.0484
13	0.0484
14	0.0569
15	0.0746
16	0.0746
17	0.0746

29, 1969, a representative smoggy day in the autumn. The concentration of CO outside the airshed is assumed to be 5 parts per million by volume (ppm) and those in the airshed at 5 a.m. are given in Table 5.1. The complete hour-by-hour intercell wind velocities and cell inversion heights will not be presented here. These were developed from the data of Roth, et al. (1971).

The airshed is described by (5.1) with $n = 1$ and $K = 20$. The deposition and reaction terms, d' and r' are zero. The source term s_k' (the subscript for pollutant is omitted since only one pollutant is involved) is

$$s_k' = (871\ \beta(t) m_k / v_k) \sum_1^p E_i s_i \qquad (5.3)$$

where 871 is a conversion factor, $\beta(t)$ is the hourly traffic distribution function, and m_k is the fraction of total vehicle miles/day traveled in cell k. m_k and $\beta(t)$ are given in Tables 5.1 and 5.2 respectively. These values were determined for 1969 and can be expected to remain fairly constant with time. The cell volumes v_k are in cubic meters in (5.3). Thus the airshed model consists of 20 coupled linear ordinary differential equations. With the source inputs specified by (5.3), these equations can be solved to give $z_k(t)$, $k = 1,..,20$, by use of the fundamental matrix. We will not detail the solution method here, as it is quite standard.

5.1.1 Definition of the problem

The problem we wish to consider is the determination of CO control strategies for 1972-1974 for the Los Angeles basin corresponding

to different prescribed levels of air quality. As noted, CO is emitted almost entirely from motor vehicles. For the purposes of control we will classify the motor vehicles in the airshed in 1972-1974 in two groups: (i) 1965 and earlier models, and (ii) 1966-1969 models. Thus p = 2 for the control problem. 1970 and later models have CO control and will not be considered as accessible to control, although this group does constitute a source of CO which must be included in the total source emissions computed from (5.3). It is the very slow rate of disappearance of used cars which leads to the necessity to control used cars.

Table 5.3 presents the projected source information for 1972-1974 for Los Angeles (Trijonis, 1972). Used in developing the estimates in Table 5.3 were an age distribution of vehicles in any given year and the estimated number of vehicle miles/day traveled by cars of different ages. Additional detail on these estimated source strengths, are given by Trijonis (1972).

Table 5.4 presents the control methods that we will consider for the 1969 and older used cars as given by Downing et al.(1970). Each of the four methods is a device which can be installed on a used car. Only one device can be used per car, however. The reductions are given in Table 5.4 as fractional reductions π_{ij}, that is, the fraction of the uncontrolled CO emission eliminated with the device. The problem, then, is to find the optimal allocation of these four devices among the used car population in 1972-1974.

The costs shown in Table 5.4 are given in $/day. The manner

Table 5.3 Cabon monoxide source emission projections for Los Angeles

(1972-1974)

Year		p = 1 1965 and earlier models	p = 2 1966-1969 models	1970 and later models
1972	(a)	1.978	1.4835	0.8385
	(b)	2.3585	3.7689	3.2336
1973	(a)	1.672	1.4835	1.2445
	(b)	1.8854	3.1773	4.5020
1974	(a)	1.3725	1.4835	1.6440
	(b)	1.4556	2.7282	5.5457

(a)—Total number of motor vehicles projected ($X10^{-6}$)

(b)—Total daily mileage traveled ($X10^{-7}$)

NOTE: Uncontrolled emission levels of CO in grams/mile for the three classes of vehicles are assumed to be:

1965 and earlier	80
1966-1969	34
1970-	23

Source of data in Table—Trijonis (1972)

Table 5.4 Control methods and costs (Downing, 1970)

s_1 = number of thousand vehicle miles/day traveled by 1965 and earlier motor vehicles

s_2 = number of thousand vehicle miles/day traveled by 1966-1969 motor vehicles

d_{12}, d_{22} = number of catalytic reactors/1000 vehicle miles of source 1 and 2.

d_{11}, d_{21} = number of flame afterburners/1000 vehicle miles of source 1 and 2.

d_{13} = number of smog package tuneups/1000 vehicle miles of source 1

d_{14} = number of spark advance systems/1000 vehicle miles of source 1

(NOTE—d_{13} and d_{14} applicable only to 1965 and earlier motor vehicles)

π_{ij} = fraction of uncontrolled CO emissions reduced by control method j on source i

	π_{ij}			
ij	1	2	3	4
1	0.97	0.95	0.15	0.50
2	0.97	0.95	N/A	N/A

C_{ijt} = $/day for control method j for source i in year t

	1972	1973	1974
$C_{11} = C_{21}$	0.2948	0.3895	0.6466
$C_{12} = C_{22}$	0.2347	0.2809	0.4197
C_{13}	0.0340	0.0350	0.0360
C_{14}	0.1258	0.1424	0.1922

of determination of these costs is of some importance in the final results. There are basically two alternatives in assigning the cost of a control method. The first is to ascribe the complete cost of the device to the year in which it was purchased. The second is to amortize the cost in some manner over some subsequent time period, perhaps the life of the device. The method of assigning the costs will have important ramifications in the optimal policy over a t-year period. Consider the situation, for example, in which control costs are amortized over, say, 10 years and the period over which legislation is to be enacted is T= 3 years. For our cost function, we only take into account those costs incurred over the three years. Assuming the air quality criterion can be met each year, certainly the least cost policy would be to install all the controls in the third year, thereby incurring only one-tenth of the total costs in the three-year period. Clearly we need a way to account properly for the total cost of control in the T-year period. We have chosen to determine the daily costs C_{ijt} in the following way. For those controls installed in year 1 (1972), their cost is allocated equally over the entire three years; for those in year 2 (1973), equally over the last two years; and for those in year 3 (1974), entirely in that year. Thus the total control costs are confined to the three-year period. (equivalently, set $\mu(1) = 1$, $\mu(2)= 2$, $\mu(3) = 3$)

As the air quality criterion we will select the 4-hour average CO concentration from 5 a.m. to 9 a.m. in cell 20. Cell 20 includes south-central Los Angeles, where CO concentrations are traditionally among the highest in the basin. In addition, the period of highest CO

concentrations in this area are experienced in the morning. Thus,

$$g(z) = \frac{1}{T}\int_0^T z_{20}(t)dt \qquad (5.4)$$

where T = 4 hours. All of the elements of the problem are now available to solve the optimal control problem.

5.1.2 Solution of the problem

The relationship between a set reduction Δx in mass emissions of CO in any of the three years and the minimum cost set of controls to achieve this reduction is easily obtained by linear programming, using the data in Table 5.3 and 5.4. To insure that at most one device be prescribed per vehicle, we need the following constraints in the linear programming solution,

$$d_{ijt} \leq d_{it}^* \qquad (5.5)$$

where $q_{11} = 4 \quad q_{12} = 2 \quad q_{13} = 2$

$q_{21} = 2 \quad q_{22} = 2 \quad q_{23} = 2$

and $d_{11}^* = 84 \quad d_{12}^* = 87 \quad d_{13}^* = 94$

$d_{21}^* = 39 \quad d_{22}^* = 47 \quad d_{23}^* = 54$

This constraint is necessary because the source unit used is 1000 vehicle miles. Thus d_{11}^* indicates that, in 1972, 1000 miles/day traveled by 1965 and earlier vehicles is equivalent to 84 vehicles, and that the sum of devices/1000 vehicle miles cannot exceed 84.

(5.5) is just one form of the limited supply constraint. The source constraint is also satisfied once (5.5) is satisfied.

Since n = 1 and p = 2 the state vector is $E_t = (E_{1t}, E_{2t})$ and the control variable is $\Delta x_t, t = 1,2,3$. At year 3, in principle, Δx_t for each E_{12} and E_{22}, the state variables "leaving" year 2 and "entering" year 3, we would be required to determine $f_1(E_{12}, E_{22})$ minimizing $C_3(E_{12}, E_{22}, \Delta x_3)$ by choice of Δx_3, in (3.25). However, for each E_{12} and E_{22}, if we set $g(z) = g_3^*$, which will be the least cost policy (i.e., reduce only to the standard but not below it), Δx_3 will be automatically determined. If g_3^* can be met with no reduction necessary in year 3, the $\Delta x_3 = 0$. Then, given Δx_3, E_{12}, and E_{22}, the minimum cost and corresponding control allocation can be determined by linear programming and stored in a table. We denote this minimum cost by $C_3^*(E_{12}, E_{22})$. Clearly, C_3^* only includes the cost of controls instituted in year 3.

Proceeding backward to year 2, we want to determine

$$f_2(E_{11}, E_{21}) = \min_{\Delta x_2}\left\{ C_2(E_{11}, E_{21}, \Delta x_2) + f_1(E_{12}, E_{22})\right\}$$

$$+ B_2(E_{10}, E_{20}, E_{11}, E_{21}, \Delta x_1) = \min_{\Delta x_2}\left\{ C_2(E_{11}, E_{21}, \Delta x_2)\right.$$

$$\left. + C_3^*(E_{12}, E_{22}) + B_3(E_{10}, E_{20}, E_{11}, E_{21}, \Delta x_1, \Delta x_2)\right\}$$

$$+ B_3(E_{10}, E_{20}, \Delta x_1) \qquad (5.6)$$

We note that B_3 is that portion of the cost of controls for

years 1 and 2 that is to be paid for in year 3. B_3 can be decomposed into

$$B_3(E_{10}, E_{20}, E_{11}, E_{21}, \Delta x_1, \Delta x_2) = B_2(E_{10}, E_{20}, \Delta x_1)$$

$$+ H_3(E_{11}, E_{21}, \Delta x_2) \qquad (5.7)$$

where $B_2(E_{10}, E_{20}, \Delta x_2)$ is the contribution of the cost of controls instituted in year 1 to the total cost in year 3, and $H_3(E_{11}, E_{21}, \Delta x_2)$ is the contribution of the cost of controls instituted in year 2 to the total cost in year 3. Since we have assumed that the cost of controls instituted in year 1 is to be evenly allocated over the 3-year period, $B_2(E_{10}, E_{20}, \Delta x_1)$ is the contribution from year 1 controls to the costs of year 2 as well as year 3. We note also that $H_3(E_{11}, E_{21}, \Delta x_2)$ is simply equal to $C_2(E_{11}, E_{21}, \Delta x_2)$. Thus, (5.7) becomes

$$f_2(E_{11}, E_{21}) = \min_{\Delta x_2} \left\{ 2C_2(E_{11}, E_{21}, \Delta x_2) + C_3^*(E_{12}, E_{22}) \right\}$$

$$+2B_2(E_{10}, E_{20}, \Delta x_1) \qquad (5.8)$$

Corresponding to an assumed pair (E_{11}, E_{21}) we perform a minimization over Δx_2. We then store the control and cost values in the year-2 table. The procedure for year 1 is identical to year 2, except that only one pair (E_{10}, E_{20}) need be considered. These are, of course, the uncontrolled levels of emission in year 1. For those used cars that leave the system by attrition, we assume that a device on the car can be salvaged such that the car owner would no longer have to pay for the device in future years.

5.1.3 Discussion of results

Given a particular set of original emission levels in year 1, E_{10} and E_{20}, and the three air quality levels, g_1^*, g_2^*, g_3^*, the results obtained are: (a) the optimal reduction in source emissions over the three years, (b) the optimal allocation of controls over the three years, (c) the control cost associated with the three-year policy, and (d) the actual air quality achieved each year as a result of the control policy. To conserve space we will not present the dynamic programming tables developed in the solution.

Figure 5.2 shows the total three-year cost in dollars/day for $E_{10} = 8 \times 10^4$ and $E_{20} = 3.4 \times 10^4$ grams CO/1000 vehicle miles, $g_1^* = 9$ ppm, and various values of g_2^* and g_3^*. We see that it is relatively inexpensive to achieve $g_3^* = 6$ ppm as compared to 5ppm. It turns out that 5 ppm is about the minimum level of g_3 that can be achieved in 1974, since there will be a large proportion of 1970 and later model cars which are not being controlled beyond their original emissions. The cost of control increased rapidly with increase in either g_2^* or g_3^*. For $g_1^* = 9$ no control is required in 1972.

Table 5.5 shows the optimal three-year strategy for $g_1^* = 8$, $g_2^* = 6$, and $g_3^* = 5.5$. In this case, control is required in all three years. We note that the optimal policy involves only flame afterburners. Because the 1966-1969 cars account for more mileage individually than those of 1965 and earlier, the optimal policy is to control them sooner and more heavily than the 1965 and earlier cars. This is perhaps counter to our intuition, which would be to control the oldest

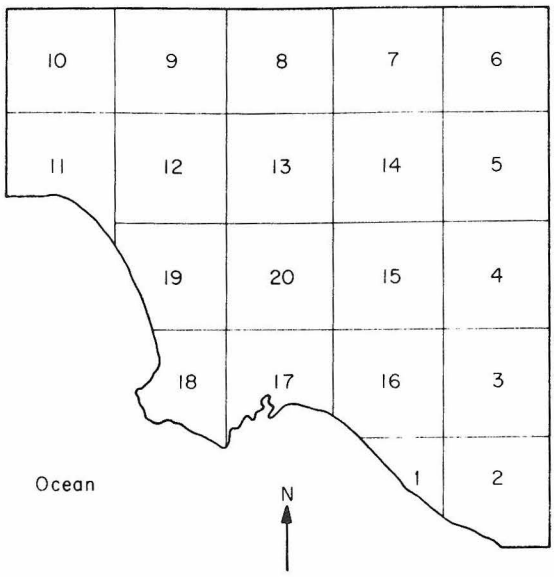


Figure 5.1 The Los Angeles basin divided into 20 cells

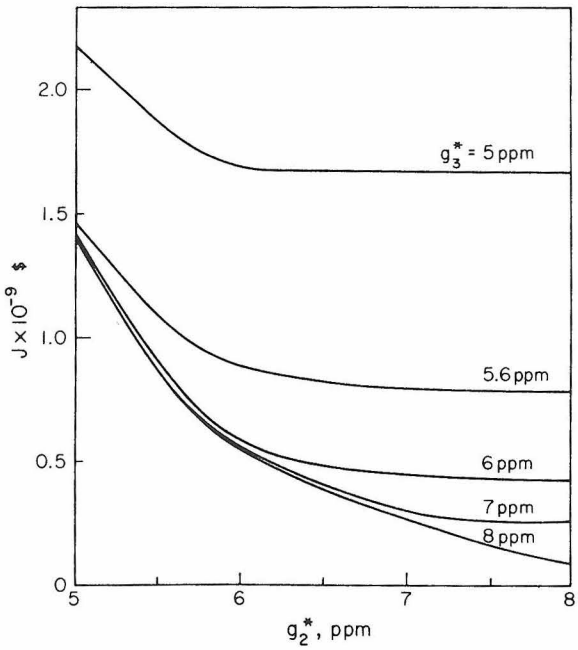


Figure 5.2 The minimum three-year total control cost J as a function of air quality in years 2 and 3

cars first. The total cost of the policy in Table5.5 is $1.158 billion dollars over the three-year period.

Table 5.5 Sample control policy for carbon monoxide for Los Angeles in 1972-1974

	1972	1973	1974
g^*, ppm	8	6	5.5
E_1, grams CO/ 1000 miles	8×10^4	8×10^4	3.395×10^4
E_2, grams CO/ 1000 miles	3.4×10^4	2.373×10^4	0.71×10^3
Δx tons/day	430	1812	223
d_{ij}, control methods	31.2% of the 1966-1969 cars to install flame afterburners	59.3% of the 1965 and earlier cars and 99.3% of the remaining 1966-1969 cars to install flame afterburners	42% of the remaining 1965 and earlier models to install flame afterburners
J, cost in $/day for each year	1.753×10^5	13.86×10^5	16.11×10^5

5.2 Air pollution Control Strategies for the Los Angeles basin from 1973 to 1975

In this section, a detailed study of the control of photochemical smog in the Los Angeles basin from 1973 to 1975 is presented. The specific mathematical formulation and method of solution for the Los-Angeles problem are given in section 5.2.1. Section 5.2.2 summarizes the inventory of sources, control methods and related data. The results of the three-year control study for the Los Angeles basin are presented in section 5.2.3.

5.2.1 Mathematical formulation and method of solution for the Los Angeles problem

The airshed simulation model to be used here is a statistical airshed model developed by Trijonis (1972). It is the only available statistical air pollution model for the Los Angeles basin. This model was developed by a statistical analysis of the actual measurement of air pollution data (pollutant concentration, source emission level etc.) in the Los Angeles basin and is therefore expected to give a more practically relevant results. The detailed development of this model can be found in Trijonis (1972). For our purpose, it suffices to note that the net result of Trijonis' model can be presented as in Figure 5.3, where the number of days per year that both the NO_2 and O_3 standards in downtown Los Angeles can be expected to be violated are plotted as a function of the levels of NO_x and RHC emissions in tons/day. Equivalently, Figure 5.3 be represented by the following algebraic expressions.

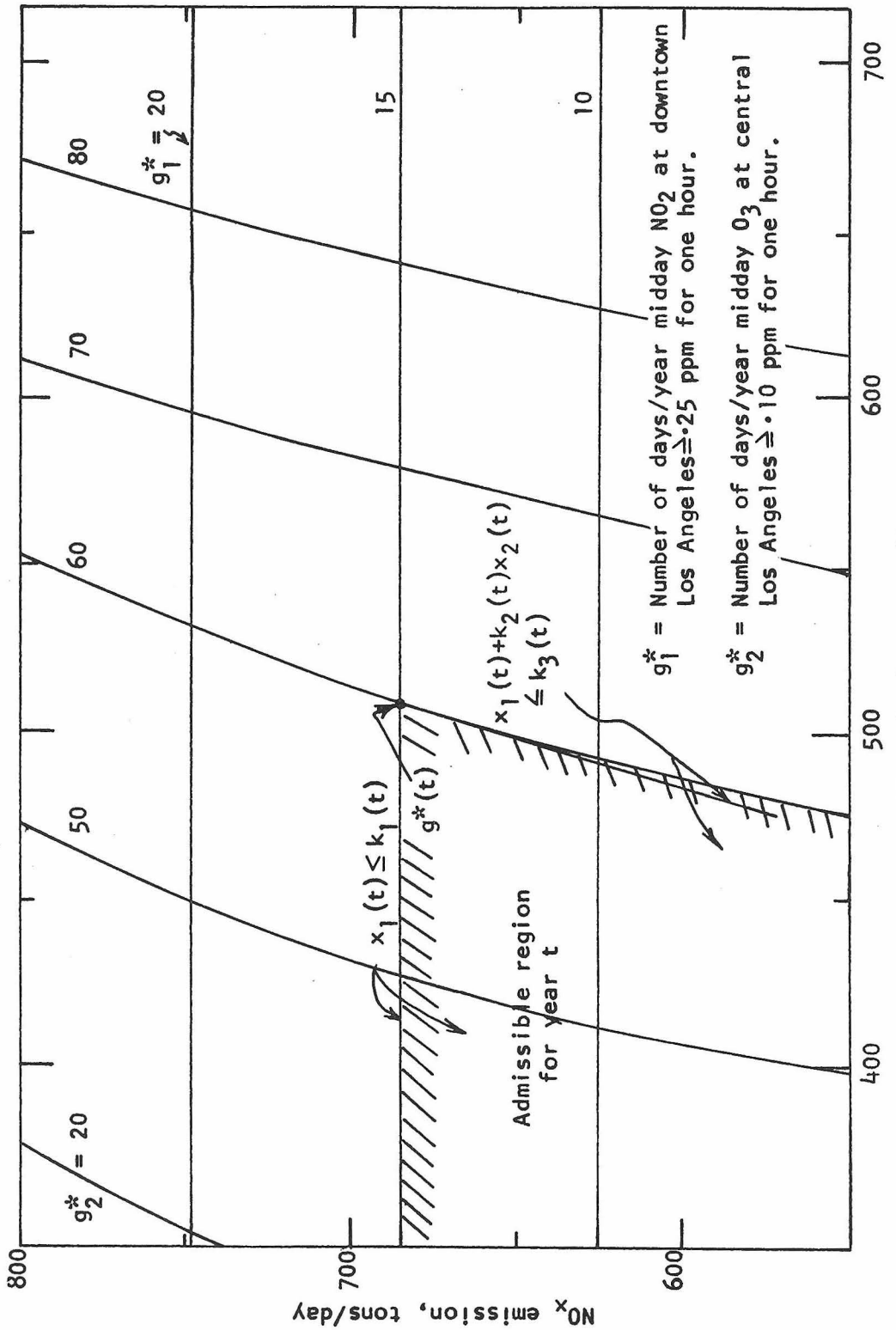


Figure 5.3 Air quality emission relations for Los Angeles basin (Trijonis' statistical airshed model, 1972)

$$g_{NO_2} = g_1 (x_1(t)) \tag{5.9}$$

$$g_{O_3} = g_2 (x_1(t), x_2(t)) \tag{5.10}$$

where g_{NO_2} = air quality for NO_2 defined as the number of days per year midday NO_2 at downtown Los Angeles $\geq$ 25 pphm (parts per hundred million)for one hour

g_{O_3} = air quality for O_3 defined as the number of days per year O_3 at central Los Angeles $\geq$ 10 pphm for one hour

$x_1(t)$ = Emission of NO_x in year t from all the sources in the Los Angeles basin ,tons/day

$x_2(t)$ = Emission of RHC in year t, from all the sources in the Los Angeles basin, tons/day

$g_1(x_1(t))$ = A known algebraic function of $x_1(t)$

$g_2(x_1(t), x_2(t))$ = A known algebraic function of $x_1(t)$ and $x_2(t)$

Thus Trijonis model as represented by (5.9) and (5.10) predicts the air quality directly, given emission levels of NO_2 and RHC.

The general multi-year formulation (3.3) to (3.11) takes the following form for the three-year Los Angeles problem (with year 1 = 1973, year 2 = 1974 and year 3 = 1975)

$$\underset{w(t),t=1,2,3}{\text{Minimize}} \quad J = \sum_{t=1}^{3} \left[\bar{\mu}(t)\, \bar{C}^T(t)\, \bar{w}(t) + \bar{\bar{\mu}}(t)\, \bar{\bar{C}}^T(t)\, \bar{\bar{w}}(t) \right] \tag{5.11}$$

subject to

$$x(t) + R(t)w(t) + \sum_{i=1}^{t-1} \bar{\bar{R}}(i)\bar{\bar{w}}(i) = y^o(t) \tag{5.12}$$

$$A(t)w(t) + \sum_{i=1}^{t-1} \overline{\overline{A}}(i)\overline{\overline{w}}(i) \leqq s(t) \tag{5.13}$$

$$D(t)w(t) + \sum_{i=1}^{t-1} \overline{\overline{D}}(i)\overline{\overline{w}}(i) \leqq \ell(t) \tag{5.14}$$

$$w(t),\ x(t) \geqq 0 \tag{5.15}$$

and $$g(x(t)) \leqq g^*(t) \tag{5.16}$$

$t=1,2,3$

where $g(x(t))$ on the left hand side of (5.16) has the form of (5.9) and (5.10). Since Trijonis'model describes air qualities merely as a function of total emissions of pollutant from all the sources in the airshed, (3.8) to (3.11) collapse to (5.16).

The most crucial parameters of the whole system are g_1^*(air quality criterion for NO_2) and g_2^*(air quality criterion for ozone). If g^* is larger then the current air quality, then no control is necessary and optimal w is identically zero. Equivalently, if constraint (5.16) is absent, clearly solution to (5.11) and (5.15) is $x(t) = y^o(t)$ with $w(t) = 0$, $t = 1, 2, 3$, giving $J = 0$

Each prescribed values for $g^*(t)$, $t=1, 2, 3$ make up an air quality path $g^*(1)$, $g^*(2)$, $g^*(3)$ describing the levels of air quality we are interested in achieving from year 1 to year 3. Now, for each value of g^* for year t, it specifies a point in Figure 5.3 which is the vertex of a region where controlled emission levels $x_1(t)$, $x_2(t)$ must lie as illustrated in Figure 5.3. We can expect the optimally controlled emission levels x(t) to be as close as possible to the vertex point, according to Observations 1 and 2 in chapter 3.

We shall make one useful obsevation on Figure 5.3. The iso-air-quality curves bounding an admissible emission region are approximately straight, especially in the vicinity of the vertex point where the two iso-air-quality curves, corresponding to specified air quality standards, intercept. Since we can expect the optimal $x(t)$ to lie on or in the vicinity of the vertex point, we can safely replace the air quality constraints (5.16) by the following linear algebraic expressions:

$$x_1(t) = k_1(t) \tag{5.17}$$

$$x_1(t) + k_2(t)x_2(t) = k_3(t) \tag{5.18}$$

where $k_i(t)$ are constants for year t so that (5.17) and (5.18) represent the iso-air-quality criteria curves appropriately. (As illustrated in Figure 5.3)

Now, the system (5.11) to (5.15) with (5.17) and (5.18) constitutes a typical linear programming problem with unknowns $w(t)$, $x(t)$, $t=1,2,3$ and therefore can be solved by the Simplex method for any given air quality path (criteria) $g^*(t)$, $t=1,2,3$.

For the present Los Angeles problem, w for each year is 31-dimensional and x is 2-dimensional. The total number of variables is 99 for the three years. The dimensions of $y^o(t)$, $s(t)$, $\ell(t)$ and $g^*(t)$ are 2, 23, 1 and 2 respectively. Therefore for the three years, the number of equations are 84. It takes about one minute of computing time to solve this 84x99 system on the IBM 370/155 computer.

5.2.2 Source inventory and related data

The pertinent source data are presented by Trijonis (1972),

where detailed information on source descriptions, projections of source growth and emission characteristics are given.

Based on the figures of growth rates given by Trijonis (1972), an inventory of sources which are to be considered for institution of multi-year control measures, is given in Table 5.6. The spatial distribution of the major sources are indicated in Figures 4.1 and 4.2. More details are given in LAAPCD Profile (1969) and Roberts et al. (1971).

The Emission inventory of the source shown in Table 5.6, is given in Table 5.7.

Some source are excluded from consideration for control because of one or more of the following reasons:

(a) the sources are already under fairly strict control measures (e.g. 1971 and later new cars are already complied with Federal control requirements; petroleum refining and marketing are already controlled by LAAPCD rules)

(b) the control technology is not yet fully developed (e.g. diesel powered vehicles; catalytic regenerators)

(c) economically or politically unacceptable
(e.g. residential fuel combustion)

(d) not enough cost data available
(e.g. control for metallurgical industries)

(e) detailed information of source inventory not known
(e.g. miscellaneous manufactuning processes; miscellaneous organic solvent users; piston aircraft engines

Table 5.6 Inventory of sources to be controlled

Source No.	Description of sources	Definition of source unit (SU)	SU in 1973	SU in 1974	SU in 1975
1.	Non-power plant large boilers (>30 MBTU/Hr)	one boiler	130	135	140
2.	Medium size boilers (2 to 30 MBTU/Hr)	one boiler	5410	5630	6000
3.	Large refinery heaters (≥ 90 MBTU/Hr)	one heater	60	60	60
4.	Small refinery heaters (< 90 MBTU/Hr)	one heater	160	160	160
5.	Rule-68-complying large power plant boilers (180-350 megawatt)	one boiler	8	8	8
6.	Non-rule-68-complying large power plant boilers (220-480 megawatt)	one boiler	8	8	8
7.	Small power plant boilers (10-175 megawatt)	one boiler	37	37	37
8.	Large stationary internal combustion engines (≥ 300 HP)	one engine	140	140	140
9.	Small compressor engines (< 300 HP)	one engine	360	360	360
10.	Underground service station tanks	one tank	32,000	33,000	34,000
11.	Service station, automobile tank filling	one station	10,800	11,000	11,300
12.	Surface coating operations resulting in emission of reactive hydrocarbon	one ton/day of emitted reactive organic solvents	47	49	51

Table 5.6 continued

Source No.	Description of sources	Definition of source unit (SU)	SU in 1973	SU in 1974	SU in 1975
13.	Degreasers	one ton/day of reactive organic solvent used for degreasing	24	25	26
14.	Dry cleaners using petroleum based solvents.	one dry cleaner plant	25	25	25
15.	Pre-1966 motor vehicle exhaust emissions	one vehicle	1.69×10^6	1.38×10^6	1.10×10^6
16.	Pre-1966 motor vehicle evaporative emissions	one vehicle	1.69×10^6	1.38×10^6	1.10×10^6
17.	1966-1969 motor vehicle exhaust emissions	one vehicle	1.49×10^6	1.45×10^6	1.38×10^6
18.	1966-1969 motor vehicle evaporative emissions	one vehicle	1.49×10^6	1.45×10^6	1.38×10^6
19.	1970 model motor vehicle	one vehicle	0.40×10^6	0.40×10^6	0.39×10^6
20.	1971-1974 model year fleet vehicles suitable for conversion to gaseous fuels	one vehicle	1.76×10^5	2.58×10^5	3.45×10^5
21.	Jet aircraft - JT8D engines	one engine	2500	2500	2500
22.	Jet aircraft - other engines	one engine	2900	2900	2900
23.	Piston aircraft engines registered in Los Angeles County	one engine	6350	6650	7000

Table 5.7 Emission inventory of precontrolled sources (including only the sources to be controlled)

Source No.	Emission/SU, tons/day		1973 emission, tons/day			1974 emission, tons/day			1975 emission, tons/day		
	RHC	NO_x	SU	RHC	NO_x	SU	RHC	NO_x	SU	RHC	NO_x
s1	0	0.2	130	0	26	135	0	27	140	0	28
s2	0	6.8×10^{-3}	5410	0	36.9	5630	0	38.3	6000	0	41
s3	0	0.24	60	0	14.4	60	0	14.4	60	0	14
s4	0	0.06	160	0	9.6	160	0	9.6	160	0	10
s5	0	3.8	8	0	30.4	8	0	30.4	8	0	30
s6	0	9.5	8	0	76.	8	0	76.	8	0	76
s7	0	0.62	37	0	22.94	37	0	22.94	37	0	23
s8	0	0.18	140	0	25.2	140	0	25.2	140	0	25
s9	0	0.02	360	0	6.84	360	0	6.84	360	0	7
s10	5.8×10^{-4}	0.	3.25×10^{4}	18.85	0.	3.32×10^{4}	19.26	0	3.4×10^{4}	20	0
s11	3.6×10^{-3}	0.	1.08×10^{4}	38.88	0.	1.1×10^{4}	39.6	0	1.13×10^{4}	40	0
s12	1.0	0.	47.	47.	0.	49.	49.	0	51.	51.	0
s13	0.9	0.	24.	21.6	0.	25.	22.5	0	26.	23.	0
s14	0.2	0.	25.	5.0	0.	25.	5.0	0	25.	5.	0

Table 5.7 continued

Source No.	Emission/SU, tons/day		1973 emission, tons/day			1974 emission, tons/day			1975 emission, tons/day		
	RHC	NO_x	SU	RHC	NO_x	SU	RHC	NO_x	SU	RHC	NO_x
s15	8.5×10^{-5}	4.5×10^{-5}	1.69×10^{6}	143.7	76.1	1.38×10^{6}	112.9	59.76	1.1×10^{6}	94	50
s16	7.0×10^{-5}	0	1.69×10^{6}	118.37	0	1.38×10^{6}	96.74	0	1.1×10^{6}	77	0
s17	4.7×10^{-5}	10.4×10^{-5}	1.49×10^{6}	70.03	154.96	1.45×10^{6}	68.15	150.8	1.38×10^{6}	65	143
s18	7.0×10^{-5}	0	1.49×10^{6}	104.3	0	1.45×10^{6}	101.5	0	1.38×10^{6}	97	0
s19	5×10^{-5}	1.5×10^{-5}	$4.\times10^{5}$	20	60	$4.\times10^{5}$	20	60	3.9×10^{5}	20	58
s20	9.8×10^{-5}	1.9×10^{-4}	1.74×10^{5}	17.05	33.06	2.58×10^{5}	25.28	49.02	3.45×10^{5}	34	65
s21	5.6×10^{-3}	1.2×10^{-3}	2500	14.0	3.0	2500	14.0	3.0	2500	14	3
s22	1.4×10^{-3}	1.7×10^{-3}	2900	4.06	4.93	2900	4.06	4.93	2900	4	5
s23	1.7x10	6.0x10	6350	10.8	3.81	6650	11.3	3.99	7000	12	4
	Total			633.7	584.		589.3	582.2		556	582

Table 5.8 Emission inventory of sources not to be controlled

No.	Description of sources	reasons#	Emissions, tons/day					
			1973		1974		1975	
			RHC	NO_x	RHC	NO_x	RHC	NO_x
1.	Residential fuel combustion	c	0	25.7	0	26.7	0	28.
2.	Metallurgical industries	b,d	0	8	0	8	0	8
3.	Catalytic regenerators	b,d	0	10	0	10	0	10
4.	Small commercial and industry boilers (<2 MBTU/Hour)	c	0	8.3	0	8.7	0	9
5.	Oil well pump engines*	d,e	0	3	0	2.5	0	2
6.	Miscellaneous stationary sources**	d,e	0	3	0	3	0	3
7.	Petroleum refining	a	7	0	7	0	7	0
8.	Miscellaneous manufacturing processes	a,e	15	0	15	0	15	0
9.	Other organic solvent uses	a,e	13	0	13.5	0	14	0
10.	1971-74 nonfleet vehicles	a	34.8	62.6	51.7	93	69	124
11.	Diesel power vehicles	b,d	0	16.6	0	17.3	0	18
12.	Piston aircrafts not registered in Los Angeles County	e	4.5	1.8	4.7	1.9	5	2
	Total		74.3	139	91.9	171	110	204

a,b, etc. under the column "reasons" correspond to the listings of why these sources are not to be controlled as discussed in the text.

* Diminishing minor source.

** Minor source.

Table 5.9 Precontrolled emissions in Los Angeles from all sources (to be controlled and not to be controlled)

Pollutants	1973	1974	1975
RHC, tons/day	708	681	666
NO_x, tons/day	723	753	786

Table 5.10 Supply limits of natural gas

Projection of natural gas availability	1973	1974	1975
1. Total amount of natural gas available,bbls/year	1.16×10^8	1.13×10^8	1.10×10^8
2. Residential and nonpower plant use, bbls/year	8.23×10^7	8.55×10^7	8.9×10^7
3. Amount available to power plants and fleet vehicles	3.37×10^7	2.75×10^7	2.1×10^7

Table 5.11 Control methods, cost and emission reduction characteristics

Control No.	Description of one unit of control method	Source applicable	Redn./unit control tons/day		Annual $/unit control	Source units/ unit control	Limited supply/ control
			RHC	NO_x			
w1	Low excess air firing(LEAF) to one large boiler	s1	0	0.08	1000	1	0
w2	Low excess air firing and flue gas recirculation (FGR) to one large boiler	s1	0	0.14	10500	1	0
w3	LEAF to one medium boiler	s2	0	2.7×10^{-3}	1270	1	0
w4	LEAF and FGR to one medium boiler	s2	0	4.7×10^{-3}	2700	1	0
w5	LEAF to one large heater	s3	0	9.6×10^{-2}	950	1	0
w6	LEAF to one small heater	s4	0	0.024	1900	1	0
w7	Substitution of 2.6×10^{6} equivalent bbls/year of gas for fuel oil	s5	0	1.3	0	1	2.8×10^{6}
w8	Substitution of 3.9×10^{6} equivalent bbls/year of gas for fuel oil	s6	0	3.2	0	1	3.9×10^{6}
w9	Advanced conbustion modification	s6	0	3.8	2.2×10^{5}	1	0

Table 5.11 continued

Control No.	Description of one unit of control method	Source applicable	Redn./unit control tons/day		Annual cost $/unit control	Source units/ unit control	Limited supply/ unit control
			RHC	NO_x			
w10	Advanced combustion modofi-cation and substitution of 3.9×10^6 equivalent bbls/year of gas for fuel oil	s6	0	5.7	2.2×10^5	1	3.9×10^6
w11	LEAF for one small power plant boiler	s7	0	0.19	1.4×10^4	1	0
w12	LEAF and FGR for small power plant boiler	s7	0	0.31	51000	1	0
w13	Water injection (WI) or exhaust gas recirculation (EGR) to one large engine	s8	0	0.14	680	1	0
w14	WI or EGR to one small engine	s9	0	o.014	190	1	0
w15	Vapor recycle system(VRS) for one tank truck and 90 underground gas tanks	s10	0.052	0	3900	90	0
w16	VRS for one service station nozzle	s11	2.7×10^{-3}	0	330	1	0
w17	Further restrictions on reactive solvents in surface coating	s12	38	0	4.5×10^6	51	0

Table 5.11 continued

Control No.	Description of one unit of control method	Source applicable	Redn./unit control tons/day		Annual cost $/unit control	Source units/ unit control	Limited supply/ unit control
			RHC	NO_x			
w18	Substitution of 0.75 tons/ day of 1,1,1-trichloroethane for 1 ton/day of trichloro-ethylene	s13	0.9	0	-3280	1	0
w19	Activated carbon system to one dry cleaner plant	s14	0.19	0	3300	1	0
w20	Capacitor discharge-ignition optimization system (CDIOS) to one vehicle	s15	5.1×10^{-5}	1.6×10^{-5}	9	1	0
w21	EGR and control spark retar-dation (CSR) to one vehicle	s15	1.3×10^{-5}	2.5×10^{-5}	45	1	0
w22	Evaporative control retrofit (ECR) to one vehicle	s16	6.0×10^{-5}	0	78	1	0
w23	CDIOS to one 1966-1969 vehicle	s17	4.7×10^{-5}	5.7×10^{-5}	1	1	0
w24	Vacuum spark advance dis-connect and tuning adjust-ment	s17	1.4×10^{-5}	4.2×10^{-5}	6	1	0
w25	ECR to one 1966-1969 vehicle	s18	6.0×10^{-5}	0	50	1	0

Table 5.11 continued

Control No.	Description of one unit of control method	Source applicable	Redn./unit control tons/day		Annual cost $/unit control	Source units/ unit control	Limited supply/ unit control
			RHC	NO_x			
w26	CDIOS to one 1970 vehicle	s19	0	8.7×10^{-5}	0	1	0
w27	Conversion of one fleet vehicle to operate on natural gas	s20	9.2×10^{-5}	1.5×10^{-4}	30	1	22.5
w28	Conversion of one fleet vehicle to operate on liquid propane gas	s20	8.0×10^{-5}	1.3×10^{-4}	130	1	0
w29	Combustion chamber re-design on one JT8D engine	s21	5.3×10^{-3}	-2.4×10^{-4}	4000	1	0
w30	Combustion chamber re-design on one non-JT8D engine	s22	1.3×10^{-3}	-3.4×10^{-4}	4000	1	0
w31	After-burner on one piston aircraft engine	s23	1.3×10^{-3}	0	350	1	0

Table 5.11a Data for the system (5.11) to (5.16)#

Temporary control $\bar{w} = (w7, w8, w18)^T$

$\bar{\mu}(1) = \bar{\mu}(2) = \bar{\mu}(3) = 1$; $\bar{\bar{\mu}}(1) = 3$, $\bar{\bar{\mu}}(2) = 2$, $\bar{\bar{\mu}}(3) = 1$

Parameters for the air quality paths

Path		1973	1974	1975
1	g_1^*	19	15	10
	g_2^*	68	59	50
	(5.17)	$x_1(1) = 733$	$x_1(2) = 678$	$x_1(3) = 626$
	(5.18)	$-x_1(1)+3.45x_2(1)$ $= 1276$	$-x_1(2)+3.34x_2(2)$ $= 976$	$-x_1(3)+4.0\ x_2(3)$ $= 1014$
2 (EQL)	g_1^*	13	12	10
	g_2^*	79	63	50
	(5.17)	$x_1(1) = 670$	$x_1(2) = 646$	$x_1(3) = 626$
	(5.18)	$-x_1(1)+4.38x_2(1)$ $=2065$	$-x_1(2)+3.53x_2(2)$ $=1170$	$-x_1(3)+4.0\ x_2(3)$ $= 1014$
3	g_1^*	10	10	10
	g_2^*	80	64	50
	(5.17)	$x_1(1) = 626$	$x_1(2) = 626$	$x_1(3) = 626$
	(5.18)	$-x_1(1)+4.22x_2(1)$ $=2014$	$-x_1(2)+4.04x_2(2)$ $= 1464$	$-x_1(3)+4.0\ x_2(3)$ $= 1014$
4	g_1^*	20	15	10
	g_2^*	50	50	50
	(5.17)	$x_1(1) = 748$	$x_1(2) = 685$	$x_1(3) = 626$
	(5.18)	$-x_1(1)+3.03x_2(1)$ $= 610$	$-x_1(2)+4.15x_2(2)$ $= 1084$	$-x_1(3)+4.0\ x_2(3)$ $= 1014$

The other quantities in (5.11) to (5.16) such as the matrices A, D, R etc. which are large dimensional and can be readily constructed from the data given in Tables 5.6 to 5.11 are not explicitly shown.

not registered in Los Angeles county).

The sources (not to be controlled) are given in table 5.8. Combining tables 5.7 and 5.8 gives the levels of the original source emissions in the Los Angeles basin in 1972, 1973, and 1974 (before institution of controls considered in this study) which are given in Table 5.9. The growth rates of Table 5.8 are similar to those used in Table 5.6

The control methods for the sources listed in Table 5.6, are given in Table 5.11. Detailed descriptions of each of the control methods can be found in references cited by Trijonis (1972).

5.2.3 Results of solving the Los Angeles problem

The results obtainable from the solution of a long-term problem, such as the Los Angeles problem being considered, are the following:

(1) the sources to be controlled, amount of control on the various sources and their controlled status (controlled emission characteristics)

(2) the set of control methods to be used.

(3) the cost incurred by each of the sources and the total control cost

(4) the allocation of the control methods or the time-table of the institution of the control measures over the various years

(5) the pollutants to be preferentially controlled

(6) the effect of the air quality path on the cost

The three-year Los Angeles problem was solved for various air quality paths as shown in Figure 5.4

Some of the results along air quality path 3 and path 2 which approximately corresponds to the EQL uniform reduction strategy are shown in Tables 5.12 to 5.15.

In Table 5.12, the control methods to be instituted on the various sources in each of the three years are shown. The emission reductions of the primary pollutants associated with each of the control methods used, are also indicated. From the given cost data, the cost incurred by the various sources and their controlled status can be easily computed. The yearly cost and the total (from all sources in the Los Angeles basin) emission reductions in the various years are shown in Table 5.13.

In Figure 5.4 , the cost associated with the three principal air quality paths are shown. Air quality path 4 involves a major reduction in RHC emission in the first year whereas air quality path 3 advocates a major NO_x reduction in the first year. Paths 4 and 3 represent two extremes, with paths 1 and 2 intermediates between them. As indicated in Figure 5.4 the control cost incurred along air quality path 3 is the smallest among all the paths shown. The heavy cost associated with path 4 can be attributed to two factors: (1) the percentage reduction in RHC emission is quite large in year 1973, resulting in institution of more costly controls[#] and (2) RHC emission,

---

e.g. along path 4, control of pre-1966 motor vehicles are called for in 1973, whereas along path 3, it is not called for.

without any additional control, decreases by itself because of attrition of older car and the introduction of new (cleaner) cars. Thus, it does not seem advantageous to control RHC heavily in year 1973.

A comparison of EQL strategy # 1 designated as path 2 (uniform reduction strategy) and path 3 are shown in Table 5.14. The entries in Table 5.14 under EQL strategy # 1 were calculated assuming a uniform rate of reduction of RHC and NO_x emissions along 1973 to 1975. It can be seen in Table 5.15, that the cost incurred by EQL strategy # 1 is only slightly greater than that of path 3. Therefore, the EQL strategy # 1 appears to be a very good multi-year air pollution control strategy.

The control methods used along air quality path 3 can be categorically grouped as:

(1) Fleet vehicles to burn natural gas

(2) Controlling automobiles' RHC and NO_x emissions by

  (a) adding capacitor discharge ignition optimization system (CDIOS) to pre-1966 vehicles.

  (b) adding vaccum spark advance disconnect and tuning system and evaporative control retrofit to 1966-1969 vehicles.

  (c) adding CDIOS to 1970 vehicles.

(3) Use of natural gas in power plants

(4) Use of low excess air in industrial boilers and heaters

(5) Reduction of reactive organic vapors by solvent users

(6) Vapor recirculation systems for gasoline tank trucks and gas station nozzles and underground tanks.

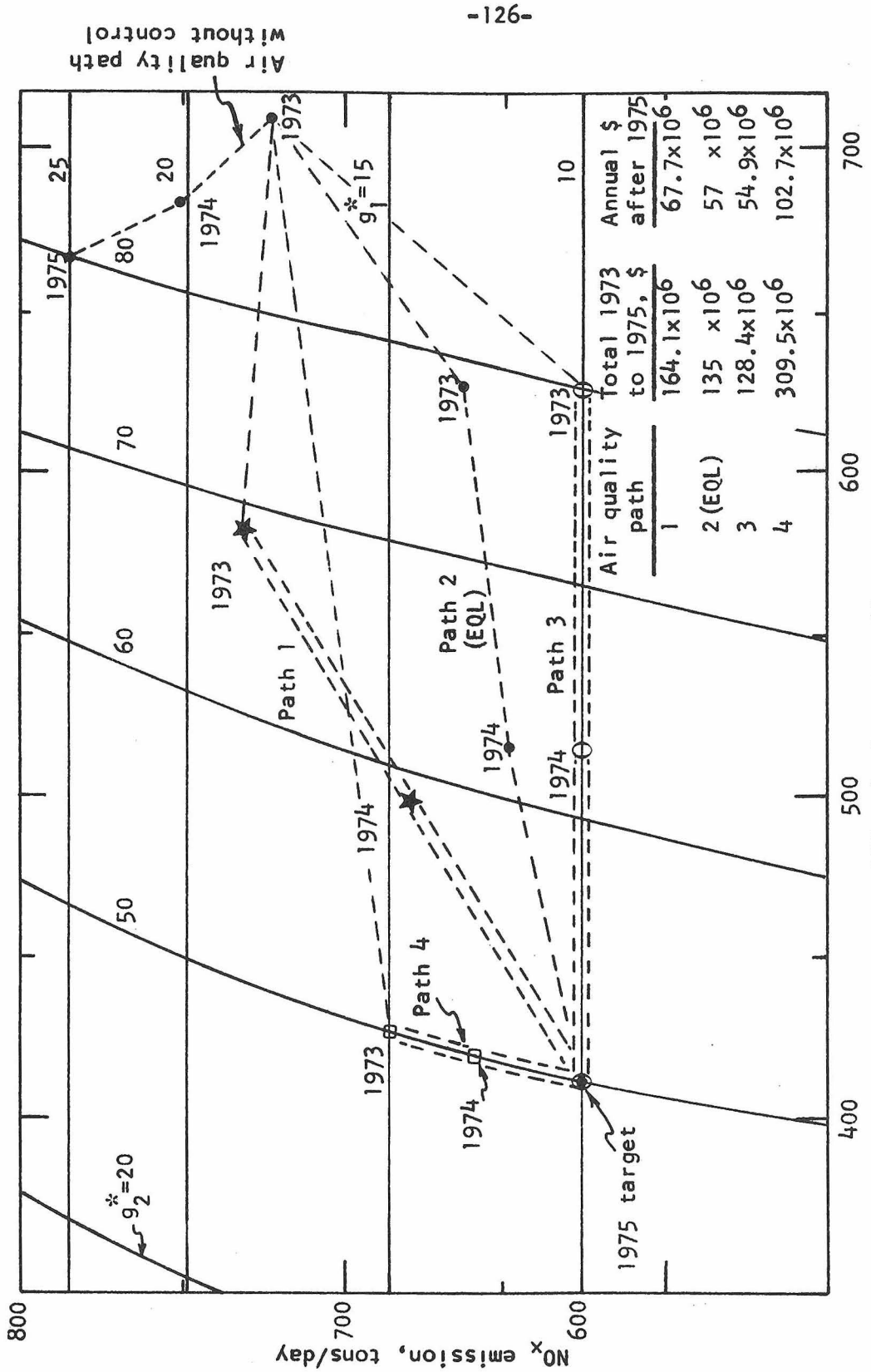


Figure 5.4 Multi-year air quality paths and their associated costs.

Table 5.12 Distribution of control effect on a yearly basis along air quality path 3

New controls instituted in each year

1973			1974			1975		
Control methods units	NO_x reductions tons/day	RHC reductions tons/day	Control methods units	NO_x reductions tons/day	RHC reductions tons/day	Control methods units	NO_x reductions tons/day	RHC reductions tons/day
w7=8	10.4	0	w5=34	3.27	0	w5=25	2.5	0
w8=1.89	6.06	0				w8=2.9	9.2	0
w15=52.8	0	2.75	w15=162	0	8.4	w15=163	0	8.5
w18=24	0	21.6	w18=25	0	22.5	w16=1.13 $\times 10^4$	0	30.5
w19=25	0	4.75				w17=1	0	37.6
			w20=1.1 $\times 10^6$	17.6	56.1	w18=26	0	23.4
w24=1.38 $\times 10^6$	58	19.3						
w25=1.29 $\times 10^5$	0	7.74						
w26=1.15 $\times 10^5$	10	0						

Table 5.12 Continued

New controls instituted in each year

1973			1974			1975		
Control methods units	NO_x reductions tons/day	RHC reductions tons/day	Control methods units	NO_x reductions tons/day	RHC reductions tons/day	Control methods units	NO_x reductions tons/day	RHC reductions tons/day
						w26=7.5 $\times 10^4$	6.5	0
w27=1.74 $\times 10^5$	26.1	16	w27=8.4 $\times 10^4$	12.6	7.73	w27=8.7 $\times 10^4$	13.1	8
w29=2.5 $\times 10^3$	-0.6	13.2				w31=350	0	1

Note:- The above table lists the fresh control methods instituted in each year. The total amount of control measures in effect in each year consists of the fresh control and the control of the permanent type instituted in the prior years.

Table 5.13 Yearly cost and reduction along path 3.

	1973	1974	1975
Yearly cost	$30x10^6$	$43x10^6$	$55x10^6$
Yearly NO_x reduction	116	127	160
Yearly RHC reduction	86	165	256

total 3-year cost = $128.0x10^6$

Table 5.14 Comparison of path 2(EQL) and path 3

		Air quality path 3			EQL strategy #1 (path 2)		
		1973	1974	1975	1973	1974	1975
Yearly new reductions tons/day	NO_x	114	23	33	53.3	53.3	53.3
	RHC	86	100	113.5	85.3	85.3	85.3
Cumulative yearly reduction, tons/day	NO_x	114	127	160	53.3	106.6	160
	RHC	86	165	256	85.3	170.6	256
Reduction emission levels tons/day	NO_x	609	626	626	669.7	646.4	626
	RHC	622	516	410	623	510.4	410
Air quality as defined in Fig. 5.3	NO_x	10	10	10	13	12	10
	RHC	80	63	50	78	62	50

Table 5.15 Costs along path 2(EQL) and path 3

	Total 3-year cost	Yearly cost after 1975
Air quality path 3	$ 128×10^6	$ 55×10^6
EQL strategy #1 (path 2)	$ 135×10^6	$ 57×10^6

From Figure 5.4 , we can see that ozone air quality is most efficiently controlled by decreasing RHC emission. Thus RHC emission is the most sensitive to improve ozone air quality, as was found in the real-time control study described in Chapter 4.

From Table 5.12, one can obtain the information on which sources and what control methods are of the major importance with regard to reduction of a given pollutant emission at least cost.

For major NO_x reduction (Table 5.12, in year 1973 where major NO_x reduction is called for) and major RHC reduction (Table 5.12, year 1973, 1974 and 1975), the important sources and control methods are given below:

Table 5.16 Major sources for control.

Major pollutant reduction	Major sources to be controlled	Major control methods
NO_x	1.1966-1969 vehicles	Vacuum spark advance disconnect and tuning
	2.1971-1974 fleet	fleet vehicle to burn natural gas
	3.power plants	to burn natural gas
RHC	1.pre-1966 vehicles	capacitor discharge ignition optimization
	2.organic solvent users	further restriction on organic solvent users
	3.service stations	vapor recycle systems
	4.degreasers	substition of 1,1,1, Trichloro-ethane for Trichloroethane.

We see from the above table that automobiles are the most important source to be controlled with respect to NO_x and RHC emission reductions. The entries in Table 5.16 under "major sources" are listed in decreasing order of importance.

5.2.4 Comment on the results of the long-term control

The validity of the results of the long-term control obtained on the previous section depends heavily on the accuracy of the simulation model and the emission inventory and related data. Thus, a brief discussion on the applicability of the statistical airshed model is given below. The discussion of the emission inventory can be found in Trijonis (1972).

The major assumption used in the development of Trijonis' statistical model is that the control measures apply uniformly throughout the whole airshed (Trijonis, 1972). Thus for those control measures that can meet this assumption, the model and hence the long-term control results can be expected to be reasonably reliable. Control measures on motor vehicles (and on service stations) appear to meet this homogeneity assumption since motor vehicles can be considered uniformly distributed throughout the airshed. Since the long-term control strategy obtained in the previous section primarily calls for the control of the motor vehicles, our results can be expected to have a reasonable validity.

5.2.5 Accuracy and sensitivity analysis of the Los Angeles problem

To facilitate a discussion of the accuracy and sensitivity of the results of the Los Angeles problem, the various parameters involved are listed below:

(1) Source parameters: Source units in each year (including source growth rate projections); spatial and temporal source inventory; source emission characteristics.

(2) Control method parameters: Control cost per unit of control; reduction of pollutant emission by each control method; number of source units controllable by one unit of control method; number of units of limited supply consumed by one unit of control method.

(3) Meteorological (and kinetic) parameters: Wind speeds and directions; inversion heights; solar irradiation; atmospheric mixing parameters; (kinetic parameters).

In a decreasing order of accuracy by which these parameters can be estimated are: control method parameters, source parameters and meteorological parameters. Control method parameters can be accurately obtained from manufacturer's informations or can be ascertained by some actual testing. The source parameters can be obtained with reasonable accuracy by a suitable source modeling (e.g. Roberts et al.1971) given enough manpower. However, there is some uncertainty in the estimation of source growth projections for the future years. The parameters that are the least accurate are the meteorological parameters due to the random nature of the atmosphere. These data are also difficult and

expensive to measure, since they have to be measured not only on a region-wide basis, but also on a day in and day out or a year in and year out basis.

The sensitivity of the control results (e.g. optimal cost of control, optimal control strategy) is a very difficult question to address quantitatively, since the number of parameters involved is very large and the system structure is very complex. Any quantitative evaluation of the sensitivity of the results to these parameters will be prohibitively expensive. Therefore, only a qualitative discussion is given below:

As was pointed out in chapter 3, the solution of a given problem is most sensitive to the air quality standard. Specification of a given level of the air quality standard, in fact, specifies the problem and its solution. The cost of optimal controls increases with a decrease in the level of the air quality standard and increases sharply for a slight decrease in the level of the air quality standard when the standard becomes quite strict (as can be seen in Figure 5.2).

The parameters of the cost per unit control method appear in equation (5.11) defining the cost functional J. The source parameters appear on the righthand side of the linear programming constraints (5.12) to (5.14). The control method parameters, except the cost per unit control method, appear on the lefthand side of the linear programming constraints. It is well known in the theory of linear programming (Gale, 1960) that the minimal cost J is a continuous function of the costs per unit control method and the righthand side of the linear

programming constraints, but it needs not depend continuously on the parameters appearing on the lefthand side of these constraints (i.e. the coefficient matrix of the linear programming constraints). Therefore, a slight change in the value of one source parameter (say, the number of 1970-model motor vehicles) will only produce a slight change in the minimal control cost J. However, a slight change in one of the control parameters (say, the reduction of NO_x emission from one pre-1966 motor vehicle by the institution of one vacuum spark advance disconnect and tuning device) may produce a significant change in the minimal control cost or may have no effect at all on the minimal cost. Therefore, from the point of view of sensitivity, the control method parameters are the most important (given a specified level of air quality standard). Fortunately, the control method parameters are probably the most accurate data available.

At the present stage of airshed modeling, no definite information is available relating to the sensitivity of the simulation results to the airshed parameters (source parameters, meteorological parameters and kinetic parameters). However, we would imagine that the simulation results are more sensitive to the meteorological parameters which constitute the coefficient matrix of the airshed (dynamic) modeling equations. In comparing alternative control strategies, the same set of the meteorological data was employed. Thus, we are looking at relative effects of the strategies. This diminishes the influence of inaccuracies in the meteorological inputs.

In summary, the parameters in a decreasing order of importance

relating to the sensitivity of the results of the long-term problem appear to be as indicated below:

For the control-method-emission problem:

(a) Level of air quality standard.

(b) Control method parameters.

(c) Source parameters.

For the emission-air-quality problem:

(a) Level of air quality standard.

(b) Meteorological parameters.

(c) Source parameters (emissions).

For the overall problem, it is not quite clear how meteorological parameters would compare with the other parameters relating to the sensitivity problem. However, since the control-method-emission problem, in effect, inputs a certain set of emissions (corresponding to an optimal strategy) into the emission-air-quality problem and since meteorological parameters appear to be more important than the source parameters relating to the sensitivity of the emission-air-quality problem, we would expect meteorological parameters to have an over-riding importance on the final results of the overall problem.

In conclusion, it appears that sensitivity should not be a problem for the control-method-emission problem, since the more sensitive parameters can be more accurately estimated. The performance of the overall problem should be enhanced with a better airshed model and a more accurate or representative set of meteorological data.

CHAPTER 6 CONCLUSIONS

We summarize here the results of the previous chapters.

In chapter 2, the general concepts of an airshed system and its control were introduced. The general problem structure for air pollution control and its decomposition into two simpler problems were delineated. The concepts of long-term and real-time controls were introduced, together with the notions of feedforward and feedback controls for air pollution.

In chapter 3, the problem of long-term air pollution control was formulated mathematically. The problem considered is to determine the set of control measures that minimizes the total cost of controls over a multi-year period while maintaining a specified level of air quality each year. Various computational methods were developed for the solution of this problem based on mathematical programming techniques. The general mathematical formulation and the computational methods are applicable to any airshed and any airshed simulation model.

In chapter 4, real-time air pollution control for an urban airshed was posed as selecting those measures from among all those possible such that the air quality is maintained at a certain level over a period of several hours to several days and the total control imposed is a minimum. A computational algorithm was then developed for solving the class of control problem that results. The control strategy is assumed to be enforced over a certain time period, say, one hour, based on meteorological predictions made at the beginning of the period. The strategy for each time period for the various locations

of an airshed can be determined by an air pollution control agency by means of a computer implementing the algorithm developed. The theory is illustrated by an application to a hypothetical study of implementations of the optimal control on September 29,1969 in the Los Angeles basin.

In chapter 5, the theory of the long-term control is illustrated by a detailed study for the evaluation of optimal control strategies for the control of NO_2 and O_3 air qualities for 1973 to 1975 in the Los Angeles basin. An analysis of the accuracy and sensitivity of the results of the long-term control problem to the various parameters involved was attempted.

REFERENCES

Altshuller, A. P. and J. J. Bufalini, "Photochemical Aspects of Air Pollution: A Review," Environmental Science and Technology, 5, 39 (1971).

Babcock, L. R. Jr. and N. L. Nagda, "Cost Effectiveness of Emission Control," JAPCA, 23(3), 173-179 (March 1973).

Barret, L. B. and T. E. Waddell, "The Cost of Air Pollution Damages: A Status Report," U. S. Dept. of Health, Education, and Welfare, NAPCA (USPHS), (July 1970).

Bibbero, R. J. "Systems Approach toward Nationwide Air Pollution Control 1. The Problem, the System, the Objective," IEEE Spectrum, 8(10), 20-31 (October 1971).

Bibbero, R. J. "Systems Approach toward National Air Pollution Control, 2. The Technical requirements," IEEE Spectrum, 8(11), 73-81 (November 1971).

Bibbero, R. J. "Systems Approach toward Nationwide Air Pollution Control 3. Mathematical Models," IEEE Spectrum, 8(12), 47-58 (December 1971).

Bounds, J. M. "The Formulation of Emission Control Strategies Using a Cubic Polynomial Approximation," JAPCA, 21(4), 210-213 (April 1971).

Burton, E. S. and W. Sanjour, "A Simulation Approach to Air Pollution Abatement Program Planning," Socio-Econ. Plan. Sci. 4, 147-159 (1970).

Croke, E. J. and S. G. Booras, "The Design of an Air Pollution Incident Control Plan," APCA Paper NO. 69-99, Presented at the 1969 Annual APCA Meeting, New York.

Croke, K. G. and E. J. Croke, "The Influence of Natural Gas Availability on Air Pollution Episode Control Feasibility," JAPCA, 20(10), 649-652 (October1970)

Croke, E. J. and J. J. Roberts," Air Resource Management and Regional Planning," Bull. Atomic Scientists, PP. 8-12, (February 1971).

Downing, P. B. and Stoddard, L., "Benefit/Cost Analysis of Air Research Reports, Volume 3, University of California, September 1, 1970.

Eschenroeder, A. Q. and J. R. Martinez, "Concepts and Applications of Photochemical Smog Models," Technical Memorandum 1516, General Research Corporation, Santa Barbara, California (1971).

Farmer, J. R., P. J. Bierbaum and J. A. Tikvart, "Proceeding from Air Quality Standards to Emission Standards," APCA Paper No. 70-85, Annual Meeting of the Air Pollution Control Association, St. Louis (1970).

Fensterstock, J. C., J. A. Kurtzweg and G. Ozolins, "Reduction of Air Pollution Potential through Environmental Planning," JAPCA, 21(7), 395-399 (July 1971).

Frankel, R. J., "Problems of Meeting Multiple Air Quality Objectives for Coal-fired Utlity Boilers," JAPCA, 19(1), 18-23 (January 1969).

Friedlander, G. D., "Computer-controlled Vehicular Traffic," IEEE Spectrum,6(2), 30-43 (February 1969).

Friedlander, S. K. and J. H. Seinfeld, "A Dynamic model of Photochemical Smog," Env.Sci. Technol. 3, 1175-1181 (1969).

Hamburg, F. C. and F. L. Cross, Jr., "A Training Exercise on Cost-effectiveness Evaluation of Air Pollution Control Strategies," JAPCA, 21(2), 66-70 (February 1971).

Hecht,T. A. and J.H. Seinfeld, "Development and Validation of a Generalized Mechanism for Photochemical Smog," Environmental Science & Technology, 6, 47-57 (1972).

Herzog, Henry W. Jr., "The Air Diffusion Model as an Urban Planning Tool," Socio-Econ. Plan. Sci. 3, 329-349 (1969).

Johnston, H. S., J. N. Pitts, Jr., J. Lewis, L. Zafonte and T. Mottershead, "Atmospheric Chemistry and Physics," Project Clean Air, Vol. 4, University of California (1970).

Kamrany, N. M., "Economic Growth and Environmental Impact: Evaluating Alternatives," Socio-Econ. Plan. Sci., 7(1), 37-53 (February 1973).

Kleinman, F. K., "The Regional Approach to Air Pollution Control," JAPCA, 21(2), 71-75 (February 1971).

Kohn, R. E., "A Linear Programming Model for Air Pollution Control in the St. Louis Airshed," Ph.D. Thesis, Department of Economics, Washington University, St. Louis (1969).

Kohn, R. E., "A Linear Programming Model for Air Pollution Control: A Pilot Study of the St. Louis Airshed," Journal of Air Pollution Control Association, 20, 78-82 (1970).

Kohn, R. E., "Abatement Strategy and Air Quality Standards, in Developmant of Air Quality Standards," edited by A. Atkisson and R. Gaines, pp. 103-123. Charles E. Merrill (1970).

Kohn, R. E., "A Cost Effectiveness Model for Air Pollution Control with a Single Stochastic Variable," J. Amer Statistical Assoc., 67, 337, 19 (1972).

Kyan, C. P. and J. H. Seinfeld, "Determination of Optimal Multi-Year Air Pollution Control Policies," J. Dyn. Syst. Meas. Control., 94G, 266 (1972).

Lamb, R. G., "Numerical Modeling of Urban Air Pollution," PhD thesis,Department of Meteorology, UCLA, September (1971).

Lamb, R. G. and M. Neiburger, "An Interim Version of a Generalized Urban Air Pollution Model, " Atmospheric Environment, 5, 239-264 (1971).

Leavitt, J. M., S. B. Carpenter, J. B. Blackwell and T. L. Montgomery, "Meteorological Program for Limiting Power Plant Stack Emission," JAPCA, 21(7), 400-405 (July 1971).

Lemke, E. E., "Profile of Air Pollution Control," Los Angeles County Air Pollution Control District, (1971).

MacCracken, M. D., T. V. Crawford, K. R. Peterson and J. B. Knox, "Development of a Multibox Air Pollution Model and Initial Verification for the San Francisco Bay Area," Lawrence Radiation Laboratory Report UCRL-73348, Livermore, Calif. (1971).

Malone, D. W., "Modeling Air Pollution Control as a Large Scale, Complex System," Soci-Econ. Plan. Sci., 6(1), 69-85 (February 1972).

Monin, A. S. and Yaglom, A. M., Statistical Fluid Mechanics, MIT Press, Cambridge (1971).

Morgenstern, P., and K. A. Hagg, "A System for Abatement Control Strategy Evaluation," JAPCA, 22(10), 774-778 (October 1972).

Muller, F., "An Operational Mathematical Programming Model for the Planning of Economic Activities in Relation to the Environment," Socio-economic Planning Sciences 7(2), 123-138 (April 1973).

Parson, D. O., and E. J. Croke, "An Economic Evaluation of SO_2 Air Pollution Incident Control," APCA Paper 69-20, N. Y., N. Y., June 22-26, (1969).

Pasquill, F., "Atmospheric Diffusion," D. Van Nostrand, London, (1962).

Randerson, D., "A Numerical Experiment in Simulating the Transport of Sulfur Dioxide Through the Atmosphere," Atmospheric Environment, 4, 615-632 (1970).

Reynolds, S. D., M. Liu, T. A. Hecht, P. M. Roth and J. H. Seinfeld, "Evaluation of a Simulation Model for Estimating Ground Level Concentrations of Photochemical Pollutants," Systems Applications, Inc., (1973).

Reiquam, H., "An Atmospheric Transport and Accumulation Model for Airshed," Atmospheric Environment, 4, 233-247 (1970).

Reiquam, H., "A method for Optimizing Pollutant Emissions in an Airshed," Atmospheric Environment, 5(1), 57-64 (January 1971).

Roberts, P. J. W., M. Liu and P. M. Roth, "A Vehicle Emissions Model for thr Los Angeles Basin - Extension and Modifications," Systems Applications, Inc., Beverly Hills, California, R72-8 (1972).

Roberts, P. J. W., P. M. Roth and C. L. Nelson, "Contaminant Emissions in the Los Angeles Basin - Their Sources, Rates and Distribution, "Appendix A of Development of a Simulation Model for Estimating Ground Level Concentrations of Photochemical Pollutants, System Applications, Inc., Beverly Hills, California, 71SAI-6 (1971).

Roberts, P. J. W., M. K. Lin, S.D. Reynolds and P. M. Roth, "Extension and Modifications of a Contaminant Emissions Model and Inventory for Los Angeles-Appendix A of Further Development and Validation of a Simulation Model for Estimating Ground Level Concentrations of Photochemical Pollutants," SAI Report R73-15, Systems Applications, Inc, Beverly Hills, California (January 1973).

Roberts, S. M., "Dynamic Programming in Chemical Engineering and Process Control," Academic Press, New York (1964).

Rossin, A. D., and J. J. Roberts, "Episode Control Criteria and Strategy for CO," JAPCA, 22(4), 254-259 (April 1972).

Roth, P. M., S. D. Reynolds, P. J. M. Roberts and J. H. Seinfeld, "Development of a Simulation Model for Estimating Ground-Level Concentrations of Photochemical Pollutants," Report 71SAI-21, Systems Applications, Inc., Beverly Hills, California (1971).

Savas, E. S., "Computers in Urban Air Pollution Control Systems," Socio-Econ. Plan. Sci. 1, 157-183 (1967).

Savas, E. S., "Feedback Controls on Urban Air Pollution," IEEE Spectrum, 6(7), 77-81 (July 1969).

Seinfeld, J. H., "Mathematical Models of Air Quality Control Regions," in Development of Air Quality Standards, edited by A. Atkisson and R. Gaines, pp. 169-196. Charles E. Merrill (1970).

Seinfeld, J. H., "Modeling Problems in Air Pollution, in Mathematical Models of Public Systems— Proceedings of the 1970 National Seminar on Advanced Simulation, edited by G. Bekey. Simulation Councils, Inc. and the University of Southern California (1971).

Seinfeld, J. H., T. A. Hecht and P. M. Roth, "A Kinetic Mechanism for Atmospheric Photochemical Reactions," Appendix B of Development of a Simulation Model for Estimating Ground Level Concentrations of Photochemical Pollutants, Systems Applications, Inc., Beverly Hills, California, 71SAI-9 (1971).

Seinfeld, J. H., and C. P. Kyan, "Determination of Optimal Air Pollution Control Strategies," Socio-economic Planning Sciences, 5, 173-190 (1971).

Seinfeld, J. H., and L. Lapidus, "Aspects of the Forward Dynamic Programming Algorithm," I&EC Process Design and Development, 7, 475-478 (1968).

Seinfeld, J. H., P. M. Roth and S. D. Reynolds, "Simulation of Urban Air Pollution," Advances in Chemistry, in Press (1972).

Shepard, D. S., "A Load Shifting Model for Air Pollution Control in the Electric Power Industry," JAPCA, 20(11), 756-761 (Nov. 1970).

Sklarew, R. C., "A New Approach: The Grid Model of Urban Air Pollution," Air Pollution Control Association Paper No. 70-79 (1971).

Slade, D. H., ed., Meteorology and Atomic Energy 1968, U. S. Atomic Energy Commission, TID-24190, (July 1968).

Smith, D. R., N. G. Edmisten and D. J. deRoeck, "System for Implementing Comprehensive Air Pollution Control Programs," JAPCA, 22(12), 943-949 (Dec. 1972).

Sporn, P., "Our Environment - Options on the way into the Future," IEEE Spectrum, 8(5), 49-58 (May 1971).

Sussman, V. H., "New Priorities in Air Pollution Control," JAPCA, 21, 201-203 (April 1971).

Trijonis, J., "An Economic Air Pollution Control Model-Application: Photochemical Smog in Los Angeles County in 1975," Ph.D. Thesis, California Institute of Technology, Pasadena, Calif. (1972).

Turner, D. B., Workbook of Atmospheric Dispersion Estimates, U. S. Dept. of H.E.W., (1969).

Ulbrich, E. A., "Adapredictive Air Pollution Control for the Los Angeles Basin," Socio-economic Planning Sciences, 1, 423-440 (1968).

Wayne, L. G., R. Danchick, M. Kokin and A Stein, "Modeling Photochemial Smog on a Computer for Decision-making," Journal of Air Pollution Control Association, 21, 334 (1971).

Wilde, D. J., "Optimum-Secking Methods," Prentice-Hall, Inc., Englewood Cliffs, N. J. (1964).

Wilson, R. D. and D. W. Minnotte, "A Cost Benefit Approach to Air Pollution Control," JAPCA, 19(5), 303-308 (May 1969).

Supplementary References

Bryson, A. E. Jr. and Y.-C. Ho,"Applied Optimal Control" Ginn and Company, Waltham, Massachusetts, 1969.

Dantzig, G.,"Linear Programming and Extensions", Princeton University Press, Princeton, N. J., 1963.

Dreyfus, S. E.,"Dynamic Programming and the Calculus of Variations", Academic Press, New York, 1965.

Gale, D.,"The Theory of Linear Economic Models", McGraw-Hill Book Company, New York, 1960.

Kyan, C. P. and J. H. Seinfeld,"Real-time Control of Air Pollution", accepted for publication in A.I.Ch.E. Journal,1973.

APCD Profiles, published by the Los Angeles County Air Pollution Control District, 1969 and 1971.

Reynolds, S. D., Ph. D. Thesis, California Institute of Technology, Pasadena, California, in preparations (1973).

Proposition I

Analysis and Control of a Batch Polystyrene Reactor

(Accepted in Candidacy Examination, February 15, 1971)

Abstract

The analysis and computer simulation study are carried out for a batch polystyrene reactor which is typical of that class of chemical systems having large heats of reaction and poor transport properties and thus rendering their operation and control rather difficult. In this proposition, techniques for estimating stable operating conditions are presented from which feasible control methods are developed.

Table of Content

Introduction

Many chemically reacting systems of industrial importance are difficult to maintain at their favorable operating conditions due to the large heat of reactions and the poor transport properties of the systems. These systems are characterised by the common behavior that they either hardly react at lower temperature or proceed so fast at the favorable reacting temperature that the reactions experience a "runaway" or "explosion" and end up with undesirable products. Typical examples are polymerization reactions and catalytic gas-phase reactions. It will be of interest to study the behavior of these systems especially those governed by partial differential equations and to investigate possible methods for controlling them at their favorable operating conditions with stable reactor performance.

In this proposition, the batch production of impact polystyrene is chosen as a specific case study. Several points of interest with respect to stability and possible control methods are to be treated.

The kinetics and physical data were summarized from a report by Dr. Seinfeld.

I. Kinetics of Thermal Polymerization of Styrene

The polymerization of pure styrene monomer has been extensively studied (1,2).

The thermal polymerization reactions are initiated by the formation of free radicals on the rupture of the double bond of styrene monomer. This has been accepted to be caused predominantly by a bimolecular collision mechanism (2).

$$M + M \xrightarrow{k_i} \cdot M + M\cdot$$

The propagation step in free radical polymerization proceeds as

$$M_x^{\cdot} + M \xrightarrow{k_p} M_{x+1}^{\cdot}$$

Where $M_x^{\cdot}$ is an active polymer molecule of length x.

Termination reaction is caused by combination of free radicals

$$M_x^{\cdot} + M_y \xrightarrow{k_t} M_{x+y}$$

and by disproportionation

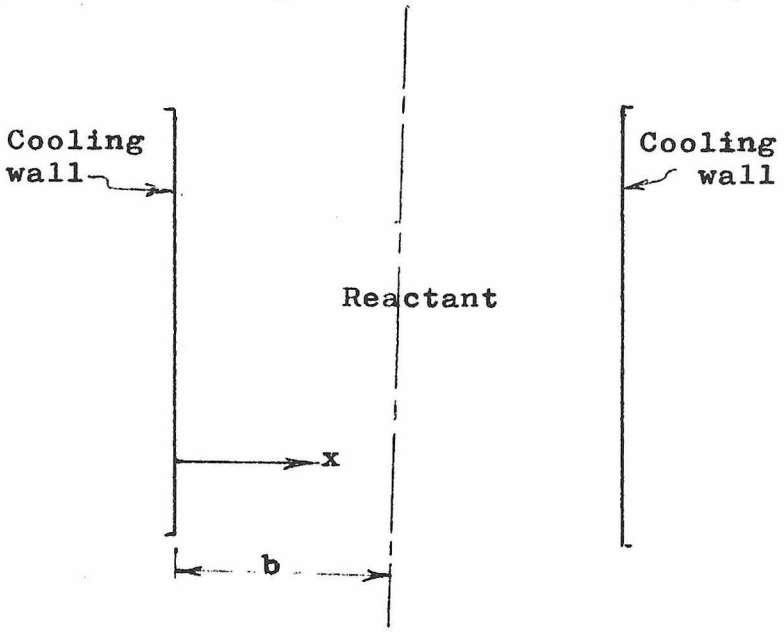
$$M_x^{\cdot} + M_y^{\cdot} \xrightarrow{k_t} M_x + M_y$$

In solution polymerization, a second order kinetics is well accepted. But the bulk-phase thermal polymerization is believed to follow a first order kinetics(3). However, many of the data given in the literature for the overall rate constant are obtained from solution polymerization

experiments. Because of the lack of data for the first order kinetic rate constant, conversion vs. time data obtained from curves in Boundy and Boyer at 100°C, 140°C were analysed in terms of a first order mechanism (3) and it was found that the reaction could be considered first order up to 40% conversion at 100°C, up to 60% at 120°C and up to 80% conversion at 140°C. The first order reaction rate constant obtained, $A\exp(-E/RT)$, had the values $A=5.68\times10^{6}$ sec^{-1} and E=20·33 kcal/mole.

II. System Equations

For the purpose of simulation, the batch reactor for the production of impact polystyrene is to be represenented by the following two-dimensional configuration.



During the period of heating the reactant from the ambient temperature to the reaction temperature, the fluidity of the reactant permits adequate stirring to avoid large temperature gradient in the reactor. After the heat-up period, as soon as significant polymerization begins, there is a sharp increase in the viscosity of the reacting medium and adequate agitation is no longer possible and the major heat transfer mechanism is conduction(4). It is during and after this period that the control of the reactor to prevent any unstable temperature (runaway temperature) becomes critical. Thus, the simulation is for the case of pure conduction alone, ignoring the agitation effect.

By energy and material balances, the system equations are:

$$\frac{\partial T(x,t)}{\partial t} = \frac{k}{\rho C_p} \frac{\partial^2 T(x,t)}{\partial x^2} + \frac{\Delta H}{\rho C_p} A \exp(-E/RT(x,t)) C_M(x,t) \tag{1}$$

$$\frac{\partial C_M(x,t)}{\partial t} = -A \exp(E/RT(x,t)) C_M(x,t) \tag{2}$$

$$0 \leq x \leq b, \quad 0 \leq t$$

with boundary conditions

$$T(0,t) = g(t) \quad , \; t>0 \tag{3}$$

$$\frac{\partial T(b,t)}{\partial x} = 0 \tag{4}$$

and initial conditions

$$T(x,0) = T_o \tag{5}$$

$$C_M(x,0) = C_{Mo} \tag{6}$$

Introducing the following dimensionless quantities, with T_r being some convenient reference temperature,

$$\xi = x/b \tag{7}$$

$$\tau = kt/ C_p b^2 \tag{8}$$

$$\Theta = T/T_r \tag{9}$$

$$G = g/T_r \tag{10}$$

$$\Gamma = C_M/C_{Mo} \tag{11}$$

$$\varepsilon = E/RT_r \tag{12}$$

$$\beta = Ab^2\rho C_p/k \tag{13}$$

$$\phi = \Delta HAC_{Mo}b^2/kT_r \tag{14}$$

equations (1) to (6) become

$$\frac{\partial\theta}{\partial\tau} = \frac{\partial^2\theta}{\partial\xi^2} + \phi\Gamma\exp(-\varepsilon/\theta) \tag{15}$$

$$\frac{\partial\Gamma}{\partial\tau} = -\beta\Gamma\exp(-\varepsilon/\theta) \tag{16}$$

with boundary conditions

$$\theta(0,\tau) = G \tag{17}$$

$$\frac{\partial\theta}{\partial\xi}(1,\tau) = 0 \tag{18}$$

and initial conditions

$$\theta(\xi,0) = \theta_o \tag{19}$$

$$\Gamma(\xi,0) = 1 \tag{20}$$

III. Solution of the System Equations

Equations (15) to (20) were solved numerically using Crank-Nicholson scheme. The iterative solution of the resulting nonlinear algebraic equations was found difficult to converge using Newton Raiphson's iterative method when the system is subjected to step-wise change in the boundary temperature and near the runaway temperature. However, a simple successive approximation in the sense of Picard's method was found to work satisfactorily and gave convergence even during reactor "runaway".

The parameters β , ε and ϕ are calculated using the following data (4).

$A = 5.68 \times 10^{6}$ sec^{-1}

$b = 4.7$ cm

$C_p = 0.43$ cal/gram °C

$E = 20330$ cal/mole

$\Delta H = 17500$ cal/mole

$k = 0.43 \times 10^{-3}$ cal/cm sec°K

$\rho = 1$ gram/ml (average of pure styrem and pure polyslyrene)

$T_r = 300$ °K

$C_{Mo} = 8.15 \times 10^{-3}$ moles/ml

giving

$$\varepsilon = 34.1 \quad , \quad \beta = 1.255 \times 10^{11} \quad , \quad \phi = 1.388 \times 10^{11}.$$

IV. Stability and Control of the Reactor

The stability of the reactor is to be investigated using a simple order of magnitude analysis with the reactor subjected to constant wall temperature. The results of the analysis are then utilized to develop feasible control methods.

A. Operation of the Reactor with Constant Wall Temperature

To investigate the dynamic nature of the system when it is subjected to a constant wall temperature, equations (15) to (20) are to be analysed with $G = \theta_c$, a constant wall temperature. The purposes are

(i) to explore the existence and the nature of the runaway temperature and

(ii) to find the wall temperature corresponding to any intial temperature such that the runaway temperature can be prevented from occurring to insure stable reactor performance.

Then, we may hopefully utilize the above information to develop fesaible methods for controlling the reactor.

A1. Estimation of the Runaway Temperture

Eqn. (15) indicates that without the source term $\phi \rho e^{-\frac{E}{\theta}}$, the system will always reach a steady-state without the occurrance of any runaway temperature. Thus we may pin-point the occurrance of runaway temperature being due to the source term. It is intuitively clear that when the source term generates sufficiently more heat than the

conductive term can transport it, a runaway temperature can be expected to occur. Thus, we may make the following order of magnitude analysis.

$$\frac{\partial\theta}{\partial\tau} = \frac{\partial^2\theta}{\partial\xi^2} + \phi\Gamma e^{-\frac{\varepsilon}{\theta}}$$

Accumulation Conduction Source (generation)

$\mathcal{O}(1)$ $\mathcal{O}(X)$

The conduction term has order one since the favorable reacting temperature is around 1.2933 and ξ is the normalized space coordinate. Therefore, when the order of the source term is much greater than one, we may expect an unstable temperature to occur. Let X be the order of the source term above which we may have a runaway temperature. Then, we can estimate the potential temperature θ_p that could cause a "runaway" by solving the following equation.

$$\phi\Gamma_p e^{-\frac{\varepsilon}{\theta_p}} = X \qquad (23)$$

where subscript p indicates values near unstable or runaway condition.

Eqn. (23) gives

$$\theta_p \simeq \frac{\varepsilon}{\log\phi\Gamma - \log X} \qquad (24)$$

Since ϕ is of $\mathcal{O}(10^{11})$, the effect of the decreasing in the concentration, Γ_p on θ_p will not be significantly felt unless Γ_p is very small. For example, recalling $\phi = 1.388\times10^{11}$

for $\Gamma_p = 1$, $\log\phi\Gamma_p = 25.64$

for $\Gamma_p = 0.1$ $\log\phi\Gamma_p = 23.33$

Thus we may take Γ_p to be one and eqn. (24) gives

$$\theta_p \cong \varepsilon / \log(\phi/X) \qquad (25)$$

To see how the magnitude of X effects θ_p, we construct the following table.

X	θ_p
1	1.33
10	1.462
20	1.505
10^2	1.625
10^3	1.82

We see that although X varies over 1 to 10^4, θ_p increases only by 0.74. Thus, even the exact value of X at which runaway phenomenon occurs is not known before hand, the above table indicates that θ_p should be above 1.33 and is most probably less than 1.82. Figure 1 favorably supports this estimates. If runaway temperature is to be defined for the time being as the temperature after which drastic temperature increase with time, occurs, then figure 1 shows that θ_b lies around 1.5, implying that tolerable magnitude of X is about 20.

A2. Runaway Phenomenon Related to the Shape of the Temperature Profile

Let us consider the phenomenon of a runaway temperature by imagining how the temperature profiles should look like in the absence of, as well as near the runaway condition.

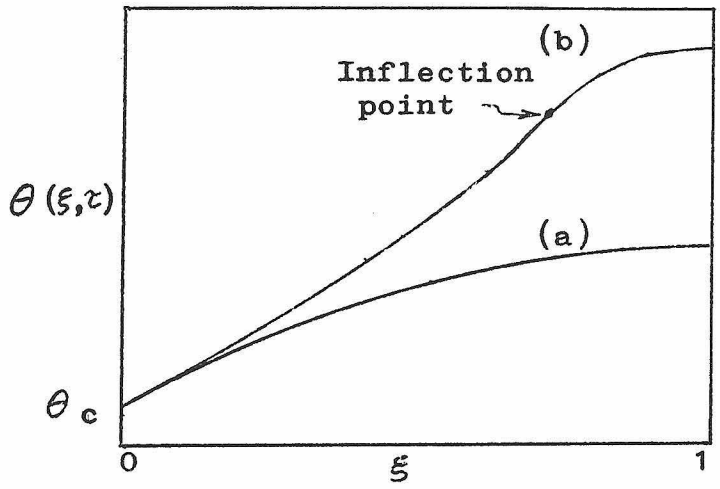


Figure 2. (a) Temperature profile without "runaway".

(b) Temperature profile near "runaway".

There shall not be any runaway phenomenon if there is no hot spot in the reactor. This can be insured if the heat flux at $\xi = \xi_1$ is at least equal to the heat flux at $\xi = \xi_2$ for any $\xi_1 < \xi_2$ after $\tau > \tau_d$

$$\text{i.e.,} \quad \left.\frac{\partial\theta}{\partial\xi}\right|_{\xi_1} \geq \left.\frac{\partial\theta}{\partial\xi}\right|_{\xi_2} \quad , \quad \text{all} \quad \xi_1 < \xi_2 \, , \; \tau > \tau_d \tag{26}$$

Condition (26) implies that for the operation of the reactor without a runaway phenomenon, well developed temperature profile should look like Figure 2 (a), representing a parabolic temperature profile.

If for $\tau > \tau_d$, at some point ξ, the reverse of (26) should happen, we can expect a runaway temperature. Near the runaway condition, the temperature profile will evolve from the shape of Figure 2 (a) to a shape such that for some point ξ_2

$$\left.\frac{\partial\theta}{\partial\xi}\right|_{\xi_1} < \left.\frac{\partial\theta}{\partial\xi}\right|_{\xi_3} \quad , \quad \xi_1 < \xi_2 < \xi_3 \tag{27}$$

Since by boundary condition (18)

$$\frac{\partial\theta(1,\tau)}{\partial\xi} = 0 \quad , \text{ for all } \tau,$$

condition (27) implies that at some point ξ'

$$\left.\frac{\partial^2\theta}{\partial\xi^2}\right|_{\xi'} = 0 \quad , \quad 0 < \xi' < 1 \tag{28}$$

The temperature profile will look like Figure 2 (b), having an inflection point.

Thus, it may be concluded that after some time τ_d beyond which the temperature profile has been well developed,

1) if condition (26) holds, there will not be a runaway temperature and a stable reactor performance can be assured of.

2) if condition (28) holds, there will be a runaway temperature, giving unstable reactor performance.

Conclusion 2) is equivalent to the order of magnitude argument leading to eqn. (24).

Numerical solutions of system equations (15) to (20) for various initial temperatures and wall temperatures agree with the above conclusions. Some sample results are shown in Figures 3 and 4. The runaway phenomenon is always preceded by the formation of inflexion points in the temperature profiles for a short period of time.

Now we may define the runaway temperatures as the value of the centre line temperature[#] where the well developed temperature profile starts to have inflection point.

With the above definition, the runaway temperature for all the computed results lies within 1.45 to 1.5

A3. Estimation of the Stable Wall Temperature

A stable behavior of the reactor can be achieved if condition (28) can be avoided near the occurrence of the runaway temperature. This can be obtained by imposing a suitable θ_c such that the overall conductive effect be at least equal to any local conductive effect especially the ones at the inflection points which have the largest temperature gradients. Therefore, from the numerical solution with runaway behavior for a given initial data and an

---

#$\theta(1,\tau)$ is chosen since it is the hottest spot in the reactor.

arbitrary wall temperature θ_{ca}, we can estimate the stable wall temperature θ_c by solving#

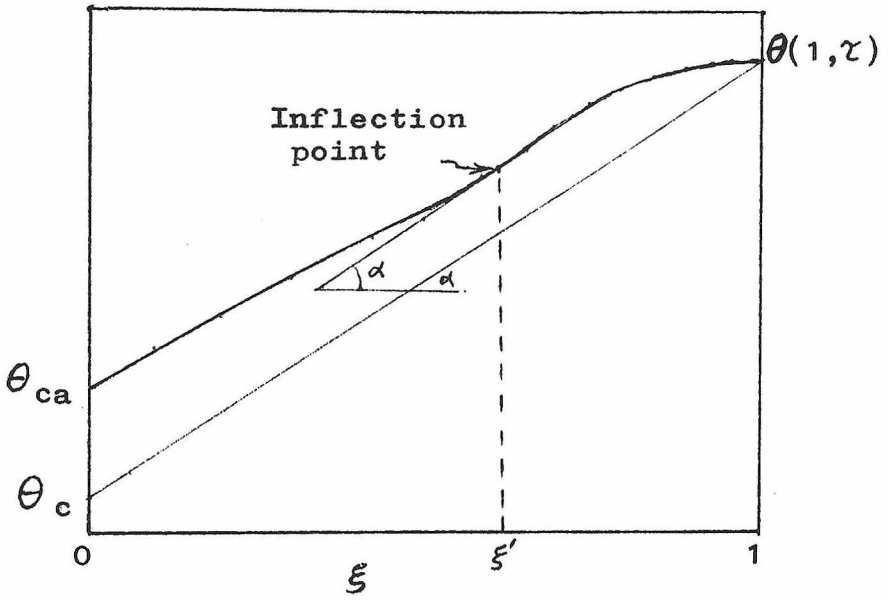


Figure 5. Estimation of stable θ_c from potential runaway numerical result.

$$\frac{\theta_p - \theta_c}{1} \geq \max_{\xi} \frac{\partial \theta}{\partial \xi} = \left.\frac{\partial \theta}{\partial \xi}\right|_{\xi'} = \alpha$$

giving

$$\theta_c \leq \theta_p - \alpha \qquad (29)$$

The graphical estimation of stable θ_c is also shown in Figure 5.

Applying the above technique to Figures 3 and 4 yield the following prediction for stable θ_c.

#. In fact from condition (26) and boundary conditions (17) and (18), it can be shown that for a stable reactor performance, $\theta_c \leq \theta_p - M$, $M \geq 0$ such that $\frac{\partial^2 \theta}{\partial \xi^2} \geq -M$.

Figure 3: $\theta_o = 1.29333$, $\theta_{ca} = 1.16667$;

Figure 4: $\theta_o = 1.25$, $\theta_{ca} = 1.2$.

Fig.	Curve	Inflection point		τ_p	θ_p	$\theta_c \leq \theta_p - \alpha$
		ξ'	α			
3	A1	0.75	0.405	0.125	1.495	1.09
3	A2	0.85	0.65	0.13	1.607	0.957
4	A1	0.75	0.382	0.44	1.512	1.13
4	A2	0.78	0.66	0.445	1.613	0.953

Direct search for a stable θ_c gave.

$\theta_o = 1.29333$ stable $\theta_c = 0.95$

$\theta_o = 1.25$ stable $\theta_c = 1.167$

The predicted values of θ_c agree roughly with the actual values. The temperature responses to these stable θ_c are shown in Figures 6 and 7.

The above technique may be used iteratively to determine a stable θ_c for any given θ_o, thus generating the relation of $\theta_c = \theta_c(\theta_o)$ for stable operation, as shown in Figure 8.

For constant wall temperature operation, there should be an optimum choice for θ_o. A higher starting temperature has to be operated with a lower wall temperatur to prevent instability and thus resulting in a poorer conversion in region near the wall.

The operating scheme with $\theta_o = 1.25, \theta_c = 1.167$ is actually better than the scheme with $\theta_o = 1.2933, \theta_c = 0.95$ as can be seen in Figures 6 and 7.

A4. Disadvantages of Operation with Constant Wall Temperature

Referring to Figures 6 and 7, stable operation of the reactor with constant wall temperature has the following defects:

1. In order to avoid the runaway behavior, the wall temperature has to be so low that the useful reacting time is only limited to the initial period. Figure 6 shows that the reactor is practically non-productive for $\tau > 0.5$.

2. There is a large temperature gradient in the reactor at all time. Therefore the reaction does not proceed uniformly enough to give satisfactory product distribution. Figure 7 shows that at $\tau = 2$, although concentration has dropped to 0.65 at $\xi = 1$ (centre of reactor), the concentration at $\xi = 0.5$ is still 0.79. The situation is even worse for the run shown in Figure 6.

In view of the above defects, better methods for operating the reactor are to be investigated.

B. Proportional Control

Proportional control through varying the wall temperature by comparing the deviation of the centre line temperature from a certain favorable centre line temperature seems promising to achieve a better reactor performance with respect to preventing an unstable temperature and for achieving a more uniform reaction throughout the reactor.

The system equations (15) to (20) are to be integrated with

$$G = \theta_c(\tau) = \theta_{co} + K(\theta^d - \theta(1,\tau)) \qquad (30)$$

where θ^d = desired temperature with respect to favorable reaction rate and stability.

K = proportional gain.

For a given initial reactor temperature θ_o, stable values for K and θ^d have to be tried out. The stable θ_c given in Figure 8 for constant wall temperature case may be used for θ_{co}.

Since practical limitation imposes constraint on the maximum and minimum values for $\theta_c(\tau)$, we have

$$*\theta_c \leq \theta_c(\tau) \leq \theta_c^* < \theta_p \qquad (31)$$

because θ_c^* should be well below θ_p. $*\theta_c$ may usually be the available cooling water temperature which is approximately 1 in dimensionless temperature scale. Hence, $*\theta_c$ and θ_c^* do not differ by a large range and thus it makes the choice for

K not critical. If K is too large, the proportional control will approximate bang-bang control.

Thus the critical parameter for a stable operation is a correct choice for θ^d , since the system is very sluggish, a large θ^d will tend to blow up the system before the effect of control through cooling at the wall reaches the centre line of the reactor. θ^d should be searched around 1.29333, the desired reaction temperature of the system.

For θ_o = 1.25, the stable value for θ^d =1.27. The computed results are shown in Figure 9.

The results shown in Figure 9 have the following improvement over those shown in Figure 7.

1) More uniform temperature and faster reaction. For comparison, at τ = 2, the concentrations and temperature at selected values of ξ are given below for θ_o = 1.25

	Constant wall temperature		Proportional control	
	Γ	θ	Γ	θ
ξ = 0.1	0.928	1.174	0.8835	1.201
ξ = 0.5	0.7908	1.196	0.7354	1.248
ξ = 1.0	0.6523	1.207	0.5888	1.267

2) Furthermore at τ = 2, the reaction is still going on in the case of using proportional control, whereas the reaction using constant wall temperature is almost at a standstill.

C. Feedforward Control

Although the feedback proportional control has advantage over the operation with constant wall temperature, it can be seen in Figure 9 that almost during half of the time, the reactor is at a lower temperature not favorable for reaction. It, of course, requires control mechanism making the equipment cost higher than the operation with constant wall temperature. Thus if a suitable policy for the wall temperature can be prescribed and applied in a feed-forward fashion, we may expect a better reactor performance without incurring additional cost for control mechanism.

We shall make a computer search for a stable wall temperature policy in the form of a step function of time which if necessary may be generated into a smooth function.

The technique for finding a stable constant wall temperature mentioned in section IV. A3 may be used to develop a wall temperature policy in combination with computed numerical results as outlined below:

Step 1. For any given θ_o, use the corresponding stable θ_{c1} from Figure 8. From the centre-line temperature response of the numerical result, locate the time τ_1 at which the temperature starts to drop below a certain desired value (e.g. the favorable reaction temperature).

Step 2. Increase the wall temperature to a new θ_{c2} (e.g. the optimum starting temperature) to reheat the reacting

mass to decrease the large temperature gradient created in step 1. In so doing, the centre line temperature will start to increase again to a desired temperature at time τ_2.

Step 3. At τ_2, decrease the wall temperature to θ_{c3} the stable θ_c for the $\theta_o = \theta_{c2}$. Locate the time τ_3, at which the increasing centre-line temperature starts to drop below a desired level.

Step 4. Repeat step 2 with wall temperature θ_{c4} and repeat the procedure to obtain

$$\theta_c = \theta_{c1} \quad , \quad 0 \leq \tau \leq \tau_1$$

$$= \theta_{c2} \quad , \quad \tau_1 < \tau \leq \tau_2$$

$$= \theta_{c3} \quad , \quad \tau_2 < \tau \leq \tau_3$$

$$\cdots \qquad \cdots$$

$$= \theta_{cn} \quad , \quad \tau_{n-1} < \tau < \tau_n$$

The result of the feedforward control of the system treated in Figure 6 is shown in Figure 10.

The temperature and concentration profiles at $\tau = 2$ are

ξ	0.1	0.5	1.0
θ	1.285	1.321	1.324
Γ	0.801	0.582	0.170

which are better than any of the previous cases.

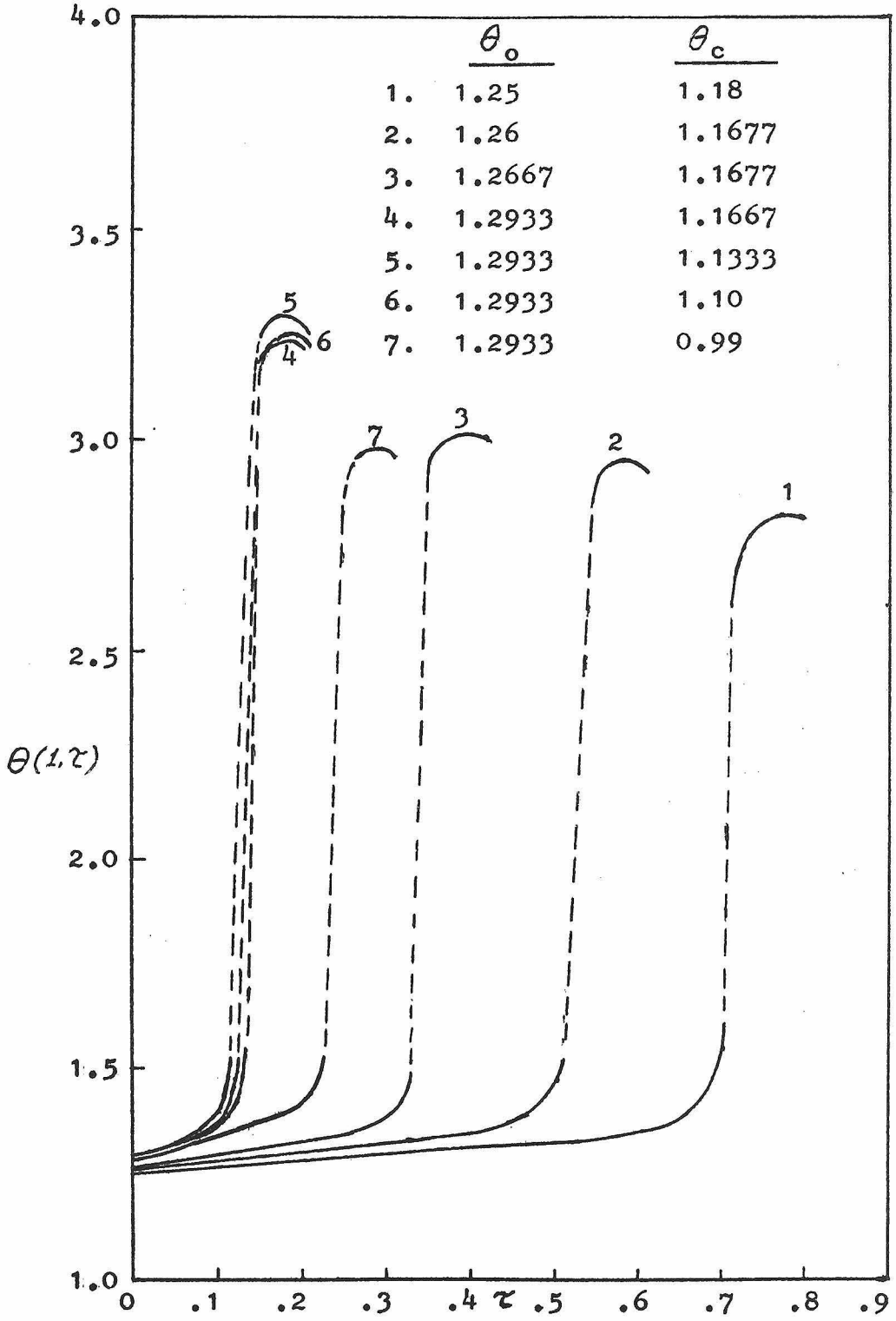


Figure 1. Occurrence of reactor "runaways" within a close temperature range.

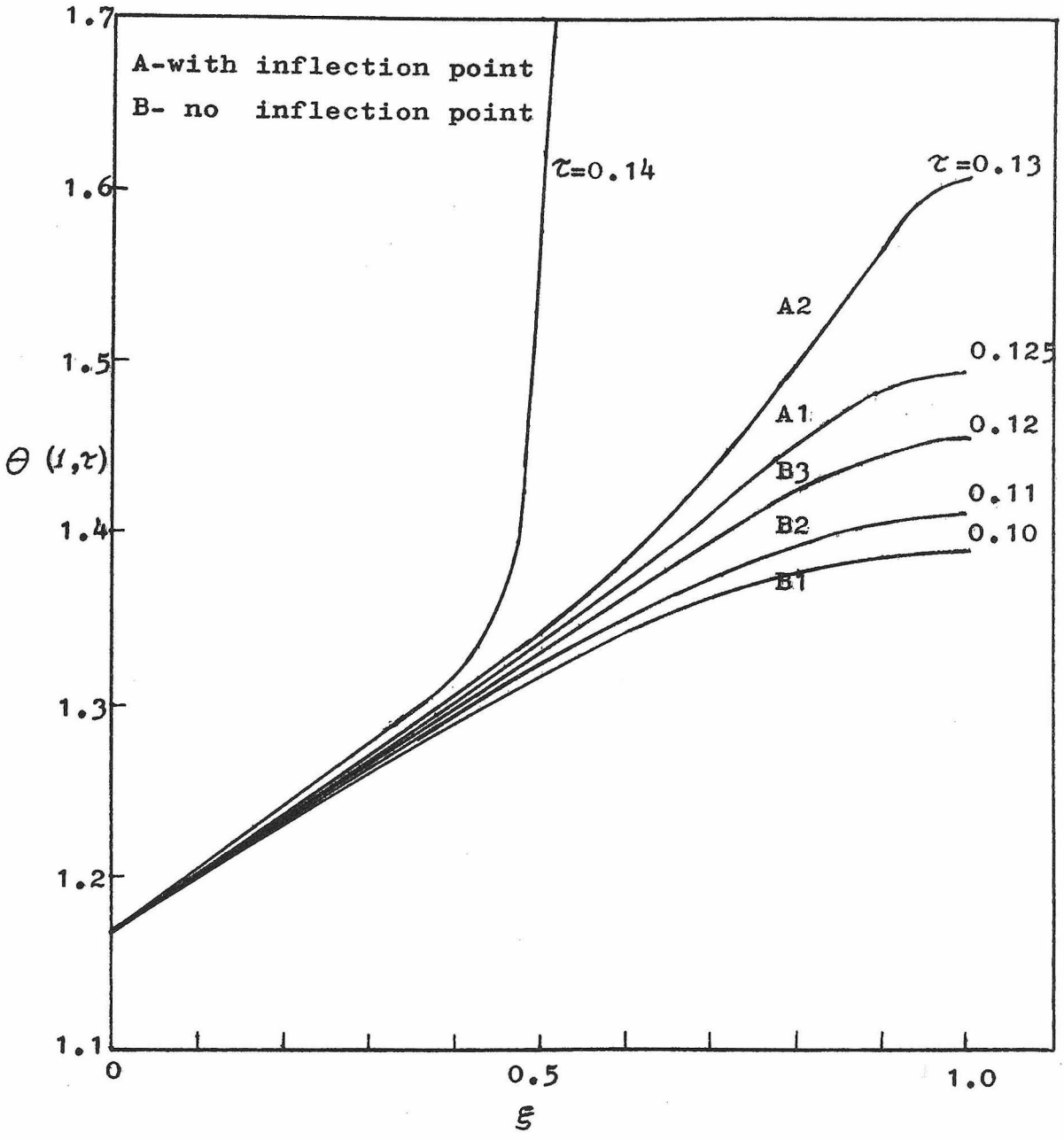


Figure 3. Evolution of temperature profiles leading to a runaway temperature with θ_o=1.29333, θ_c=1.1667.

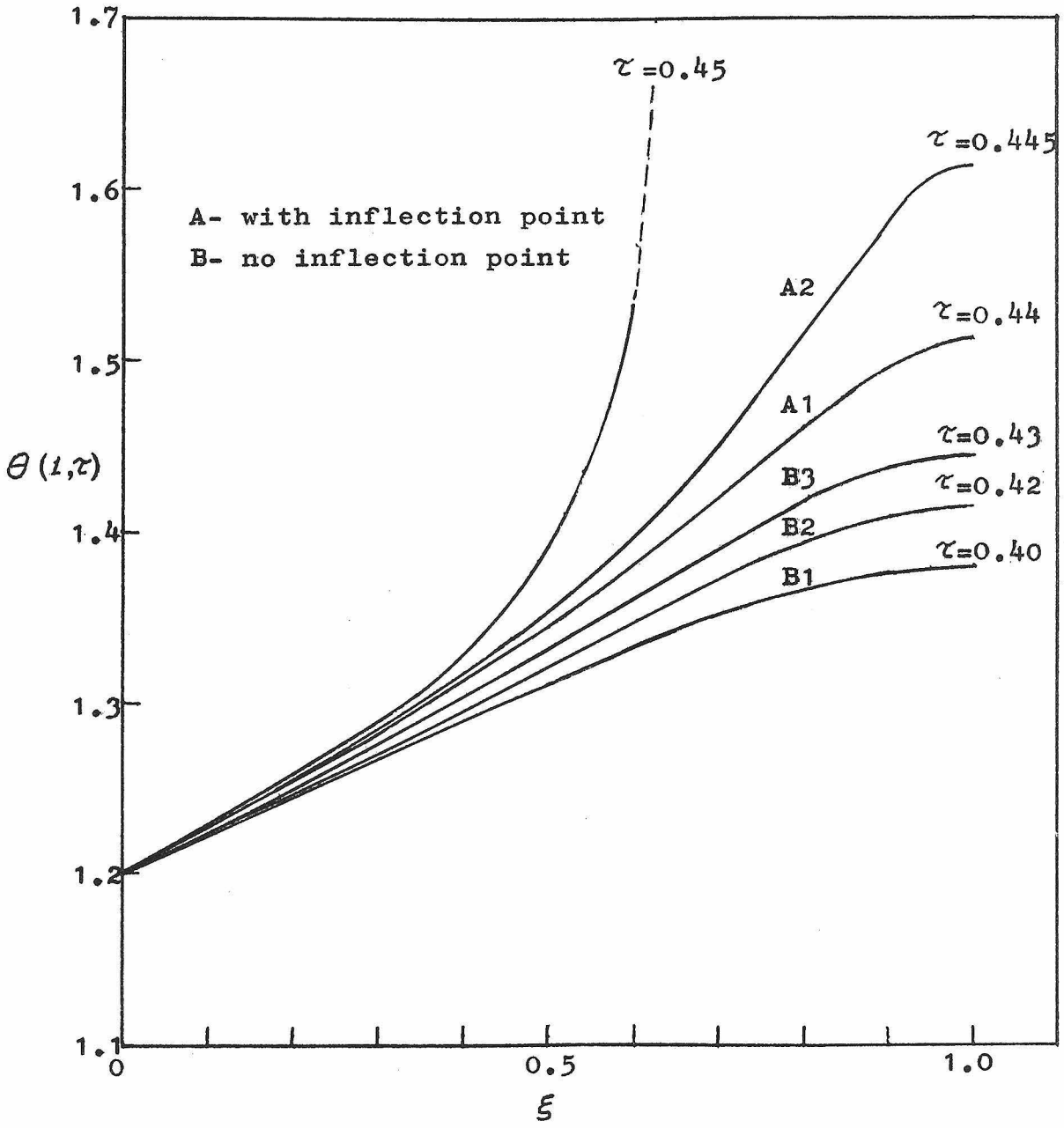


Figure 4. Evolution of temperature profiles leading to a runaway temperature with $\theta_o=1.25, \theta_c=12$

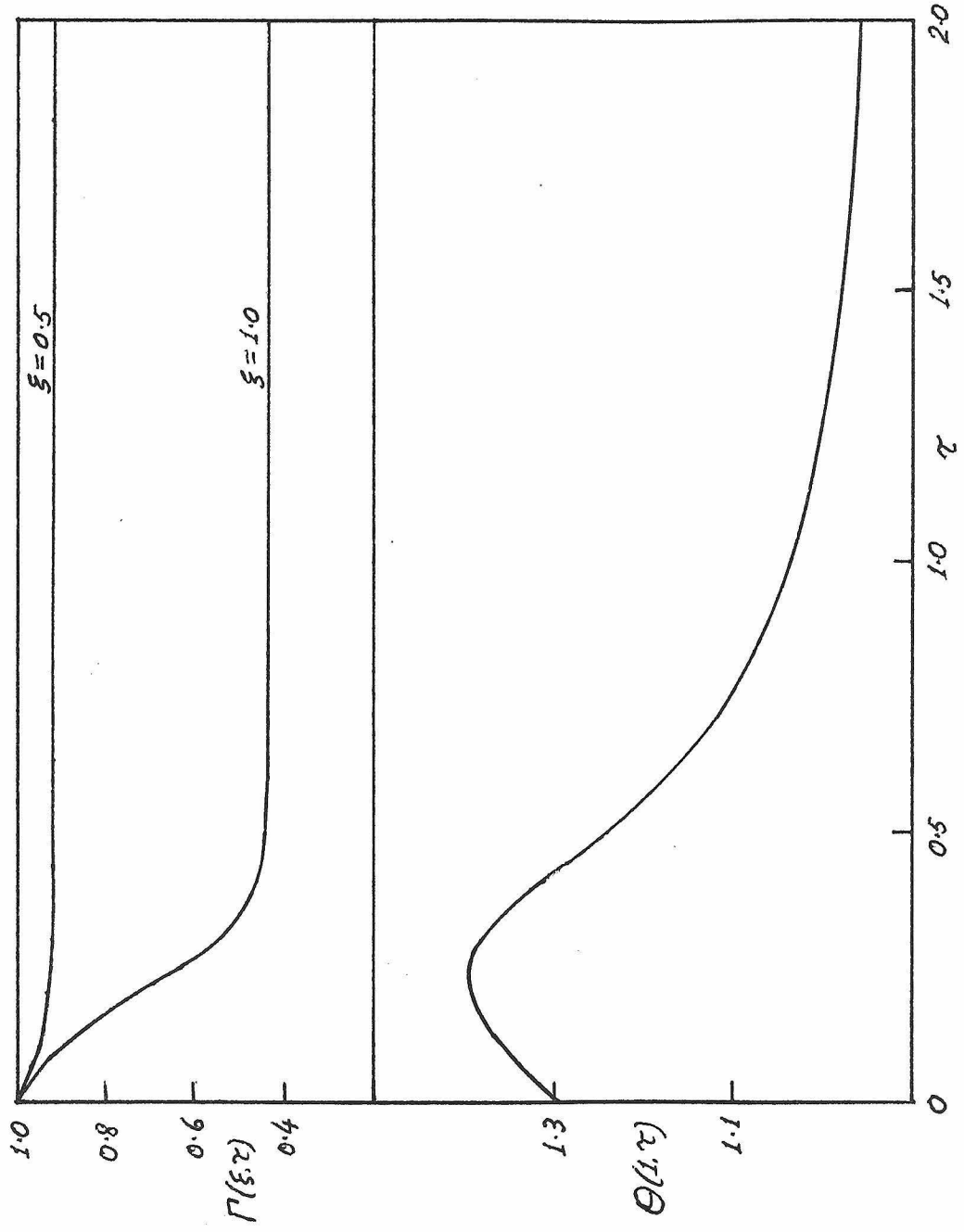


Figure 6. Stable temperature response with $\Theta_0 = 1.29333$, $\Theta_e = 0.95$.

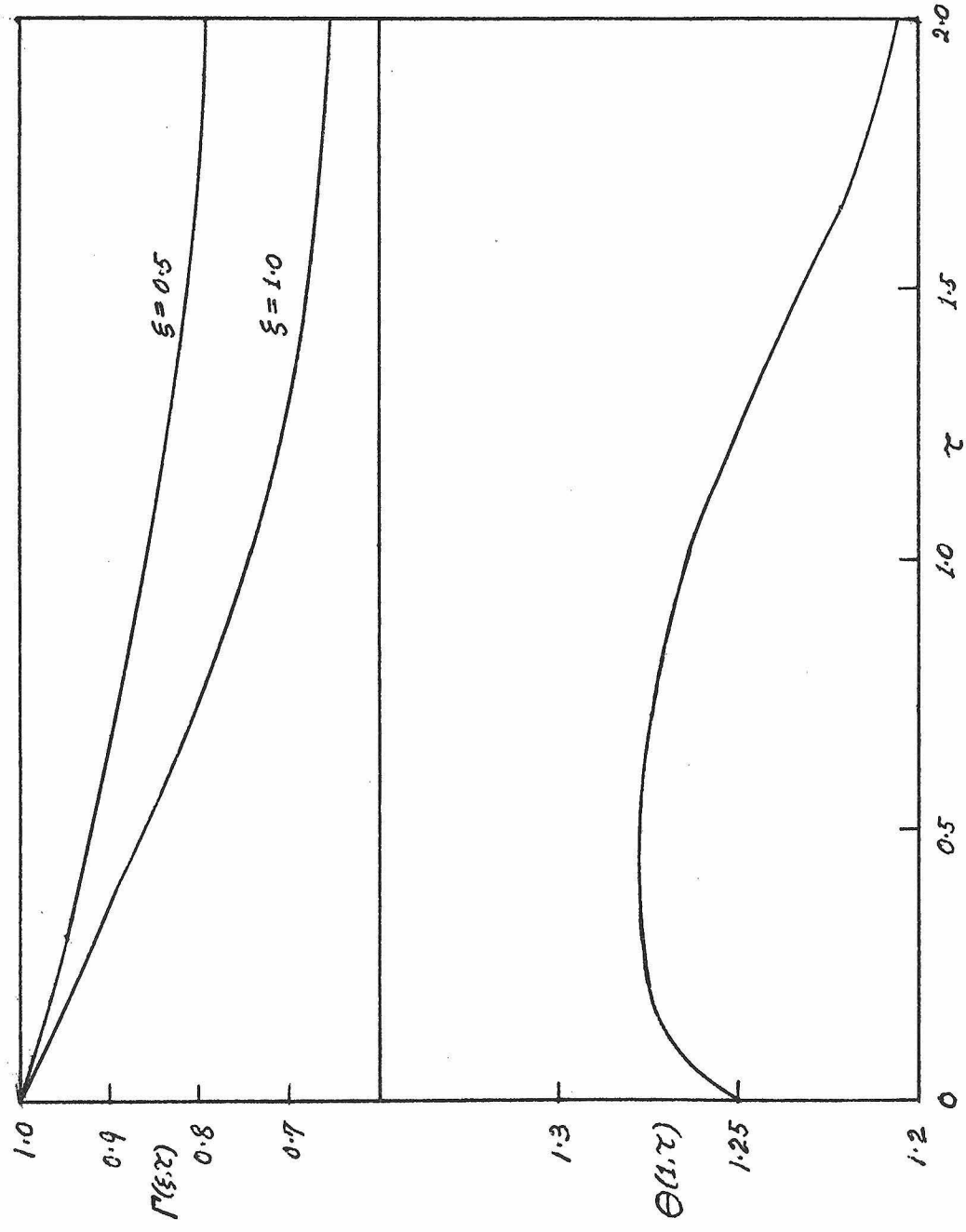


Fig. 7. Stable temperature response with $\Theta_0 = 1.25$, $\Theta_c = 1.1667$

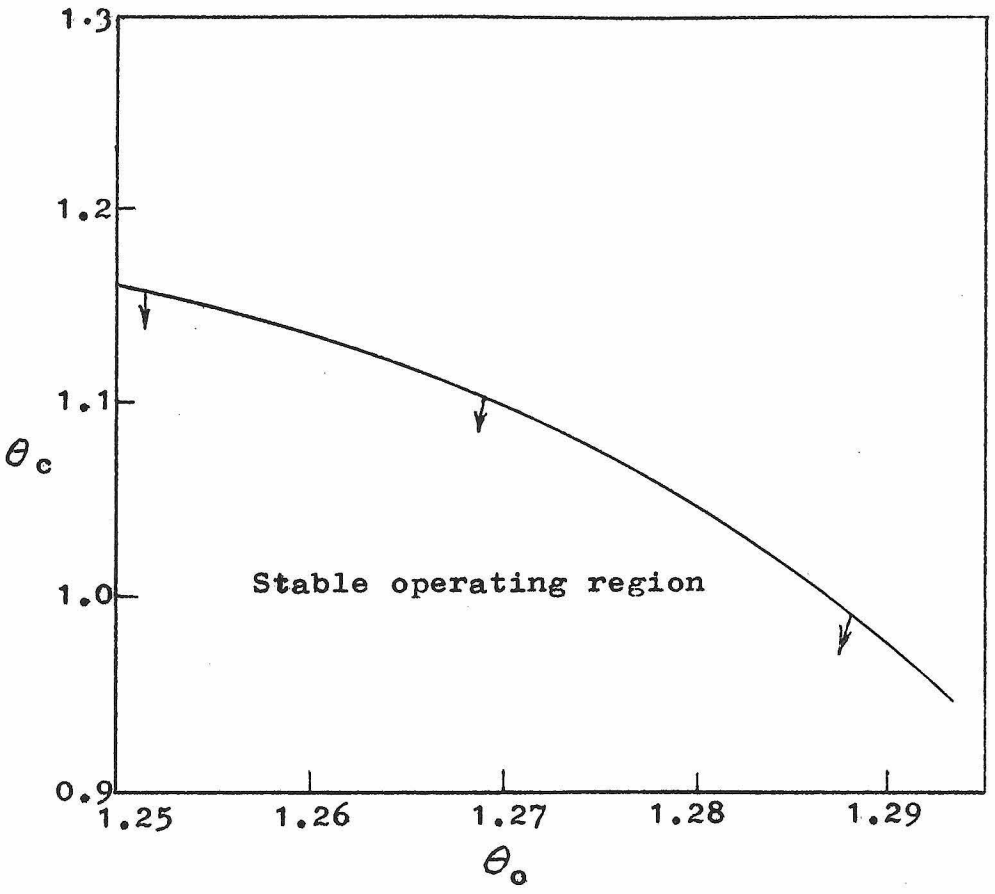


Figure 8. Stable θ_c for constant wall temperature operation with initial temp. θ_o.

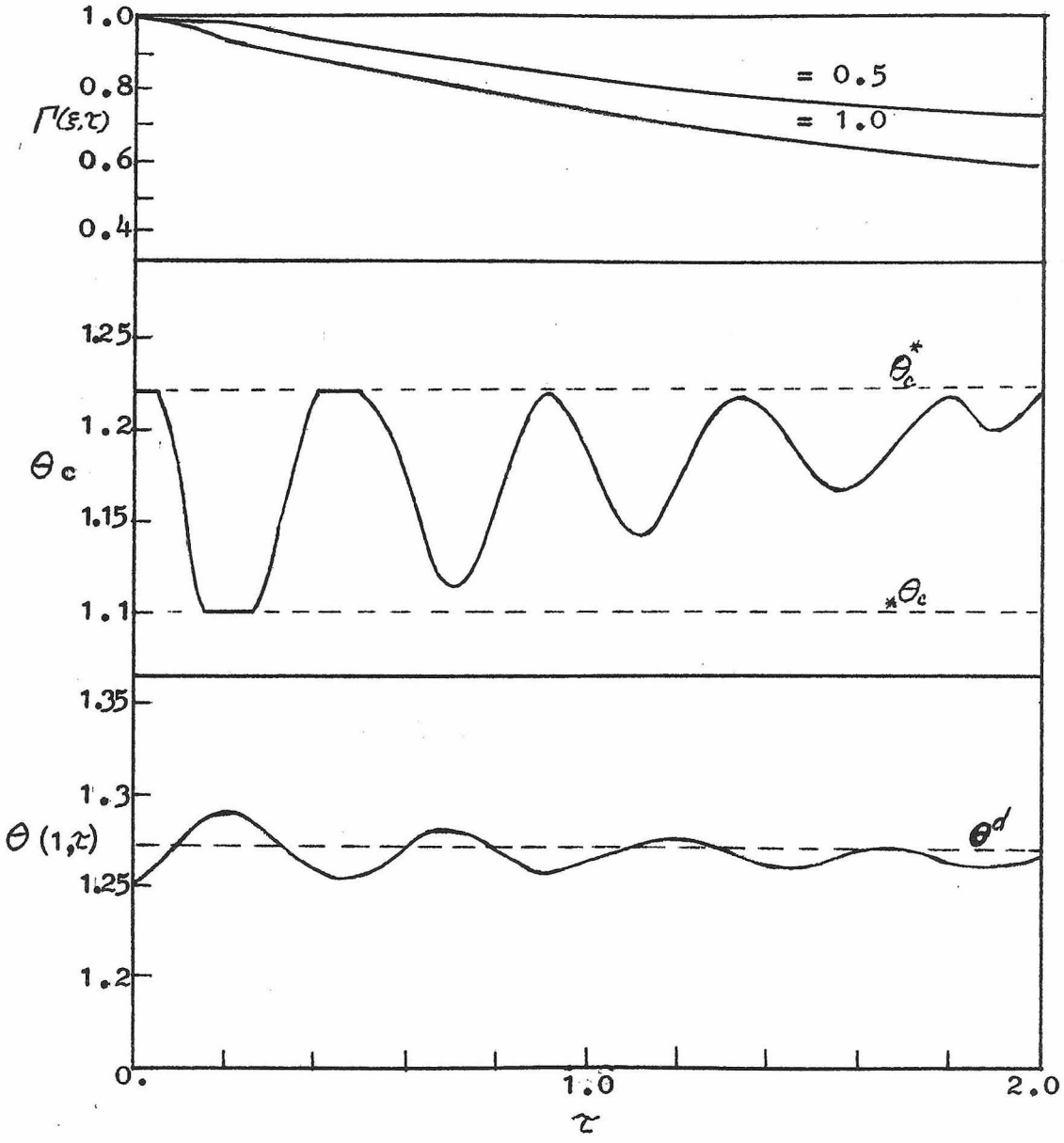


Figure 9. Proportional control of the reactor with $\theta_0 = 1.25$, $K = 0.5$

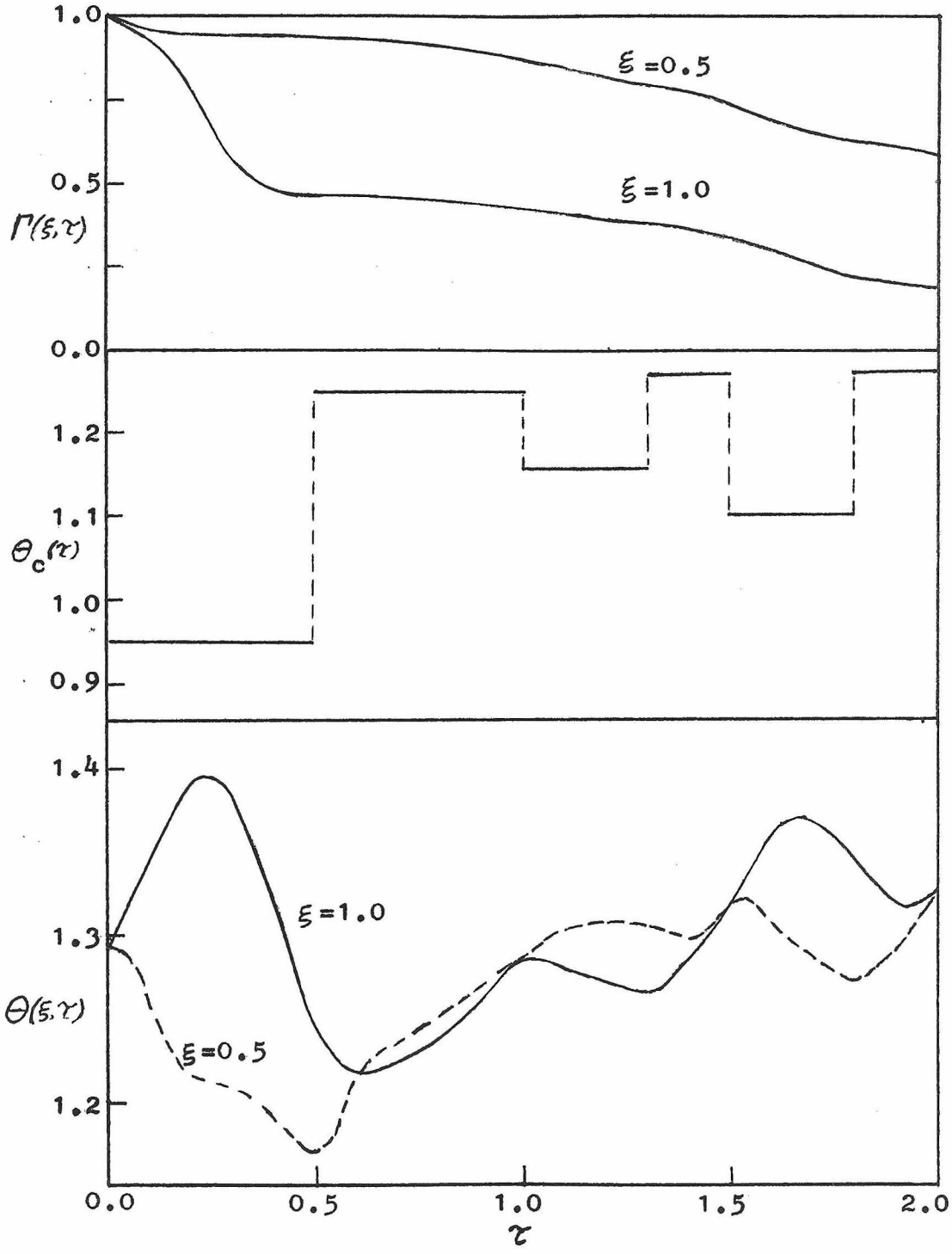


Figure 10. Feed-forward control with θ_o= 1.29333.

V. Conclusions

A batch polystyrene reactor was analysed using order of magnitude analysis in combination with computed numerical results. The following conclusions may be drawn:

1. The runaway temperature of the reactor occurs at about $\theta_p = 1.45$ which may be regarded as the characteristics of the system and is independent of the operating conditions.

2. Order of magnitude analysis of the governing system equations leads to simple formulae for the estimation of the reactor runaway temperature and the stable operating conditions.

3. A stable reactor performance is characterized by the temperature profiles always being parabolic, whereas an unstable reactor performance is always preceded by the formation of inflection point in the temperature profile.

4. Feasible feed-back and feed-forward control schemes were developed for a better operation of the reactor with stable performance.

Nomenclature

A	Frequency factor.
b	Half thickness of the batch reactor, cm.
C_M	Concentration of monomer, mole/ml.
C_p	Heat capacity, cal/gm °C.
E	Activation energy, cal/mole.
g	Wall temperature function.
G	Dimensionless g as defined by eqn.(10).
H	Heat of reaction, cal/mole.
k	Thermal conductivity, cal/cm sec °C.
K	Proportional gain.
M	Monomer.
R	Universal gas constant.
t	Time in hour.
T	Absolute temperature, °K .
Td	Desired temperature.
Tr	Referance temperature, °K .
X	Order of magnitude of heat generation to cause unstable temperature.

Greek Letters

α	Slope of temperature profile at inflection point.
β	Dimensionless parameter, $Ab^2\rho C_p/k$
θ	Dimensionless temperature, T/T_r.
θ_c	Dimensionless wall temperature.
$*\theta_c$	Minimum feasible θ_c.

θ_c^*	Maximum feasible θ_c.
θ_{ca}	Arbitrary θ_c.
θ^d	Dimensionless desired temperature.
ε	Dimensionless activation energy, E/RT_r .
ξ	Dimensionless space coordinate x/b.
Γ	Dimensionless concentration, CM/C_{Mo}.
τ	Dimensionless time coordinate, $kt/\rho C_p b^2$.
τ_d	Dimensionless time after which temperature is well developed.
ϕ	Dimensionless parameter, $\Delta HAC_{Mo}b^2/kT_r$.

Subscript

p	Indicates value near runaway.
o	Indicates initial value.

Reference

1. Boundy,R.H. and R.F. Boyer, Styrene:Its Polymers, Copolymers and Derivatives, American Chemical Society Monograph Series, Hafner, New York, 1961.

2. Flory, P.J., Polymer Chemistry, Cornell University Press, Ithaca, 1953.

3. Gee, J.C., "Kinetics of the Polymerization of Styrene", Amoco Chemicals Memorandum 4450.8,October 1, 1969.

4. Seinfeld, J.H.,"Batch Production of Impact Polystyrene", December 15, 1969.

Proposition II

Numerical Evaluation of the Transition Matrix of a Time-varying System

CHAPTER 1. INTRODUCTION

In solving a dynamical system described by a set of linear ordinary differential equations, it is important to be able to evaluate effeciently the transition matrix associated with the system. The evaluation of the transition matrix for a time-invariant system can be easily done in many ways (Koppel, 1968). However, for a time-varying system, it is not always easy to evaluate its transition matrix, especially when the dimension of the system is large, the coefficient matrix ill-conditioning or only known at a discrete set of time instants. It is specially for such cases that a computational scheme for the evaluation of the transition matrix of a time-varying system is developed in this proposition.

1.1 A brief review of the solution of a lumped-parameter dynamical system

A dynamical system described by a set of linear ordinary differential equations such as

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \tag{1.1}$$

$$x(t_o) = x^o, \text{ known} \tag{1.2}$$

where $x(t)$ = a n-dimensional state vector

$u(t)$ = a p-dimensional control vector

$A(t)$ = a nxn time-varying coefficient matrix

$B(t)$ = a nxp matrix

t = time variable with t_o being initial time

has the following solution (Koppel, 1968):

$$\dot{x}(t) = \Phi(t,t_o)x(t_o) + \int_v^t \Phi(t,v)B(v)u(v)dv \qquad (1.3)$$

where $\Phi(t,t_0)$ is given by

$$\Phi(t,t_o) = A(t)\Phi(t,t_o) \qquad (1.4)$$

$$\Phi(t_o,t_o) = I \qquad (1.5)$$

It can also be shown (Amundson,1966) that $\Phi(t,t_o)$ is given by the following matrizant:

$$\Phi(t,t_o) = I + \int_{t_o}^{t} A(v_1)dv_1 + \int_{t_o}^{t} A(v_1)dv_1 \int_{t_o}^{v_1} A(v_2)dv_2 +$$

$$\int_{t_o}^{t} A(v_1)dv_1 \int_{t_o}^{v_1} A(v_2)dv_2 \int_{t_o}^{v_2} A(v_3)dv_3 + \cdots\cdots\cdots \qquad (1.6)$$

Thus, when the dimension of A(t) is large, ill-conditioned or only known at a discrete set of time instants, it can be seen that the evaluation of $\Phi(t,v)$, $t_o \leq v < t$ by the integration of (1.4) is no longer convenient and may even be very difficult. In the next chapter, we shall see that this difficulty can be overcome by direct numerical integrations of the matrizant (1.6).

CHAPTER 2. NUMERICAL EVALUATION OF MATRIZANT

As noted in the last chapter,

$$\Phi(t,t_o) = I + \int_{t_o}^{t} A(v_1)dv_1 + \int_{t_o}^{t} A(v_1)\int_{t_o}^{v_1} A(v_2)dv_2dv_1 +$$

$$\int_{t_o}^{t} A(v_1)\int_{t_o}^{v_1} A(v_2)\int_{t_o}^{v_2} A(v_3)dv_3dv_2dv_1 + \cdots\cdots \qquad (2.1)$$

We shall evaluate the integrals in (2.1) term by term, by a combination of Simpson' and trapezoidal rules (Lapidus,1962). We shall assume that $A(v)$ are known for an even-interval discrete set of time instants $v = t_o, t_1, \ldots, t_m$.

Simpson's rule is

$$\int_{t_o}^{t} A(v)dv = (h/3)\left\{A(t_o) + 4\left[A(t_1) + A(t_3) + \cdots + A(t_{m-1})\right]\right.$$
$$\left. + 2\left[A(t_2) + A(t_4) + \cdots + A(t_{m-2})\right] + A(t_m)\right\}$$
$$- O(h^5) \qquad (2.2)$$

where m = an even integar

$h = (t-t_o)/m$

$t_i = t_o + ih$, $\quad i = 1,\ldots,m$

$t_m = t$

Trapezoidal rule is

$$\int_{t_o}^{t_i} A(v)dv = h'\left[A(t_o)/2 + A(t_1) + \cdots + A(t_{n-1}) + A(t_n)/2\right] - O(h'^3) \qquad (2.3)$$

where $h' = (t_i - t_o)/n$, n = an even or odd integer

$t = t_o + jh'$, $j = 1,\ldots,n$

$t_n = t_i$

The general numerical scheme is as follows:

1. For simple integral or the outer integral of a composite integral, Simpson's rule is to be used.

2. For the inner integrals of a composite integral, trapezoidal rule is to be used.

Schematically, the above numerical scheme is the following:

$$\underbrace{\int_{t_o}^{t} A(v_1)dv_1}_{\text{Simpson's rule}}$$

$$\underbrace{\int_{t_o}^{t} A(v_1)\underbrace{\int_{t_o}^{v_1} A(v_2)\int_{t_o}^{v_2} A(v_3)dv_3dv_2}_{\text{Trapezoidal rule}}\,dv_1}_{\text{Simpson's rule}}$$

There is no specific reason for the choice of Simpson's rule except for its accuracy and common use. Trapezoidal rule is chosen for its simplicity and the flexibility of choosing n which may be even or odd without any effect on the accuracy. As will be seen later, the numerical scheme as mentioned above can be used with any type of integration formulae.

2.2 Derivation of the numerical scheme

We shall derive recursive relationships for the evaluation of each term of (2.1). Then, these are combined to give a final compact relation for the evaluation of $\Phi(t,t_o)$.

Applying Simpson's rule (2.2),

$$\int_{t_o}^{t} A(v_1)dv_1 = (t-t_o/3m)\Big\{A(t_o) + 4\Big[A(t_1) + A(t_1) + \ldots + A(t_{m-1})\Big] + 2\Big[A(t_2) + A(t_4) + \ldots + A(t_{m-2})\Big] + A(t_m)\Big\} \tag{2.4}$$

where $t_i = t_o + (t-t_o)i/m$, $\quad i = 0, 1, 2, \ldots, m$

$t_m = t$

Applying Simpson's rule (2.2),

$$\int_{t_o}^{t} A(v_1)\int_{t_o}^{v_1} A(v_2)dv_2dv_1 = (t-t_o)/3m\Big\{A(t_o)\int_{t_o}^{t_o} A(v_2)dv_2 + 4\Big[A(t_1)\int_{t_o}^{t_1} A(v_2)dv_2 + A(t_3)\int_{t_o}^{t_3} A(v_2)dv_2 + \ldots + A(t_{m-1})\int_{t_o}^{t_{m-1}} A(v_2)dv_2\Big] + 2\Big[A(t_2)\int_{t_o}^{t_2} A(v_2)dv_2 + A(t_4)\int_{t_o}^{t_4} A(v_2)dv_2 + \ldots + A(t_{m-2})\int_{t_o}^{t_{m-2}} A(v_2)dv_2\Big] + A(t_m)\int_{t_o}^{t_m} A(v_2)dv_2\Big\} \tag{2.5}$$

Applying trapezoidal rule (2.3) to each of the integrals

$\int_{t_o}^{t_i} A(v_2)dv_2$ in (2.5),

$$\int_{t_o}^{t_i} A(v_2)dv_2 = (t_i-t_o)/i\left[A(t_o)/2 + A(t_1) +\ldots+ A(t_{i-1}) + A(t_i)/2\right] = K_1(t_i, t_o) \qquad (2.6)$$

$i = 1, 2,\ldots, m.$

(2.5) and (2.6) give

$$\int_{t_o}^{t} A(v_1) \int_{t_o}^{v1} A(v_2)dv_2dv_1 = (t-t_o)/3m\left[4\left\{A(t_1)K_1(t_1, t_o) + A(t_3)K_1(t_3, t_o) +\ldots+ A(t_{m-1})K_1(t_{m-1}, t_o)\right\} + 2\left\{A(t_2)K_1(t_2,t_o) + A(t_4)K_1(t_4,t_o) +\ldots+ A(t_{m-2})K_1(t_{m-2},t_o)\right\}+ A(t_m)K\ (t_m,t_o)\right] \qquad (2.7)$$

Similarly, applying Simpson's rule:

$$\int_{t_o}^{t} A(v_1) \int_{t_o}^{v_1} A(v_2) \int_{t_o}^{v_2} A(v_3)dv_3dv_2dv_1 = (t-t_o)/3m\left\{A(t_o) \int_{t_o}^{t_o} A(v_2) \int_{t_o}^{v2} A(v_3)dv_2 + 4\left[A(t_1) \int_{t_o}^{t_1} A(v_2) \int_{t_o}^{v_2} A(v_3)dv_3dv_2 + A(t_3) \int_{t_o}^{t_3} A(v_2) \int_{t_o}^{v_2} A(v_3)dv_3dv_2+\ldots+ A(t_{m-1}) \int_{t_o}^{t_{m-1}} A(v_2) \int_{t_o}^{v_2} A(v_3)dv_3dv_2\right]+\right.$$

$$2\left[A(t_2)\int_{t_o}^{t_2}A(v_2)\int_{t_o}^{v_2}A(v_3)dv_3dv_2+\ldots+ \right. \tag{2.8}$$

$$\left. A(t_{m-2})\int_{t_o}^{t_{m-1}}A(v_2)\int_{t_o}^{v2}A(v_3)dv_3dv_2\right]+A(t_m)\int_{t_o}^{t_m}A(v_2)\int_{t_o}^{v2}A(v_3)dv_3dv_2\Bigg\}$$

Trapezoidal rule on the $\int_{t_o}^{t_i}A(v_2)\int_{t_o}^{v_2}A(v_3)dv_3dv_2$ gives

$$\int_{t_o}^{t_i}A(v_2)\int_{t_o}^{v_2}A(v_3)dv_3dv_2= (t_i-t_o)/i\left[\left(A(t_o)/2\right)\int_{t_o}^{t_o}A(v_3)dv_3 + \right.$$

$$A(t_1)\int_{t_o}^{t_1}A(v_3)dv_3 + \ldots+ A(t_{i-1})\int_{t_o}^{t_{i-1}}A(v_3)dv_3 + \tfrac{1}{2}A(t_i)$$

$$\left.\int_{t_o}^{t_i}A(v_3)dv_3\right]$$

Inview of (2.6), the above expression becomes

$$\int_{t_o}^{t_i}A(v_2)\int_{t_o}^{v_2}A(v_3)dv_3dv_2 = (t_i-t_o)/i\left[A(t_1)K_1(t_1,t_o) + \right.$$

$$\left.A(t_2)K_1(t_2,t_o) + \ldots+ A(t_{i-1})K_1(t_{i-1},t_o) + \tfrac{1}{2}A(t_i)K_1(t_i,t_o)\right]$$

$$= K_2(t_i,t_o)\ ,\quad i = 1, 2,\ldots, m. \tag{2.9}$$

or

$$K_2(t_i,t_o) = (t_i-t_o)/i\left[A(t_1)K_1(t_1,t_o) + A(t_2)K_1(t_2,t_o) +\ldots\right.$$

$$\left.+ A(t_{i-1})K_1(t_{i-1},t_o) + A(t_i)/2\ K_1(t_i,t_o)\right] \tag{2.10}$$

(2.8) and (2.10) give

$$\int_{t_o}^{t} A(v_1) \int_{t_o}^{v_1} A(v_2) \int_{t_o}^{v_2} A(v_3) dv_3 dv_2 dv_1 = \left((t-t_o)/3m\right)\Big\{4\Big[A(t_1) K_2(t_1,t_o) + A(t_3)K_2(t_3,t_o) + \ldots + A(t_{m-1})K_2(t_{m-1},t_o)\Big] + 2\Big[A(t_2)K_2(t_2,t_o) + \ldots + A(t_{m-2})K_2(t_{m-2},t_o)\Big] + A(t_m)K_2(t_m, t_o)\Big\} \tag{2.11}$$

with K_2 given by (2.10) recursively.

Generalizing, we have

$$\int_{t_o}^{t} A(v_1) \int_{t_o}^{v_1} A(v_2) \int_{t_o}^{v_2} A(v_3) \int_{t_o}^{v_{n-1}} A(v_n) dv_n dv_{n-1}, \ldots dv_1$$

$$= \left((t-t\ )/3m\right)\Big\{4\Big[A(t_1)K_{n-1}(t_1,t_o) + A(t_3)K_{n-1}(t_3,t_o) + \ldots + A(t_{n-1})K_{n-1}(t_{n-1},t_o)\Big] + 2\Big[A(t_2)K_{n-1}(t_2,t_o) + A(t_4)K_{n-1}(t_4,t_o) + \ldots + A(t_{m-2})K_{n-1}(t_{m-2},t_o)\Big] + A(t_m)K_{n-1}(t_m,t_o)\Big\} \tag{2.12}$$

Where $K_n(t_i,t_o)$ are given by the recursive relationship,

$$K_n(t_i,t_o) = \left((t_i-t_o)/i\right)\Big[A(t_1)K_{n-1}(t_i,t_o) + A(t_2)K_{n-1}(t_2,t_o) + \ldots + A(t_{i-1})K_{n-1}(t_{i-1},t_o) + A(t_i)/2\ \ K_{n-1}(t_i,t_o)\Big]$$

$$n = 2, 3, \ldots \tag{2.13}$$

with $K_1(t_i,t_o) = \left((t_i-t_o)/i\right)\left[A(t_o)/2 + A(t_1) + \ldots + A(t_{i-1}) + A(t_i)/2\right]$, $i = 1, 2, \ldots, m.$ (2.14)

Substituting (2.4) and (2.12) to each terms on the right of equation (2.1) gives

$$\Phi(t,t_o) = 1 + \left((t-t_o)/3m\right)\left\{A(t_o) + 4\left[A(t_1) + A(t_3) + \ldots + A(t_{m-1})\right] + 2\left[A(t_2) + A(t_4) + \ldots + A(t_{m-2})\right] + A(t_m)\right\} +$$

$$\left((t-t_o)/3m\right)\left\{4\left[A(t_1)K_1(t_1,t_o) + A(t_3)K_1(t_3,t_o) + \ldots + A(t_{m-1})K_1(t_{m-1},t_o)\right] + 2\left[A(t_2)K_1(t_2,t_o) + A(t_4)K_1(t_4,t_o) + \ldots + A(t_{m-2})K_1(t_{m-2},t_o)\right] + A(t_m)K_1(t_m,t_o)\right\} + \left((t-t_o)/3m\right)\cdot$$

$$\left\{4\left[A(t_1)K_2(t_1,t_0) + A(t_3)K_2(t_3,t_o) + \ldots + A(t_{m-1})K_2(t_{m-1},t_o)\right] + 2\left[A(t_2)K_2(t_2,t_o) + \ldots + A(t_{m-2})K_2(t_{m-2},t_o)\right] + A(t_m)K_2(t_m,t_o)\right\} + \ldots + (t-t_o)/3m\ \left\{4\left[A(t_1)K_{n-1}(t_1,t_o) + A(t_3)K_{n-1}(t_2,t_o) + \ldots + A(t_{m-1})K_{n-1}(t_{m-1},t_o)\right] + 2\left[A(t_2)K_{n-1}(t_2,t_o) + \ldots + A(t_{m-2})K_{n-1}(t_{m-2},t_o)\right] + A(t_m)K_{n-1}(t_m,t_o)\right\} + \ldots\ldots$$ (2.15)

Collecting the coefficients of $A(t_i)$, $i=o, 1, \ldots, m$,

$$\Phi(t,t_o) = 1 + \left((t-t_o)/3m\right)\left\{A(t_o) + 4\left[A(t_1)\left(1 + K_1(t_1,t_o)\right.\right.\right.$$

$$+ K_2(t_1,t_o) + \ldots + K_{n-1}(t_1,t_o) + \ldots \Big) + A(t_3) \Big(1 + K_1(t_3,t_o) + K_2(t_3,t_o) + \ldots + K_{n-1}(t_3,t_o) + \ldots \Big) + \ldots + A(t_{m-1}) \Big(1 + K_1(t_{m-1},t_o) + K_2(t_{m-1},t_o) + \ldots + K_{n-1}(t_{m-1},t_o) + \ldots \Big) \Big]$$

$$2 \Big[A(t_2) \Big(1 + K_1(t_2,t_o) + K_2(t_2,t_o) + \ldots + K_{n-1}(t_2,t_o) + \ldots \Big) + A(t_4) \Big(1 + K_1(t_4,t_o) + K_2(t_4,t_o) + \ldots + K_{n-1}(t_4,t_o) + \ldots \Big) + \ldots + A(t_{m-2}) \Big(1 + K_1(t_{m-2},t_o) + K_2(t_{m-2},t_o) + \ldots + K_{n-1}(t_{m-2},t_o) + \ldots \Big) \Big] + A(t_m) \Big(1 + K_1(t_m,t_o) + K_2(t_m,t_o) + \ldots + K_{n-1}(t_m,t_o) + \ldots \Big) \Big\} \tag{2.16}$$

or

$$\Phi(t, t_o) = 1 + \Big((t-t_o)/3m \Big) \Big\{ A(t_o) + 4 \big[A(t_1) S(t_1,t_o) + A(t_3) S(t_3,t_o) + \ldots + A(t_{m-1}) S(t_{m-1},t_o) \big] + 2 \big[A(t_2) S(t_2,t_o) + A(t_4) S(t_4,t_o) + \ldots + A(t_{m-2},t_o) S(_{m-2},t_o) \big] + A(t_m) S(t_m,t_o) \Big\} \tag{2.17}$$

with the infinite series $S(t_i,t_o)$ defined as

$$S(t_i,t_o) = 1 + \sum_{j=1}^{\infty} K_j (t_i,t_o) \quad , \; i = 1,2,\ldots,m. \tag{2.18}$$

Thus, the evaluation of $\Phi(t,t_o)$ essentially reduces to evaluation of $S(t_i,t_o)$ with K_j given recursively by (2.13) and (2.14).

The m infinite series $S(t_i,t_o)$,i=1,...,m can be computed simultaneously. To see this, expand $S(t_i,t_o)$ as below:

$$S(t_1,t_o) = 1 + K_1(t_1,t_o) + K_2(t_1,t_o) + K_3(t_1,t_o) + \cdots\cdots$$

$$S(t_2,t_o) = 1 + K_1(t_2,t_o) + K_2(t_2,t_o) + K_3(t_2,t_o) + \cdots\cdots$$

$$\vdots \qquad (2.19)$$

$$S(t_m,t_o) + 1 + K_1(t_m,t_o) + K_2(t_m,t_o) + K_3(t_m,t_o) + \cdots\cdots$$

with $$K_n(t_i,t_o) = \left((t_i-t_o)/i\right)\Big\{A(t_1)K_{n-1}(t_1,t_o) + A(t_2)K_{n-1}(t_2,t_o)$$

$$+\cdots+ A(t_{i-1})K_{n-1}(t_{i-1},t_o) + A(t_i)/2 \;\; K_{n-1}(t_i,t_o)\Big\} \qquad (2.20)$$

$$K_1(t_i,t_o) = \left((t_i-t_o)/i\right)\Big\{A(t_o)/2 + A(t_1) +\cdots+ A(t_{i-1}) +$$

$$A(t_i)/2\Big\} \qquad i = 1,2,\ldots,m. \qquad (2.21)$$

We can therefore evaluate $S(t_i,t_o)$, i=1,...,m simultaneously as below:

(1) Initiate $S(t_i,t_o) = 1$,i=1,...,m.

(2) Evaluate $K_1(t_i,t_o)$, i=1,...,m by (2.21) and set $S(t_i,t_o) = S(t_i,t_o) + K_1(t_i,t_o)$, i=1,...,m.

(3) Knowing $K_1(t_i,t_o)$, i=1,...,m. evaluate $K_2(t_i,t_o)$, i=1,.m. by (2.20)

(4) Evaluate $S(t_i,t_o) = S(t_i,t_o) + K_2(t_i,t_o)$, i=1,...,m.

(5) Check the convergence of $S(t_i,t_o)$, i=1,...,m. If any of the series converges to the desired accuracy, they are set aside for further addition of higher order K's.

(6) Store K_2 in K_1 and repeat step 3.

Having evaluated $S(t_i,t_o)$, $i=1,\ldots,m$, $\Phi(t,t_o)$ can then be computed from (2.17).

2.3 Remarks on the numerical scheme (2.17) and (2.18).

(1) Only need to evaluate (or know) $A(t)$ at $t_o,t_i,\ldots,t_m$ or $m+1$ function evaluations of $A(t)$ and all computations are done in terms of the $(m+1)$ A's.

(2) Any integration formula can be used to develop similar numerical scheme and recursive formulae such as (2.17), (2.18), (2.20) and (2.21).

(3) The accuracy of the present scheme is that of trapezoidal rule , $o(\frac{t-t_o}{m})^3$.

(4) The series $S(t_i,t_o)$ are convergent series for finite $A(v)$, $v\in(t_o,t)$. This can be shown as follows:

Since $A(v)$ is finite for $v\in(t_o,t)$,

$$A^* = \underset{v\in(t_o,t)}{\text{Max}} /A(v)/ \qquad \text{where } /A/ = (/a_{ij}/)$$

and maximization is done element-wise of $A(v)$. Then, it can be shown (but tedious) that (2.17), upon substitution of A^* in place of $A(t_i)$, $i=0,1,..,m$, gives $I + A^*(t-t_o) + \{A^*(t-t_o)\}^2/2! + \ldots$ which is known to converge to $A^*(t-t_o)$.

Thus, convergence of (2.17) to $e^{A^*(t-t_o)} \Rightarrow S(t_i,t_o)$ converges.

Numerical experiments:

$$\text{Let} \quad A(t) = \begin{bmatrix} -4t & & \\ & -3t & 0 \\ 0 & & -2t \\ & & & -t \end{bmatrix}$$

$t \in [0,1]$

$$\text{True} \quad \Phi(t,t_o) = \begin{bmatrix} e^{-2(t^2-t_o^2)} & & & \\ & e^{-1.5(t^2-t_o^2)} & & \\ & & e^{-(t^2-t_o^2)} & 0 \\ 0 & & & e^{-0.5(t^2-t_o^2)} \end{bmatrix}$$

$$\Phi(1,0) = \begin{bmatrix} e^{-2} & & & \\ & e^{-1.5} & & 0 \\ & & e^{-1} & \\ 0 & & & e^{-0.5} \end{bmatrix} = \begin{bmatrix} 0.135 & & & \\ & 0.223 & & 0 \\ & & 0.368 & \\ 0 & & & 0.606 \end{bmatrix}$$

The numerical $\Phi(1,0)$ is

$$\Phi(1,0) = \begin{bmatrix} 0.132 & & & \\ & 0.221 & & 0 \\ & & 0.367 & \\ & & & 0.606 \end{bmatrix}$$

Using m=10 and the maximum no. of terms for the convergence of S series ($\epsilon = 10^{-5}$) is 11. The maximum pointwise error is 3% and the error norm (absolute norm) is 0.6%

A summary for $\Phi(1,0)$ using other m's is given below:

$t_o=0$, $t = 1$, convergence criterion of S series $= 1\times10^{-5}$

M	Maximum no. of term in S series for convergence	Computy time, secs, IBM 370/155	Maximum pointwise error	Error norn
4	18	0.453	18%	3.7%
6	13	0.629	8%	1.6%
8	12	0.925	4%	0.9%
10	11	1.251	2.9%	.55%

2.4 Application to ill-conditioned system.

Ill-condition is attributed to the wide spread in the values of the eigenvalues of the coefficient matrix A. This can happen if the elements of A have widely different values.

Since the series $S(t_i,t_o)$ defined by (2.18) converges for any finite $A(t)$, the numerical scheme presented in the previous sections should work, in principle, for any finite A either ill-conditioned or not. However, in actual computer computation, large magnitude of the elements of A can cause overflow before the series $S(t_i,t_o)$ converges. This can be seen by the nature of $S(t_i,t_o)$ which resembles the expansion of an exponential quantity. To render the numerical scheme applicable, methods to get around the effect of the large magnitude of some elements of A are given below:

In views of (2.20) and (2.21), we can make K_n small by either using smaller time interval $(t_i-t_o)/i = h$, or reducing the magnitude of A(t). Thus we have the following two methods for handling A(t) with large elements.

Method 1 :- Subdivesion of time interval.

Use an m large enough so that interval $h = t-t_o/m = t_i-t_o/i$ is small enough to make K as given by (2.20) decreasing fast enough. Increasing m tenfold is equivalent to decreasing the magnitude of the elements of A(t) by tenfold. Thus if m is large enough, blow-up phenomenon can be avoided when evaluating the series $S(t_i,t_o)$. Alternatively we may evaluate transition matrix for each subinterval and by the property of transition matrix,

$$\Phi(t_m,t_o) = \Phi(t_m,t_{m-1})\ \Phi(t_{m-1},t_{m-2}).... \ \Phi(t_2,t_1)\,\Phi(t_1,t_o) \tag{2.22}$$

Increasing m by tenfold, the number of series $S(t_i,t_o)$ to be evaluated also increases by tenfold. Thus, although this method is straightforward, it is computing-time consuming. To reduce the computing time requirement, method 2 in the sequel is recommended.

Method 2. Scaling down the elements of A(t)

Let $A'(t) = A(t)/c$ where $c \geq 1$ is a suitable scale factor. If we substitute A'(t) into (2.21) and (2.20), we find that

$$cK_1'(t_i,t_o) = K_1(t_i,t_o)$$

$$c^2 K_2'(t_i,t_o) = K_2(t_i,t_o)$$
$$\vdots$$
$$c^n K_n'(t_i,t_o) = K_n(t_i,t_o)$$

where K_j' is the K_j with $A(t)$ replaced by $A'(t)$.

Therefore (2.18) gives

$$S(t_i,t_o) = 1 + \sum_{j=1}^{\infty} c^j K_j'(t_i,t_o) \ , \quad i = 1,\ldots,m \tag{2.23}$$

With sufficiently large scale factor c, $K_j'(t_i,t_o)$ will not cause overflow. But c^j can still make $c^j K_j'(t_i,t_o)$ overflow before convergence of the series $S(t_i,t_o)$. To avoid this, choose an $k>0$ which is sufficiently large such that the following equivalent form of (2.23):

$$S(t_i,t_o) = 1 + c^k \sum_{j=1}^{\infty} c^{j-k} K_j'(t_i,t_o) \tag{2.24}$$

can be evaluated without overflow. k can be chosen as the number of terms that it takes $S(t_i,t_o)$ to converge to the desired degree of accuracy. Different c and k can be chosen for different time interval indicated by the index i.

Numerical Experiments:

Method 1 has been successful in evaluating the transition matrix of a 20-dimensional system obtained from the numerical linearization of a nonlinear system involving complicated chemical reactions. The diagonal elements of the coefficient matrix ranges from 10^{-3} to about 500. Since the system is large dimensional and true solution is not available for comparison, the results of the computer computations are not given. Although method 2 has not been numerically tested, it is expected to work also since method 2 merely scales down the magnitudes of the terms in evaluating the series S to avoid overflow in the computer as does method 1.

CHAPTER 3. CONCLUSION

A numerical scheme for the evaluation of the transition matrix of a time-varying dynamical system is developed by the direct numerical integrations of the matrizant. The evaluation of the transition matrix reduces to the evaluation of a finite set of algebraic series. To carry out the numerical scheme, the coefficient matrix of the dynamical system needs to be known only at a finite number of time instants.

The numerical scheme is found to be applicable to ill-conditioned system. Large dimensionality of the dynamical system will not cause any additional difficulty except involving larger computing time.

The numerical scheme developed in this proposition shall be useful for solving a set of ordinary linear differential equations obtained from linearization of nonlinear differential equations with chemical reactions, and of large dimensionality.

REFERENCES

Amundson, N. R.,"Mathematical Methods in Chemical Engineering", Prentice-Hall Inc., Englewood Cliffs, N. J. (1966).

Koppel, L. B., "Introduction to Control Theory", Prentice-Hall Inc. Englewood Cliffs, N. J. (1968).

Lapidus, L., "Digital Computation for Chemical Engineers", McGraw Hill Book Company, Inc., New York (1962).

Reprinted from I&EC FUNDAMENTALS, Vol. 9 Page 596, November 1970

PROPOSITION III

Flow of Single-Phase Fluids through Fibrous Beds

Chwan P. Kyan, Darshanlal T. Wasan,[1] and Robert C. Kintner
Department of Chemical Engineering, Illinois Institute of Technology, Chicago, Ill. 60616

A pore model for the flow of a single-phase fluid through a bed of random fibers is proposed. An effective pore number, N_e, accounts for the influence of dead space on flow; deflection number, N_d, characterizes the effect of fiber deflection on pressure drop. Experimental data were obtained with glass, nylon, and Dacron fibers of 8- to 28-micron diameter and with fluids of viscosity ranging from 1 to 22 cp. A generalized friction factor–Reynolds number equation is presented. The effects of dead space in a fibrous bed on flow and of fiber deflection on pressure drop have no parallels in a granular bed.

The flow of fluids through porous media has been a subject of investigation for many years. A considerable amount of research has been done on the flow through granular beds and many useful results have been obtained (Brownell and Katz, 1947, 1956; Ergun, 1952; Ergun and Orning, 1949). A lesser number of investigations have been done on the phenomena of flow of fluids through fibrous media, mostly in connection with aerosol filtration.

General approaches pursued by most workers on the flow of fluids through fibrous beds involved the development of theoretical pressure drop equations from either a "channel model" or a "drag model." The former was the more extensively used.

Most workers using the channel model started with the Kozeny-Carman equation,

$$\frac{\Delta P}{L} = k\mu U S^2 \frac{(1-\epsilon)^2}{\epsilon^3} \tag{1}$$

which in the friction factor form becomes

$$f = \frac{k}{kN_{Re}} \tag{2}$$

In this equation, the fact that k depends on fiber orientation and porosity had been observed and discussed by Sullivan and Hertel (Sullivan, 1941; Sullivan and Hertel, 1940) based on their experimental work. Thus Equation 1 was inadequate for pressure drop correlations. Various workers using the channel model have elaborated upon Equation 1 with modification for shape and orientation (Davies, 1952; Fowler and Hertel, 1940; Langmuir, 1942; Sullivan and Hertel, 1941).

Most workers (Chen, 1955; Iberall, 1950; Wong *et al.*, 1956) using the drag model rejected the applicability of the channel model because of the high porosity of a fibrous bed and derived a pressure drop equation by considering the drag forces due to fluid flow on the fibers.

Wong *et al.* (1956) employed an effective drag coefficient, C_{De}, to account for the fiber orientation, interference of neighboring fibers, fiber ends, and nonuniformity of fiber distribution in the bed. They concluded that the fiber volume fraction, γ, has a marked effect on C_{De}. The higher the value of γ, the higher is the neighboring fiber interference which leads to a higher C_{De}. They also noticed the leveling off of the effective drag coefficient–Reynolds number plot at Reynolds numbers greater than 6.

Gunn and Aitken (1961) in their study of the mechanism

[1] To whom correspondence should be sent.

of the flow of air and water through packed glass fibers found that pressure drop data depended on the history of previous gas and liquid flow rates through the bed.

The effect of interference by neighboring fibers has been recently explored by Spielman and Goren (1968). They considered a body damping force to be proportional to the local velocity.

No satisfactory general pressure drop (or friction factor) correlation which takes into account the nature of the fibers and the wide range of porosity for the flow of a single-phase fluid through fibrous beds has been obtained hitherto.

Development of the Model

The most peculiar thing in the case of the flow of fluid through a fibrous bed is the fact that it gives an unexpectedly high pressure drop in spite of the high porosity of the bed. The causes of this high pressure drop are postulated as follows:

Only a fraction of the free space as calculated from the bulk density of the bed is available for fluid flow, the rest being occupied by stagnant fluid.

Some energy is absorbed by deflection of individual fibers, causing an additional pressure drop other than those of a fluid dynamic nature.

In accordance with the foregoing postulations, a model is proposed. Models such as "fibers parallel to the direction of flow," "fibers normal to the direction of flow," and "gridwork" models are not used because they do not give stagnant regions and cannot be conveniently arranged to give a "normal high porosity."

The present model consists of inclined fibers intercrossed with those that are transverse to the direction of flow locked between them to make up a stable arrangement. The angle of inclination between two adjacent inclined fibers is α, with spacing between parallel fibers characterized by a model spacing number, n.

The model is diagrammatically sketched in Figures 1 and 2.

The fluid flows only through the wider spaces $ABCD$ as shown in Figure 2, the rest of the space of the elementary unit of the model being occupied by the nonflowing fluid.

The fibers in the elementary unit of Figure 2 are being bent by the drag force of the fluid, lending to a dissipation of energy due to fiber deflection.

Mathematical Relations Derived from Model. Relation of Porosity and Spacing Number, n. Based on the shaded elementary volume of the bed, as shown in Figure 3, we have

$$\text{Total volume} = (n\,D_f)\,(n\,D_f)\,(n\,D_f \sin\alpha) = n^2 D_f^3 \sin\alpha$$

Total cell fiber volume =

$$2\left[(n\,D_f)\left(\frac{\pi}{4}D_f^2\right)\right] + 4\left[(n\,D_f)\left(\frac{\pi}{8}D_f^2\right)\right] = n\,\pi\,D_f^3 \tag{3}$$

$$\epsilon = 1 - \frac{\text{fiber volume}}{\text{total volume}} = 1 - \frac{\pi}{n^2 \sin\alpha} \tag{4}$$

or

$$n = \sqrt{\frac{\pi}{(1-\epsilon)\sin\alpha}} \tag{5}$$

Effective Porosity of Bed, ϵ_e. Referring to Figures 1 and 2 for the elementary unit shown as shaded,

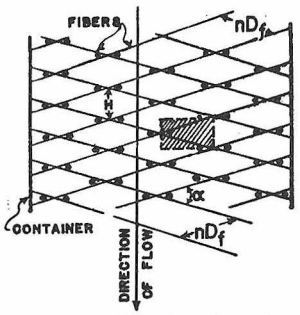


Figure 1. Front view of model

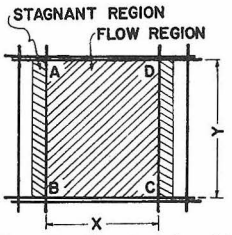


Figure 2. Plan view of model

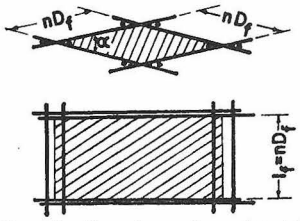


Figure 3. Elementary volume of model

Volume of

flow region $= (H)(XY) =$

$$\left(2nD_f \sin\frac{\alpha}{2}\right)(n-2)D_f(n-3)D_f\cos\frac{\alpha}{2}$$

$$= nD_f^3\,(n-2)(n-3)\sin\alpha \simeq nD_f^3(n-2.5)^2\sin\alpha \tag{6}$$

Since the value of n is usually about 8 and seldom goes below 5 (corresponding to $\epsilon = 0.75$), the error introduced by the above simplification is tolerable.

From Equations 3 and 6, we have

$$\epsilon_e = \frac{\text{volume of flow region}}{\text{total volume}} = \frac{(n-2.5)^2}{n^2} \tag{7}$$

Combining Equations 5 and 7 and defining an effective pore number,

$$N_e = \sqrt{\frac{\pi}{(1-\epsilon)\sin\alpha}} - 2.5 \tag{8}$$

we have

$$\epsilon_e = N_e^2 (1 - \epsilon) \frac{\sin \alpha}{\pi} \tag{9}$$

The stagnant space and the volume occupied by the fibers in the bed are characterized by the number 2.5 in the above relationship.

The value of α in Equations 8 and 9 can be readily obtained by applying to Equation 9 the limiting condition, $\lim_{\epsilon \to 0} \epsilon_e = 0$, yielding

$$\alpha = 30.17^\circ \approx 30^\circ \tag{10}$$

Equivalent Diameter, D_e. Considering the elementary shaded unit in Figures 1 and 2, and defining,

$$D_e = \frac{4 \text{ cross-sectional area of flow}}{\text{wetted perimeter}}$$

we have

$$D_e = \frac{2 N_e^2 D_f \cos \frac{\alpha}{2}}{(n - 3) + (n - 2)}$$

$$= 0.9656 N_e D_f \tag{11}$$

Actual Velocity through Bed.

$$U^* = \frac{U}{\epsilon_e} \tag{12}$$

Equations 9 and 12 give

$$U^* = \frac{1.9895 \pi U}{N_e^2 (1 - \epsilon)} \tag{13}$$

Pressure Drop Equation. Total pressure drop = pressure drop due to viscous flow losses + pressure drop caused by form drag + pressure drop due to deflection of fibers—i.e.,

$$\Delta P = \Delta P_{\text{flow}} + \Delta P_{\text{form drag}} + \Delta P_{\text{deflection}} \tag{14}$$

The term $(\Delta P)_{\text{flow}}$ can be accounted for using the approach of Ergun and Orning (Ergun, 1952; Ergun and Orning, 1949),

$$\frac{dP}{dL} = C_1 \frac{\mu U^*}{D_e^2} \tag{15}$$

The term on the right-hand side of Equation 15 accounts for the viscous losses only. The kinetic-loss term in the original relationship of Ergun and Orning (Ergun, 1952; Ergun and Orning, 1949) can be ignored for a fibrous bed, since it is insignificant in the normal operation.

Replacing U^* and D_e in Equation 15, using Equations 11 and 13, yields

$$\left(\frac{\Delta P}{L}\right)_{\text{flow}} = k_1 \frac{\mu U}{D_f^2 (1 - \epsilon) N_e^4} \tag{16}$$

The term $(\Delta P)_{\text{drag}}$ can be accounted for by a drag equation. Referring to the elementary unit as shown shaded in Figures 1 and 2,

$$\text{Drag force} = \frac{C_D \rho U^{*2} A_f}{2} \tag{17}$$

Work done by drag force through H per unit mass of fluid is

$$(W)_{\text{drag}} = \left(\frac{\Delta P'}{\rho}\right)_{\text{form drag}} = \frac{C_D \rho U^{*2} A_f H}{2 (XY) H \rho} \tag{18}$$

Therefore,

$$\left(\frac{\Delta P}{L}\right)_{\text{form drag}} = \left(\frac{\Delta P'}{H}\right)_{\text{drag}} = \frac{C_D \rho U^{*2} A_f}{2 HXY} \tag{19}$$

where the projected area,

$$A_f = 4 n D_f^2 \tag{20}$$

For a fibrous bed, it can be observed (Wong *et al.*, 1956) that

$$C_D = C_1 \left(\frac{\mu}{D_f U \rho}\right) \tag{21}$$

Combining Equations 19, 20, and 21 with 6 and 13, we have

$$\left(\frac{\Delta P}{L}\right)_{\text{form drag}} = k_2 \left(\frac{\rho U^2}{D_f N_e^6}\right) \left(\frac{\mu}{D_f U \rho}\right) \frac{1}{(1 - \epsilon)^2} \tag{22}$$

The term $(\Delta P)_{\text{deflection}}$ represents the possibility that the deflection of a fiber will absorb some energy of flow; it is accounted for as follows:

Referring to the elementary shaded unit in Figures 1 and 2, maximum deflection is

$$Y_{\max} = C_2 \frac{F l_f^3}{EI} \tag{23}$$

where F is the drag force acting on the fiber, l_f is the length of the fiber in the elementary unit, E is the modulus of elasticity of the fiber, and I is the volume moment of inertia of the fiber.

$$\text{Work done} = F Y_{\max} = \frac{C_2 F^2 l_f^3}{EI}$$

Work done per unit mass of fluid flowing through the elementary unit is

$$(W)_{\text{deflection}} = \left(\frac{\Delta P'}{\rho}\right)_{\text{deflection}} = \frac{C_2 F^2 l_f^3}{EI XY H \rho} \tag{24}$$

Substituting Equations 4, 6, 13, 17, 20, and 21 in Equation 24 gives

$$\left(\frac{\Delta P}{L}\right)_{\text{deflection}} = k_3 \left(\frac{\mu}{D_f U \rho}\right)^2 \frac{\rho^2 U^4}{E D_f N_e^{10} (1 - \epsilon)^{5.5}} \tag{25}$$

Therefore, combining Equations 16, 22, and 25, the total pressure drop equation becomes

$$\frac{\Delta P}{L} = k_1 \frac{\mu U}{D_f^2 N_e^4 (1 - \epsilon)} + k_2 \left(\frac{\mu}{D_f U \rho}\right) \frac{\rho U^2}{D_f N_e^6 (1 - \epsilon)^2} + k_3 \left(\frac{\mu}{D_f U \rho}\right)^2 \frac{\rho^2 U^4}{E D_f N_e^{10} (1 - \epsilon)^{5.5}} \tag{26}$$

Friction Factor Equation. Denoting

$$N_{\text{Re}} = \frac{D_f U \rho}{\mu (1 - \epsilon)}, \quad \text{Reynolds number} \tag{27}$$

$$N_a = \frac{\mu^2}{E D_f^2 \rho}, \quad \text{deflection number} \tag{28}$$

and

$$f_{fk} = \frac{\Delta P}{L} \frac{D_f}{\rho U^2} N_e^4 (1 - \epsilon)^2, \quad \text{kinetic friction factor for a fibrous bed} \tag{29}$$

Equation 26 becomes

$$f_{fk} = \left[k_1 + \frac{k_2}{N_e^2 (1 - \epsilon)}\right] \frac{1}{N_{\text{Re}}} + k_3 \frac{N_a}{N_e^6 (1 - \epsilon)^{2.5}} \tag{30}$$

Several points are worth mentioning in connection with Equation 30.

The term in parentheses which represents the coefficient of $\frac{1}{N_{Re}}$ is essentially a "permeability function," dependent upon porosity, ϵ.

The deflection number, N_d, the ratio of the viscous drag of the fluid to the elastic force of the fiber, characterizes the effect of fiber deflection on pressure drop.

The term $N_e^2 (1 - \epsilon)$ accounts for the effective porosity instead of the apparent porosity, ϵ, in the fibrous bed.

Another form of Equation 30 can be obtained after dividing it throughout by the coefficient of the term $\frac{1}{N_{Re}}$, as follows:

$$f_n = \frac{1}{N_{Re}} + f_d \tag{31}$$

where

$$f_n = \frac{\frac{\Delta P}{L}\frac{D_f}{\rho U^2} N_e^4 (1 - \epsilon)^2}{k_1 + \frac{k_2}{N_e^2 (1 - \epsilon)}} \tag{32}$$

and

$$f_d = \frac{k_3 N_d}{\left[k_1 + \frac{k_2}{N_e^2 (1 - \epsilon)}\right] N_e^6 (1 - \epsilon)^{3.5}} \tag{33}$$

Equation 31 is now simple enough for experimental verification.

Experimental Apparatus and Procedures

The equipment for the present investigation (Figures 4 and 5) consisted essentially of a 55-gallon stainless steel recycle tank, centrifugal pump driven by a 3/4-hp motor, two rotameters, a manometer, and an adjustable flanged connection to hold the fibrous bed test section. The setup also included a needle valve for adjusting flow rate through the smaller rotameter, a manometer fluid reservoir, a recycle filter, and air-seal arrangement in the downstream of the test section.

The fibrous bed test section (Figure 5) was composed of a 1 3/4-inch-i.d. 14-inch-long Plexiglas section, two coarse supporting screens, two supporting aluminum rings, and a 1/8-inch-thick fibrous pad. The functions of the pad were to damp out any sudden change in pressure to avoid sudden compacting of the fibrous bed, to normalize flow, and to act as an additional filter. Four 1/32-inch pressure taps were located 3 inches apart along the tube. Two 1/8-inch air taps were also provided. Preweighed fibers were carefully dispersed in water and the fiber suspension was filtered portion by portion into the test section under partial vacuum to make a randomly packed fibrous bed, which was secured in the desired position by supporting screens and aluminum rings. It was essential to make a bed sufficiently compact to stay rigid and stable during experimental runs. The relative positions of the fibrous bed and pressure taps are shown in Figures 4 and 5. The taps were used to measure the pressure drop through an actual 3-inch length of bed, excluding the entrance and exit losses.

During each run the rotameter reading, manometer reading, and temperature of fluid were noted. The flow rate was gradually increased to maximum flow and then gradually decreased to check the consistency of readings.

It was important to make certain in each run that the fibrous bed was rigid and stable enough and that there were no air bubbles in the line or manometer tubing.

In the present investigation, viscosities of glycerol solutions were measured with an Oswald-Cannon-Fenske viscometer. Diameters of fibers were determined with the aid of a microscope.

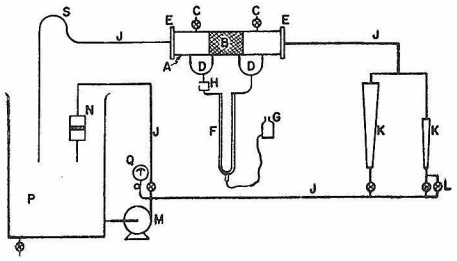


Figure 4. Schematic diagram of equipment

A. Plexiglas test section
B. Fibrous bed
C. Air taps
D. Three-way cocks
E. Adjustable flanges
F. U-tube manometer
G. Manometer fluid reservoir
H. Manometer fluid trap
J. Iron pipe
K. Rotameters
L. Needle valve
M. Centrifugal pump
N. Recycle filter
P. 55-gallon storage tank
Q. Pressure gage
S. Air seal

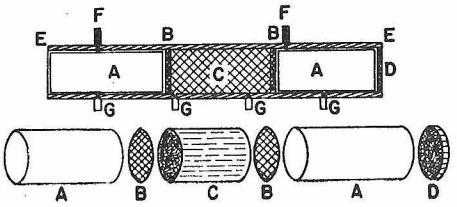


Figure 5. Fibrous bed test section

A. Supporting rings made from aluminum tube
B. Coarse wire screen
C. Fibrous bed
D. Upstream fibrous pad for filtering, flow normalizing, and damping
E. Plexiglas pipe
F. Air taps
G. Pressure taps

Experimental data were obtained with glass, Dacron, and nylon fibers of 8- to 28-micron diameters and with water and aqueous glycerol solution of a viscosity range of 1 to 22 cp.

Results and Discussion

Proportionality constants k_1, k_2, and k_3 in Equation 30 were evaluated by using two sets of experimental data plotted as f_{fr} *vs.* $\frac{1}{N_{Re}}$ on a rectangular plot. Constants k_1 and k_2 were calculated from the slopes of the best straight line fitted to the experimental data of runs 1 and 5.

k_1 and k_2 were found to be 62.3 and 107.4, respectively. Since the values for modulus of elasticity, E, for the type of fibers used were not available, the values of (k_2/Eg_c) were found from the intercepts of the best fits as 1.74 and 25.4 sec^2-ft/lb mass for glass and nylon fibers, respectively. The group (k_3/Eg_c) for Dacron fiber was estimated by using one data point in Equation 34 and found to be 29.2 sec^2-ft/lb mass.

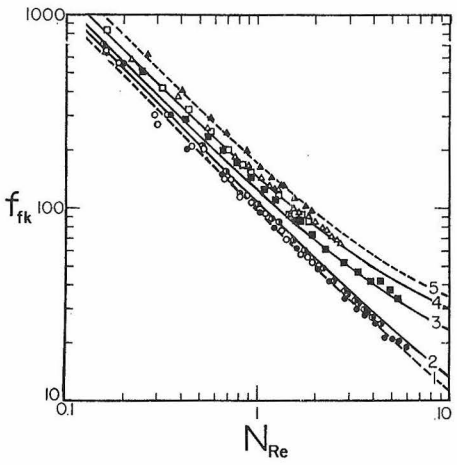


Figure 6. Friction factor plot for randomly packed fibrous beds

Run		Fiber	D_f	ϵ	Fluid	μ cp
1	●	Glass	8	0.919	Water	0.87–0.9
2	○	Glass	8	0.868	Water	0.9–0.945
3	■	Dacron	18	0.829	Water	0.9
4	△	Nylon	20	0.765	Water	0.91–0.96
5	▲	Nylon	28	0.682	Water	0.9
6	⊕	Glass	8	0.892	73% glycerol	21.82
7	⊖	Glass	8	0.901	63% glycerol	9.96
8	⊹	Glass	8	0.894	58% glycerol	7.13
9	⊷	Glass	8	0.896	50% glycerol	4.75
10	◑	Glass	8	0.895	Water	0.85–0.88
11	□	Dacron	13	0.820	Water	0.945

——— Equation 34 calculated for indicated run
– – – – Best fit to data

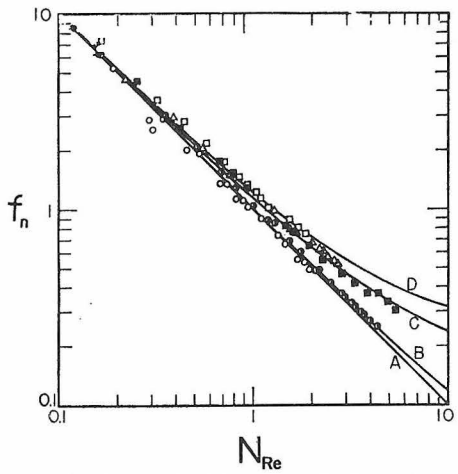


Figure 7. Normalized friction factor plot for randomly packed fibrous beds

Run		Fiber	D_f	ϵ	Fluid	μ cp	f_d
2	○	Glass	8	0.868	Water	0.9–0.945	0.0259
3	■	Dacron	18	0.829	Water	0.9	0.1060
4	△	Nylon	20	0.765	Water	0.91–0.96	0.1363
6	⊕	Glass	8	0.892	73% glycerol	21.82	11.72
7	⊖	Glass	8	0.901	63% glycerol	9.96	2.32
8	⊹	Glass	8	0.894	58% glycerol	7.13	1.238
9	⊷	Glass	8	0.896	50% glycerol	4.75	0.540
10	◑	Glass	8	0.895	Water	0.85–0.88	0.0175
11	□	Dacron	13	0.820	Water	0.945	0.2150

A = Equation 31 with $f_d = 0$ C = Equation 31 with $f_d = 0.1363$
B = Equation 31 with $f_d = 0.0175$ D = Equation 31 with $f_d = 0.2150$

Thus, for a randomly packed fibrous bed, Equation 30 becomes

$$f_{fk} = \frac{1}{N_{Re}}\left[62.3 + \frac{107.4}{N_c^2(1-\epsilon)}\right] + \left(\frac{k_3}{k_2}\right)\frac{\mu^2}{D_f^2 \rho N_c^6 (1-\epsilon)^{3.5}} \quad (34)$$

Figure 6, a comparison of the present experimental data and the results predicted from Equation 29, shows our equivalent of a Kozeny-Carman plot. It is not identical, in that our k_1 term contains a shape correction for fibers and their orientation. The solid lines represent the results calculated from the equation and the dashed lines represent the best fits to the two chosen sets of data. The proposed equation fits the data well. A pronounced dependence of the friction factor, f_{fk}, on porosity is indicated at a constant value of the Reynolds number.

Figure 7 compares the present experimental data with the results predicted from the normalized friction factor Equation 31. This figure utilizes the k_2 term to give a much more satisfying correlation of our data, thus showing the necessity of the term. The figure clearly indicates that f_n is independent of both the porosity and the deflection number at low Reynolds numbers. However, at high Reynolds numbers the curves level off, depending upon the magnitude of the deflection losses. These curves begin to deviate from a slope of -1 at different values of Reynolds numbers for different fibers and porosities. Moreover, the turbulent losses were estimated to be insignificant over the whole range of the Reynolds number encountered in the present experiments (Kyan, 1969). Thus, this behavior could not be attributed to the turbulent losses.

Comparison of Proposed Correlation with the Literature Data. Pressure drop data for the flow of a single-phase fluid through a fibrous bed are relatively scarce in the literature. Furthermore, the data were reported on short beds where the upstream and downstream disturbances and the entrance scaling effect due to the deposition of suspended foreign particles on the front face of the bed could be appreciable. In spite of these and other shortcomings, attempts were made to compare the proposed Equation 31 with the literature data.

Figure 8 shows a comparison of Equation 31 with data reported by Sherony (1969), for the flow of water through glass and nylon beads. The agreement is generally acceptable. The discrepancy between the calculated and measured values is attributed to the variations of $\Delta P/L$ with bed length.

Figure 9 compares Equation 31 with the data of Spielman (1967) and Gunn and Aitken (1961), for water and air, respectively. The data generally lie above the proposed correlation. This discrepancy between the calculated and experimental values may be attributed to the collapsing of the bed at these high porosities. If a fibrous bed is packed to a porosity higher than that at which the bed is both stable

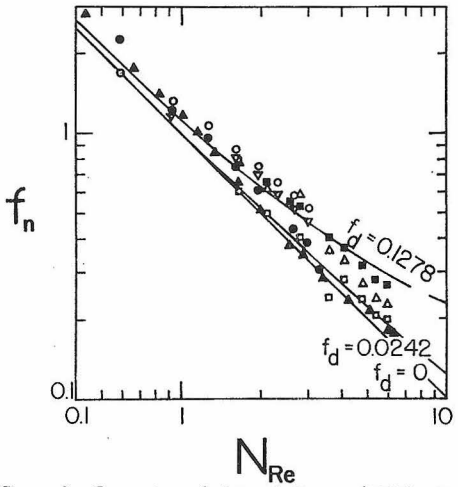


Figure 8. Comparison of data of Sherony (1969) with Equation 31

Run	Fiber	D_f	ϵ	L, Inches	f_d
○	Glass	8	0.895	2.0	0.0242
▽	Glass	8	0.894	0.5	0.0244
●	Glass	8	0.894	1.0	0.0244
△	Nylon	20	0.851	0.5	0.0778
□	Nylon	20	0.850	2.0	0.078
▲	Nylon	20	0.787	1.0	0.1278
■	Nylon	20	0.850	1.0	0.078

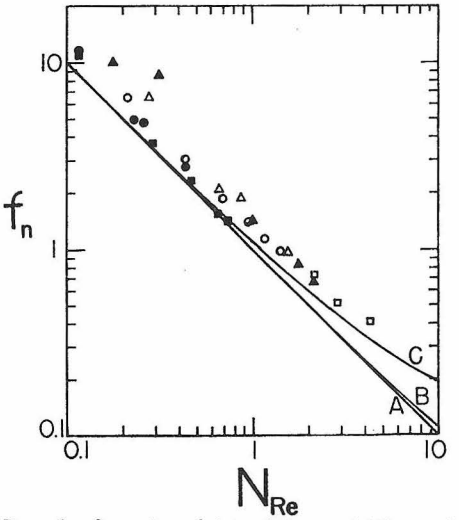


Figure 9. Comparison of data of Gunn and Aitken, and Spielman with Equation 31

Run	D_f	ϵ	L, Inches	Fluid	f_d	Source
○	6	0.942	0.30	Water	0.0326	Spielman
■	3.5	0.944	0.13	Water	0.0941	Spielman
●	3.5	0.945	0.33	Water	0.0933	Spielman
△	6	0.946	0.13	Water	0.0317	Spielman
▲	12	0.924	0.23	Water	0.0091	Spielman
□	9.73	0.95	0.81	Air	0.0029	Gunn and Aitken

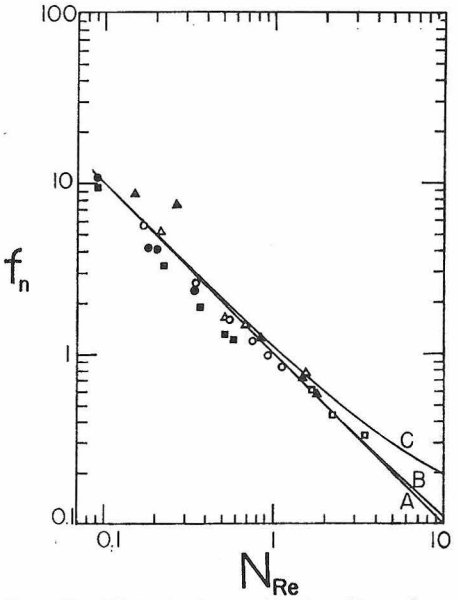


Figure 10. Effect of collapse of bed on friction factor plot

Porosity reduction of 1.5% in data of Gunn and Aitken (1961) and Spielman (1967). Symbols and data same as for Figure 9

and tight, it is liable to collapse under the action of fluid flow. Therefore, the true porosity will be lower than the apparent value. For example, if only a 1.5% reduction in porosity is assumed for data reported by Spielman and Gunn *et al.*, and friction factor values are recalculated, the data lie below the proposed correlation as shown in Figure 10.

Most of the other investigators (Brown, 1950; Lord, 1955; Wiggins *et al.*, 1939) have reported data on the flow of single-phase fluids through fibrous beds in the form of the Kozeny constant, k, as a function of bed porosity. For the sake of comparison of the present correlation with their data, Equations 1 and 34 are combined and rearranged to yield an expression for the Kozeny constant.

The Kozeny-Carman Equation 1 written for a bed of cylindrical particles takes the form

$$\frac{\Delta P}{L}\,\frac{D_f\,\epsilon^3}{\rho U^2\,(1-\epsilon)\,16\left(1+\dfrac{D_f}{2L_f}\right)^2 k} = \frac{1}{N_{Re}} \qquad (35)$$

A comparison of Equations 34 with 35 yields

$$k = \frac{[62.3\,N_e^2\,(1-\epsilon) + 107.4]\,\epsilon^3\,[1 + f_d\,N_{Re}]}{16\,N_e^6\,(1-\epsilon)^4\left(1+\dfrac{D_f}{2L_f}\right)^2} \qquad (36)$$

Equation 36 shows that k depends on ϵ, N_d, N_{Re}, and $\frac{D_f}{L_f}$. For a fibrous bed, since $L_f \gg D_f$, Equation 36 simplifies to

$$k = \frac{[62.3\,N_e^2\,(1-\epsilon) + 107.4]\,\epsilon^3\,(1 + f_d\,N_{Re})}{16\,N_e^6\,(1-\epsilon)^4} \qquad (37)$$

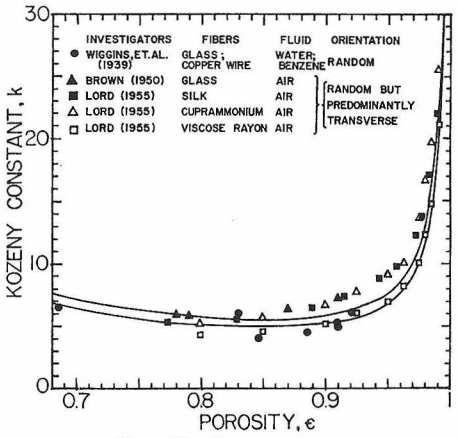


Figure 11. Kozeny constant
Comparison of data with Equation 37

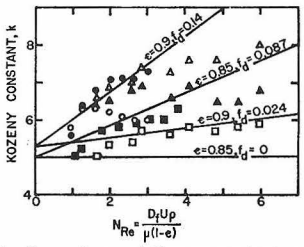


Figure 12. Dependency of Kozeny constant on Reynolds number

Malhotra (1969). Water at 68°F

Symbols	Fibers	D_f Microns	ϵ	L, Inches	f_d
●	Glass	8	0.895	2.0	0.0242
○	Glass	8	0.894	1.0	0.0244
△	Nylon	20	0.850	1.0	0.0780
▲	Nylon	20	0.851	0.5	0.0778
■	Nylon	20	0.804	1.0	0.1123
□	Nylon	20	0.850	2.0	0.0780
——.	Equation 37				

Equation 37 is plotted for $f_d = 0$ and for $f_d \cdot N_{Re} = 0.1$' and compared with the published k values in Figure 11. A good agreement is obtained between the predicted and experimental values for the randomly oriented fibers.

Equation 37 also shows that for a constant value of N_d and ϵ, k increases linearly with N_{Re}. This behavior is depicted in Figure 12 and compared with the water data reported by Malhotra (1969). The general trend appears to be correct.

The agreement between the proposed correlations as given by Equations 31 and 34 and the present and literature data appears to be satisfactory. However, it is very difficult to pack different beds to the same degree of randomness and porosity distribution. Therefore, the pressure drop is somewhat sensitive to variations in bed length. This effect was minimized in the present investigation by using a 3-inch length of bed, so that any nonuniformity was averaged out.

Conclusions

An effective pore model is proposed for flow of a fluid through a randomly packed fibrous bed. Two dimensionless parameters, N_d and N_e, were obtained as a result of a theoretical development based on the proposed model. N_d is a characteristic physical property group which is a measure of the effect of fiber deflection on pressure drop and N_e accounts for the effect of stagnant space in a fibrous bed on flow. The effects of these parameters have no parallels in a granular bed.

Friction-factor Equations 31 and 34 were developed for the flow of a single-phase fluid through a fibrous bed. The f_{fk} or f_n *vs.* N_{Re} curve correlates the data satisfactorily. The effect of N_d or f_d on pressure drop was found to be significant.

An expression (Equation 37) for the Kozeny constant, k, was obtained. It shows that k is strongly dependent on N_d, N_{Re}, and ϵ and hence the usual one-term Kozeny-Carman equation is not applicable for flow of single-phase fluids through fibrous beds.

Nomenclature

A_f = projected area of fiber in elementary unit = $4\,nD_f^2$, in^2
C_i, $i = 1, 2$ = numerical constants
C_D = drag coefficient
C_{De} = effective drag coefficient
D = diameter of bed, in
D_e = equivalent diameter of flow, in
D_f = diameter of fiber, in or microns
E = modulus of elasticity of fiber, lb/in (sec^2)
F = drag force defined in Equation 23
f = friction factor, $\frac{\Delta P}{L}\frac{1}{SU^2}\frac{3}{(1-\epsilon)}$
f_d = normalized friction factor due to deflection as defined by Equation 33
f_{fk} = kinetic friction factor for fibrous bed as defined by Equation 29
f_n = normalized friction factor for random-packed fibrous bed as defined by Equation 32
g_c = conversion constant, (lb$_m$/lb$_f$)(in/sec^2)
H = height of elementary unit of model, $2nD_f \sin 15°$, in
I = volume moment inertia of fiber = $\frac{\pi}{4}\left(\frac{D_f}{2}\right)^4$, (in)^4
k = Kozeny constant
k_i, $i = 1, 2, 3$ = numerical constant
L = length of bed, in
L_f = length of fiber, in
l = length in general, in
l_f = length of fiber in elementary unit of the model, nD_f, in
n = model spacing number
N_d = deflection number, $\frac{\mu^2}{ED_f^2\rho}$
N_e = effective pore number as defined in Equation 8
N_{Re} = Reynolds number, $D_f U\,\rho/\mu(1-\epsilon)$
$N_{Re'}$ = Reynolds number defined as $\frac{\mu S(1-\epsilon)}{\rho U}$
P = pressure, lb$_f$/in^2
ΔP = pressure drop through bed length L, lb$_f$/in^2
P' = pressure drop through bed length H, lb$_f$/in^2
S = ft^2 of packing surface per ft^3 of packed volume, ft^{-1}
U = superficial velocity through bed, $\frac{\text{volumetric flow rate}}{\text{cross-sectional area of bed}}$, in/sec
U^* = actual velocity through bed, in/sec
W = work, (lb$_f$) (in)

$\bar{W}$ = work done per unit mass of fluid, lb_f-in/lb_m
X = width of flow area of model = $(n - 3)\ D_f \cos 15°$, in
Y = length of flow area of model = $(n - 2)\ D_f$, in

GREEK SYMBOLS

α = angle of inclination between two fibers of model = 30.17°
γ = fiber volume fraction
ϵ = porosity
ϵ_e = effective porosity
μ = absolute viscosity of fluid, lb_m/in sec or cp
ρ = density of fluid, lb_m/in^3

Literature Cited

Brown, J. C., *Tappi* **33,** 130 (1950).
Brownell, L. E., Katz, D. L., *Chem. Eng. Progr.* **43,** 549–54, 601–12 (1947).
Chen, C. Y., *Chem. Rev.* **55,** 595 (1955).
Davies, C. N., *Proc. Inst. Mech. Eng. (London)* **B1,** 185 (1952).
Ergun, S., *Chem. Eng. Progr.* **48** (2), 89–94 (1952).
Ergun, S., Orning, A. A., *Ind. Eng. Chem.* **41,** 1179 (1949).
Fowler, J. L., Hertel, K. L., *Appl. Phys.* **11,** 496 (1940).
Gunn, D. J., Aitken, A. R., *Can. J. Chem. Eng.* **39,** 209–14 (1961).
Iberall, A. S., *J. Res. Nat. Bur. Stand.* **45,** 398 (1950).
Kyan, C. P., M.S. thesis, Illinois Institute of Technology, Chicago, Ill., January 1969.
Langmuir, I., Office of Scientific Research and Development, Rept. 865 (1942).
Lord, E., *J. Text. Inst.* **46,** T191 (1955).
Malhotra, R. K., M.S. thesis, Illinois Institute of Technology, Chicago, Ill., January 1969.
Sherony, D. F., Ph.D. thesis, Illinois Institute of Technology, Chicago, Ill., June 1969.
Spielman, L., Ph.D. thesis, University of California, Berkeley, 1967.
Spielman, L., Goren, L., *Environ. Sci. Technol.* **2** (4), 279–87 (1968).
Sullivan, R. R., *J. Appl. Phys.* **12,** 503–8 (June 1941).
Sullivan, R. R., Hertel, K. L., *J. Appl. Phys.* **11,** 761–5 (December 1940).
Wiggins, E. J., Campbell, W. B., Maass, O., *Can. J. Res.* **178,** 318 (1939).
Wong, J. B., Ranz, W. E., Johnstone, H. F., *J. Appl. Phys.* **27,** 161 (1956).

RECEIVED for review September 29, 1969
ACCEPTED July 22, 1970

Work supported through research grants WP 1452-01 and 12050 DRC, Federal Water Pollution Control Administration.