

A THEORY OF MAXIMUM ENTROPY  
MIXING IN ESTUARIES

Dominic M. Di Toro

A DISSERTATION  
PRESENTED TO THE  
FACULTY OF PRINCETON UNIVERSITY  
IN CANDIDACY FOR THE DEGREE  
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE BY THE  
DEPARTMENT OF CIVIL AND GEOLOGICAL ENGINEERING

May, 1967

Copyright by  
DOMINIC M. DI TORO  
1967

## DEDICATION

I dedicate this dissertation to my Mother,  
whose love and care helped me to grow;  
to my Father, whose wisdom and guidance  
formed my mind; and to my dear wife Marilyn.

## ACKNOWLEDGMENT

I wish to express my sincere appreciation and thanks to my mentor, Professor Roger J. M. DeWiest for the many expressions of his kindness and concern; for his support and encouragement throughout our association; and for his helpful guidance and advice during the period of this research.

In our many discussions of various hydrologic problems, his enlightening comments and his patient instruction have contributed much to my intellectual growth and awareness.

My interest in the analysis of the behavior of rivers and estuaries was initiated during my college years under the enthusiastic and patient tutelage of Professor Donald J. O'Connor. During our many discussions his insights and critical observations have greatly contributed to my understanding of this subject. For his guidance and his counsel, for his suggestions and comments on this thesis, I am deeply grateful.

Also I wish to express my appreciation and gratitude to Dr. Francis C. McMichael for our many interesting discussions concerning both the subject of this research and other related problems; and for his helpful and critical commentary on this thesis.

Thanks is also due Dr. David W. Moddy of the U.S. Geological Survey for his assistance in acquiring the necessary data for the numerical calculations.

This research was sponsored by the National Science Foundation, NSF Contract GK-33 and use was made of computing facilities supported in part by a National Science Foundation Grant NSF-GP 579.

## TABLE OF CONTENTS

I.	INTRODUCTION .....	1
II.	PREVIOUS THEORIES OF ESTUARINE MIXING ....	4
	A. The Convective Diffusion Equation .....	4
	B. The Reduction to One Spatial Dimension..	6
	C. The Tidal Velocity.....	7
	D. Other Theories of Estuarine Mixing.....	12
III.	MARKOV CHAIN REPRESENTATION OF MIXING IN ESTUARIES.....	17
	A. Preliminaries.....	17
	B. The Conservation Laws.....	19
IV.	THE MAXIMUM ENTROPY ESTIMATE OF THE MIXING MATRIX.....	29
	A. Statistical Mechanics and the Maximum Entropy Estimate.....	29
	B. Information Theory and the Maximum Entropy Estimate.....	36
	C. The Maximum Entropy Estimation Equations and Their Solution.....	42
V.	NUMERICAL RESULTS OF A MAXIMUM ENTROPY ESTIMATE.....	47
	A. The Data Requirements.....	47
	B. The Delaware River Model Data.....	48
	C. The Maximum Entropy Estimate of the Mixing Process.....	55
VI.	RECAPITULATION .....	73
	APPENDIX I .....	78
	BIBLIOGRAPHY.....	88
	ABSTRACT.....	95

## LIST OF TABLES

<u>Number</u>	<u>Description</u>	<u>Page</u>
1.	The Volumes of the Segments	53
2.	The Elements of the Translation Matrix $\underline{T}(q)$ for $q = 5000$ cfs and $q = 16,475$ cfs.	53
3.	Interpolated Salinity Data, Delaware River Model, High Water Slack.	54
4.	The Mixing Matrix $\underline{P}$ Estimated from the Data at $q = 5000$ cfs.	56
5.	The Mixing Matrix $\underline{P}$ Estimated from the Data at $q = 13,000$ cfs.	66

## LIST OF FIGURES

<u>Number</u>	<u>Description</u>	<u>Page</u>
1.	Cross Sectional Area Delaware River	52
2.	One Cycle Distributions of The Maximum Entropy Estimate (estimated at $q = 5000$ cfs.)	57
3.	Tidal Velocity, Channel Station 315	60
4.	Comparison of Theory (estimated at $q = 5000$ cfs.) and data $q = 5000$ cfs.	62
5.	Comparison of Theory (estimated at $q = 5000$ cfs.) and data $q = 7000$ cfs.	62
6.	Comparison of Theory (estimated at $q = 5000$ cfs.) and data $q = 9000$ cfs.	63
7.	Comparison of Theory (estimated at $q = 5000$ cfs.) and data $q = 10,600$ cfs.	63
8.	Comparison of Theory (estimated at $q = 5000$ cfs.) and data $q = 13,000$ cfs.	64
9.	Comparison of Theory (estimated at $q = 5000$ cfs.) and data $q = 16,475$ cfs.	64
10.	One cycle Distributions of the Maximum Entropy estimate (estimated at $q = 13,000$ cfs.)	67

11.	Comparison of Theory (estimated at q = 13,000 cfs.) and data q = 5000 cfs	69
12.	Comparison of Theory (estimated at q = 13,000 cfs.) and data q = 7000 cfs	69
13.	Comparison of Theory (estimated at q = 13,000 cfs.) and data q = 9000 cfs	70
14.	Comparison of Theory (estimated at q = 13,000 cfs.) and data q = 10,600 cfs	70
15.	Comparison of Theory (estimated at q = 13,000 cfs.) and data q = 13,000 cfs	71
16.	Comparison of Theory (estimated at q = 13,000 cfs.) and data q = 16,475 cfs	71
17.	Flow Graph of the Original N State Markov Chain	79
18.	Flow Graph of the Modified Markov Chain	79

## CHAPTER I

### INTRODUCTION

The analysis of the mixing which occurs in the tidal reach of an estuary has been and continues to be of concern to many scientists and engineers. For example, in order to predict the behavior of a pollutant after it has been introduced into an estuary, a characterization of the mixing to which the pollutant will be subjected is essential. Also, as first pointed out by O'Connor (1960),<sup>[1]</sup> the prediction of the dissolved oxygen distribution in an estuary involves the mixing process in a fundamental way. Hence, in order to predict the effects of pollution loadings on the water quality of an estuary, a theoretical description of the mixing process is needed.

Another phenomenon which is a result of the mixing process in an estuary is the intrusion of the salt water from the mouth of the estuary. The presence of salinity in the estuary makes it convenient to choose the concentration of salt as the dependent variable in an analysis of the mixing process. Since the only significant source of salinity is the ocean water at the mouth of the estuary, and since salt is a conservative substance, the only phenomena which must be considered in an analysis of the salinity distribution are those phenomena associated with the mixing process.

Depending on the nature of the mixing in the estuary and the resulting distribution of saline water, estuaries are classified into four categories.<sup>[2]</sup> A "vertically well-mixed" estuary exhibits no measurable difference in salinity concentration between the surface and the bed of the estuary. In a "slightly stratified" estuary a small difference in salinity exists over the vertical direction. A "highly stratified" estuary exhibits a large vertical variation in salinity. And, finally, in a salt-wedge estuary there is a clear interface between the salt and the fresh water.<sup>[3]</sup>

The stratification in an estuary is due to the density difference between fresh and salt water. As a result of this density difference, density currents are present which augment the mixing caused by the tidal action and the non-uniform velocity distribution.<sup>[4]</sup> Unfortunately, the structure of the mixing process due to the density currents must perforce be a function of the salinity concentration. Hence only a non-linear theory can account for this component of the mixing.

The degree of stratification can be used as a qualitative measure of the relative importance of density currents in the mixing process. If an estuary is highly stratified, then large density gradients are present and it would be expected that the effects of density currents in the mixing process are important. However, if the estuary is only slightly stratified, then the density gradients are small and it would be expected that other phenomena such as the tidal oscillation are the predominant cause of the mixing. Even if the density

currents play a role in the mixing process, for a small variation of salinity concentration, the resulting change in the structure of the mixing process can be shown to be small by a linearization argument. Hence the degree of stratification is an indication of the degree to which a linear theory can be expected to apply.

It will be assumed in the forthcoming analysis that the estuary being considered is slightly stratified. Hence the assumption that the structure of the mixing process is not a function of the salinity concentration is reasonable.

In the next chapter a review and a critique of the previous theories of estuarine mixing will be undertaken. These theories are also restricted to slightly stratified estuaries since the resulting equations which describe the salinity distribution are linear in the concentration of salt. The critique will be based on the mathematical assumptions that are made and their applicability to the mixing process in an estuary.

## CHAPTER II

### PREVIOUS THEORIES OF ESTUARINE MIXING

This chapter discusses the previous proposals that have been made for describing the mixing process in a tidal estuary. The assumptions which have been made are examined in terms of the physical situation and their plausibility is assessed.

#### II-A The Convective Diffusion Equation

Most of the previous theories of estuarine mixing have been based on the convective diffusion equation. It is important to realize the assumptions which are involved in the derivation of this equation and in the final simplified equation which is actually used. Toward this end the derivation given by Pritchard (1958)<sup>[5]</sup> for the equation governing the distribution of salt in an estuary will be outlined and the assumptions pointed out.

The basic equation of continuity for salt is:

$$(1) \quad \frac{\partial s}{\partial t} + \frac{\partial}{\partial x_i} (sv_i) = 0$$

where  $s$  is the concentration of salt (mass/unit volume),  $v_i$  is the component of the velocity of water in the  $i^{\text{th}}$  direction of the right-handed coordinate system  $x_1, x_2, x_3$ , and a summation on  $i$  is understood. The molecular dif-

fusivity of salt has been neglected. Following Reynolds, the velocity and salt concentration are expressed as a time mean denoted by an overbar plus a fluctuation denoted by a prime. The length of the time average is short relative to the period of the tidal cycle.

$$(2) \quad \begin{aligned} v_i &= \bar{v}_i + v_i' \\ s &= \bar{s} + s' \end{aligned}$$

Substituting these expressions in equation (1) and taking a time mean results in:

$$(3) \quad \frac{\partial \bar{s}}{\partial t} + \frac{\partial}{\partial x_i} (\bar{v}_i \bar{s}) = - \frac{\partial}{\partial x_i} \overline{v_i' s'}$$

where  $t$  is now a time scale longer than the interval of averaging. At this point a critical assumption must be made about the nature of the flux term due to the average of the product of the fluctuations. An analogy to molecular diffusion is made and it is assumed that:

$$(4) \quad \overline{v_i' s'} = - K_i \frac{\partial \bar{s}}{\partial x_i}$$

Substituting equation (4) into equation (3) results in the three dimensional convective diffusion equation:

$$(5) \quad \frac{\partial \bar{s}}{\partial t} + \bar{v}_i \frac{\partial \bar{s}}{\partial x_i} = \frac{\partial}{\partial x_i} \left\{ K_i \frac{\partial \bar{s}}{\partial x_i} \right\}$$

It should be realized that equation (4) is not a definition of  $K_i$ , the turbulent eddy diffusivity. It is a statement about a physical situation which may or may not be true. It asserts a relationship between two quantities

$\overline{v_i s}$  and  $\partial \bar{s} / \partial x_i$  which are both well defined physically and, at least theoretically, can be measured independently. Hence this equation is an assumption that such a relationship exists. It is an assumption, however, which has been questioned by Batchelor and Townsend (1956)<sup>[6]</sup>:

"... the usual argument, valid for molecular diffusion, [is] that the flux of concentration is proportional to the local concentration gradient. It is essential for this argument that the change in mean concentration over a mean free path be a small quantity, whereas in the case of turbulent diffusion no such restriction can be assumed."

## II-B The Reduction of One Spatial Dimension

The solution of the three dimensional form of the convective diffusion equation, with arbitrary velocity distributions and turbulent eddy diffusion coefficients, is unknown. In fact, as will be seen, even the one dimensional form of this equation when applied to estuaries is still too difficult to be solved analytically.

In order to obtain a one dimensional equation it is necessary to assume that  $v_1$ ,  $s$ , and  $K_1$  are not functions of  $x_2$  and  $x_3$ , i.e., that the estuary is sectionally homogeneous. For this case the resulting convective diffusion equation is:

$$(6) \quad \frac{\partial \bar{s}}{\partial t} + v_1 \frac{\partial \bar{s}}{\partial x_1} = \frac{1}{A} \frac{\partial}{\partial x_1} \left\{ A K_1 \frac{\partial \bar{s}}{\partial x_1} \right\}$$

where  $A$  is the cross sectional area. It is incorrect to

presume that this equation relates the average value of  $s$  over a section to the average values of  $v_1$  and  $K_1$ . If this equation is applied in terms of average values of  $v_1$  and  $s$  then:

"... no physical meaning can be attached to the coefficient  $K$ ... it merely becomes that function which when included in equation (6) allows that equation to describe properly the distribution in time and space of the mean salt content." [7]

The two velocity components which are predominant in an estuary are the fresh water velocity,  $Q_f/A$ , and the tidal velocity,  $Q_T/A$ , where  $Q$  is a volumetric flow rate (volume/unit time). Hence equation (6) becomes:

$$(7) \quad \frac{\partial \bar{s}}{\partial t} + \left\{ \frac{Q_f}{A} + \frac{Q_T(t)}{A} \right\} \frac{\partial \bar{s}}{\partial x_1} = \frac{1}{A} \frac{\partial}{\partial x_1} \left\{ A K_1 \frac{\partial \bar{s}}{\partial x_1} \right\}$$

Even for a simplified situation, for example a harmonic tidal velocity, constant fresh water discharge, simple geometry and simple variation of  $K_1$  as a function of  $x_1$  and  $t$ , no analytical solution is available for the appropriate initial and boundary conditions. The primary analytical difficulty is the harmonic tidal velocity and the consequent harmonic variation of  $K_1$  which would be expected.

### II-C The Tidal Velocity

Two different methods have been suggested to circumvent this difficulty. The first approach to the problem of the fluctuating tidal velocity is to average over a tidal cycle as well as over the cross sectional area. [8,9,10,11] This

argument is meticulously carried out by Okubo (1964). [12] He expresses the variables as a mean plus a fluctuation. Cross sectional averages are denoted by the brackets  $\langle \rangle$  and fluctuations by a subscript  $d$ ; temporal averages by an overbar and fluctuations by a prime. The temporal average is taken over a tidal cycle otherwise "for a time scale small in comparison with one tidal cycle, the averaged quantities would vary in time and space in ways too complicated for easy analytic manipulation." [13] As in the case of turbulent diffusion the flux terms due to the fluctuations about the cross sectional average and the tidal fluctuations are related to the gradient of the mean concentration:

$$(8) \quad -K_1 \frac{\partial \bar{s}}{\partial x_1} = \overline{v_1' s'} + \langle v_{1d} s_d \rangle$$

The "effective tidal mean" velocity is given as the net velocity over the cross sectional area, averaged over the tidal cycle:

$$(9) \quad \bar{v}_{1e} = \frac{\iint_A v_1 dA}{\bar{A}}$$

This velocity is related to the fresh water velocity since the harmonic terms in the tidal velocity tend to average out over a tidal period. Hence the resulting equation is substantially the one dimensional convective diffusion equation for the concentration averaged over a tidal period but without a tidal velocity term. However, two of the assumptions used to arrive at this equation are that the "tidal

fluctuations of the cross sectional area, concentration and density are sufficiently small compared with the respective mean values..."[14] In other words the effects of the tidal oscillations and the fluctuations about the cross sectional average are assumed to be small. And secondly the assumption embodied in the flux law (equation (8)) must also be true. As a practical matter, however, the fluctuation of salinity concentration over a tidal cycle in the tidal portion of an estuary is usually large relative to the mean [15,16] and therefore cannot be regarded as a small fluctuation. Hence if the convective diffusion equation is averaged over a tidal cycle and the various terms due to the tidal fluctuation are dropped, the equation cannot be expected to apply to tidal estuaries.

The second approach to the analytical problem of the tidal velocity is formally to ignore the tidal velocity term completely in the convective diffusion equation. [17,18] The estuary is considered only at a progression of high or low water slacks when the tidal velocity is zero. Harleman (1964) [19] gives an explanation of this approach based on the idea of a longitudinal dispersion coefficient:

"The time scale of equation (7) [is changed] to one in which the smallest unit of time is the tidal period.... In the new time scale the advective term can only account for fresh water flow rate  $Q_f$ .... The various assumptions which [have been made] imply that  $K$  can no longer be interpreted as the turbulent eddy diffusivity... $K$  is

more correctly a longitudinal dispersion coefficient which must account for 1. the dispersive effect of the vertical velocity distribution 2. the dispersive effect of the unsteady tidal velocity 3. the longitudinal dispersion due to density gradients arising from the intrusion of saline water from the ocean..."[20]

It is apparent, therefore, that the view expressed by Pritchard with regard to the fluctuations about the cross sectional average is extended by Harleman to include the effects of the tidal velocity as well as the effects due to the deviation from one dimensionality. It should also be noted that the resulting concentration profiles are for high or low water slack time only, and the concentration profiles obtained for a time other than these slack times have no physical meaning.

This is certainly a practical solution since the simplified equation can now be solved for a variety of geometries and functional forms of  $K$ . Also, if a steady-state salinity distribution is available, the dispersion coefficient can be easily calculated<sup>[21]</sup> and used to predict the salinity distributions at other fresh water flows. Furthermore the problem of obtaining enough data in order to compute averages over a tidal cycle does not arise since only the slack time distributions are considered.

It is not clear, however, that the argument which justifies the use of the eddy diffusion coefficient can be used to justify the concept of the longitudinal dispersion coefficient.

The concept of eddy diffusion as employed in the derivation of the convective diffusion equation in three dimensions states that the flux due to the fluctuations is proportional to the gradient of the average concentration. This diffusion law is conventionally employed in the analysis of heat diffusion (Fourier's Law) and molecular diffusion (Fick's Law). In both these phenomena the temporal and spatial scales of the underlying fluctuations are assumed to be small relative to the gross phenomena being considered.<sup>[22]</sup> The question is whether or not the same equation for flux is valid when the temporal and spatial scales of the fluctuations are on the same order of magnitude as the gross phenomena of interest. This is the situation when the dispersive effect of the tidal oscillation is included in the longitudinal dispersion coefficient. The time scale of the fluctuation is one tidal period and the spatial scale is one tidal excursion length.

During one tidal cycle the net flux at a point  $x_0$  is related to the concentrations within one tidal excursion length,  $L$ , of  $x_0$ . Within this region the concentrations interact and a new concentration distribution results at the next slack time. There seems to be no a priori reason to suppose that the net flux at  $x_0$  during the tidal cycle is related to the gradient of the concentration at  $x_0$  only since water with concentrations  $s(x)$  over a length  $x_0 - L \leq x \leq x_0 + L$  can be expected to mix with water at  $x_0$ . If during one tidal cycle the mixing is supposed to take place over a tidal excursion length then a more general mixing law which relates concentra-

tions at a distance must be used to describe the phenomena. During an analysis of pollution in the Thames Estuary<sup>[23]</sup> this point of view was adopted. It was also demonstrated how the convective diffusion equation results from the more general law as a first approximation. The degree of approximation depends on the length of the tidal excursion, that is, the scale of the fluctuation. More precisely it depends on how closely the flux term due to the fluctuations is approximated by the gradient of the mean concentration. Since the tidal excursion length in an estuary is usually of the same order of magnitude as the length over which the distributions of interest extend,<sup>[24]</sup> the convective diffusion equation, when interpreted in terms of a longitudinal dispersion coefficient and slack distributions, is not an adequate representation of the tidal mixing in estuaries.

#### II-D Other Theories of Estuarine Mixing

Ketchum (1951)<sup>[25]</sup> has proposed a model for the tidal mixing in an estuary which is based on the physical characteristics of the estuary. He divided the estuary into segments whose lengths are the average excursion of a particle of water on the flooding tide. He then proposed a mechanism of mixing during each tidal cycle: At high tide the water within each segment mixes completely; then there is an exchange between adjacent segments during the ebb tide. The exchange coefficients are given by the ratio of the intertidal volume to the high tide volume.

This theory is quite ambitious in that it attempts to predict the mixing phenomena to be expected from the physical characteristics of the estuary alone. The theory has been used to compute the equilibrium salinity profiles in the Delaware and Thames Estuaries but these calculated distributions do not agree with the observed distributions. [26] However, this theory was the first attempt to describe the nature of the mixing process in a tidal estuary and as such was a significant improvement over the previous tidal prism theories.

The evident complexity of the mixing process in an estuary led Preddy (1954) [27] to propose a less ambitious theory but one with more chance of success than Ketchum's attempt. His idea was to specify the form of the one cycle distribution function. Specifically he assumed that for a unit amount of water at  $x_0$ , after one tidal cycle a proportion  $P_2(x_0)$  of the water would be spread uniformly a distance  $L$  in the seaward direction, and a proportion  $P_1(x_0)$  would be spread uniformly in the landward direction a distance  $L$ , and  $1 - P_2(x_0) - P_1(x_0)$  would remain at  $x_0$ . He then applied two continuity laws:

"(i) During any period the net amount of salt carried upstream past any point is equal to the amount present above this point at the end of the period, minus the amount present at the beginning....

"(ii) The volume of water carried upstream past the point during the period must conform to a similar condition." [28]

These conditions when expressed mathematically resulted in two integral equations in  $P_1(x)$  and  $P_2(x)$ :

$$(i) \int_0^L s(x)A(x)P_2(x) \frac{L-x}{L} dx + \int_0^{-L} s(x)A(x)P_1(x) \frac{L+x}{L} dx$$

$$(10) \qquad \qquad \qquad = Q s(0) + S$$

$$(ii) \int_0^L A(x)P_2(x) \frac{L-x}{L} dx + \int_0^{-L} A(x)P_1(x) \frac{L+x}{L} dx = 0$$

where  $S$  is the net accumulation of salt above  $x = 0$  and  $Q$  is the fresh water flow. Equation (i) was then averaged over all the tidal cycles for which there is data. The integral equations were solved by approximating  $P_1(x)$  and  $P_2(x)$  by linear expressions in  $x$  and solving the resulting simultaneous linear equations by relaxation. Greater accuracy was achieved by using quadratic functions as approximations to  $P_1$  and  $P_2$ . The distance  $L$  is specified a priori. It should be of the same order of magnitude as the average tidal excursion distance. However:

"If the values of  $L$  taken lead to calculated values of  $P_1 + P_2$  which are greater than unity... it follows that the mixing in this section is substantial over a greater distance than the assumed value of  $L$ . Hence  $L$  should be increased in this section and the values of  $P_1$  and  $P_2$  are exceedingly small... it is better to decrease  $L$  and recalculate  $P_1$  and  $P_2$ ." [29]

Once  $P_1(x)$  and  $P_2(x)$  are known, they describe the mixing phenomena completely. Preddy used this formulation to

predict the change in the salinity distributions in the Thames Estuary under varying fresh water flows. The results compared very well with the measurements taken.

This formulation was also used in an extensive analysis of the pollution distribution in the Thames.<sup>[30]</sup> For this analysis, however, the estuary was segmented and the mixing parameters were averaged over each segment. It will be seen in the following chapter that Preddy's theory is more naturally expressible in a discrete spatial domain. Perhaps the unwieldy nature of equations (i) and (ii) and the resulting relaxation formulas explain in part, at least, the fact that this theory has not found favor among workers interested in the analysis of the mixing process in an estuary. However, this type of approach has been recognized as a fruitful area for further investigation.<sup>[31]</sup>

There are some objections which can be raised immediately with regard to the a priori nature of the structure of the theory. The form of the one cycle distribution function was chosen because of its simplicity. Preddy claims that: "...it is not possible to calculate the exact shape of the [distribution] curves from the consideration of the salinity balance during a long time..."<sup>[32]</sup> Perhaps not, but it will be shown subsequently that a much better estimate can be made on the basis of the salinity profiles available. Also the status of the length  $L$  is in doubt. It too is specified a priori and this is another assumption which adds to the uncertainty about the correctness of the procedure.

Nevertheless the approach taken by Preddy with some important modifications and a change in mathematical setting forms the basis of the theory of estuarine mixing to be presented in the following chapters. In the light of the objections which have been raised to theories based on simplified versions of the convective diffusion equations, it is appropriate to adopt instead the point of view expressed by Preddy as the foundation for a theory of estuarine mixing.

## CHAPTER III

### MARKOV CHAIN REPRESENTATION OF MIXING IN ESTUARIES

In this chapter the theory of Markov chains will be applied to the analysis of the mixing process in an estuary. The conservation laws which any mixing process must satisfy in order to be physically meaningful will be derived. These conservation laws can be readily interpreted in terms of the properties of a Markov chain.

#### III-A Preliminaries

The theory of Markov chains has proved to be a useful tool for the theoretical and computational analysis of diffusion phenomena.<sup>[33]</sup> In particular the analysis of random walks, for which the particle is constrained to move only to adjacent states during a unit time interval, has been used to model the convective diffusion equation.<sup>[34]</sup> The usual procedure is to assume a priori the structure of the transition probability matrix and then, starting with an initial distribution, compute the succeeding distributions.

It is interesting to note that the random walk is the discrete analog of the diffusion equation. More general Markov chains, which allow transition to any state during the unit time interval, are capable of describing more general mixing processes. This increase in generality is one reason that a Markov chain description is chosen for the

present analysis of the mixing process in an estuary.

A second reason for choosing a Markov chain is that there is an interval of time, the tidal period  $\tau$ , which usually separates the data. It is necessary, therefore, to relate conditions at succeeding slack water times. For the sake of clarity it is assumed that the slack water time at which the data is available is the high water slack. With discrete time intervals between the distributions the natural mathematical setting is within the framework of the theory of Markov chains.

Unfortunately a Markov chain relates the concentration only at discrete states so that the estuary must be segmented into discrete volumes. Hence the theory will attempt to relate the concentration of salt between discrete volumes of the estuary at successive high water slacks. In order to specify a definite concentration in each segment, it is assumed that the concentration of salt is essentially uniform within each segment. This assumption can be made more or less physically consistent with the facts by choosing the appropriate size segments. For the sake of simplicity a one dimensional segmentation of the estuary will be made. It will be clear subsequently that the theory is not limited to a one dimensional analysis. The segments are numbered sequentially in the downstream direction and the volumes at high water slack are  $v_1, v_2, \dots, v_N$  for the  $N$  segments.

The mixing process will be thought of as being the result of two phenomena: the fresh water flow; and the mixing

associated with the tidal oscillation. The mixing due to the tidal oscillation will be represented by the following mixing parameters: Let  $p_{ij}$  be the ratio of the volume of water which is transported during one tidal cycle to segment  $i$  from segment  $j$ , to the total volume of segment  $j$ . Then it is true that :

$$(11) \quad 0 \leq p_{ij} \leq 1 \quad i = 1, \dots, N; \quad j = 1, \dots, N$$

It will be assumed that these parameters do not vary from tidal cycle to tidal cycle. Let  $\underline{P}$  be the mixing matrix with elements  $p_{ij}$ . The mixing matrix represents the tidal mixing in the estuary and its numerical evaluation is the object of this analysis.

### III-B The Conservation Laws

In order for the mixing process to be physically meaningful certain conservation laws must be satisfied. These laws impose certain restrictions on the mixing matrix  $\underline{P}$ . The laws are conservation of water, conservation of the estuary geometry, and conservation of salt.

Conservation of water states that during a tidal cycle all the water from segment  $j$  must go to either another segment or remain in  $j$ . Hence it must be true that:

$$(12) \quad \sum_{k=1}^N p_{kj} = 1 \quad j = 1, \dots, N$$

In terms of the theory of Markov chains, a matrix which satisfies equations (11) and (12) is a transition probability matrix.

The conservation of the geometry of the estuary requires that there be no net accumulation of water upstream from any segment boundary over one tidal cycle. This is condition (ii) as given by Preddy. If this condition is not fulfilled then it is possible that during the tidal cycle there is a net migration of water across a segment boundary. This migration of water would change the water level in the estuary. But it is a physical fact that the water level in an estuary is constant at successive high water slack times.

The following theorem gives a more suggestive formulation of this conservation law.

Theorem: There is no net accumulation of water upstream from any segment boundary if and only if

$$(13) \quad \underline{Pv} = v$$

Proof: Let  $\phi$  be the boundary between segment  $k - 1$  and segment  $k$ . In order to evaluate the net transfer across  $\phi$  during one tidal cycle, consider a segment  $i$  upstream of  $\phi$  and a segment  $\mu$  downstream of  $\phi$ . Then the volume of water going into  $i$  from  $\mu$  is  $p_{i\mu}v_{\mu}$ . Hence the total volume of water going into  $i$  from segments downstream of  $\phi$  is:

$$(14) \quad \sum_{\mu=k}^N p_{i\mu}v_{\mu} .$$

By the same argument the total volume of water leaving segment  $i$  and going downstream beyond  $\phi$  is:

$$(15) \quad \sum_{v=k}^N p_{vi}v_i .$$

Hence the net transfer upstream into  $i$  across  $\phi$  from all segments downstream of  $\phi$  is:

$$(16) \quad \sum_{\mu=k}^N (p_{i\mu} v_{\mu} - p_{\mu i} v_i)$$

And finally the net transfer upstream across  $\phi$  during one tidal cycle is:

$$(17) \quad \sum_{i=1}^{k-1} \sum_{\mu=k}^N (p_{i\mu} v_{\mu} - p_{\mu i} v_i) = 0$$

which must be zero by hypothesis. Interchanging the order of summation and using equation (12) in the first term gives:

$$(18) \quad \sum_{\mu=k}^N v_{\mu} \left(1 - \sum_{i=k}^N p_{i\mu}\right) - \sum_{\mu=k}^N \sum_{i=1}^{k-1} p_{\mu i} v_i = 0.$$

However the second term in this equation can be rewritten as:

$$(19) \quad \sum_{\mu=k}^N v_{\mu} \sum_{i=k}^N p_{i\mu} = \sum_{i=k}^N \left( \sum_{\mu=1}^N p_{i\mu} v_{\mu} - \sum_{\mu=1}^{k-1} p_{i\mu} v_{\mu} \right)$$

by interchanging the order of summation and adding and subtracting  $\sum_{\mu=1}^{k-1} p_{i\mu} v_{\mu}$ . But the second term on the right-hand side of equation (19) is the same as the third term in equation (18). Hence equation (18) becomes:

$$(20) \quad \sum_{\mu=k}^N v_{\mu} - \sum_{i=k}^N \sum_{\mu=1}^N p_{i\mu} v_{\mu} = 0$$

or

$$(21) \quad \sum_{i=k}^N \left[ v_i - \sum_{\mu=1}^N p_{i\mu} v_{\mu} \right] = 0$$

If this equation is true for a boundary at  $k$  it must also be true for a boundary at  $k + 1$ . Hence:

$$(22) \quad \sum_{i=k+1}^N \left[ v_i - \sum_{\mu=1}^N p_{i\mu} v_{\mu} \right] = 0$$

and subtracting equation (22) from equation (21) gives:

$$(23) \quad \sum_{\mu=1}^N p_{k\mu} v_{\mu} = v_k$$

Since this equation is true for an arbitrary  $k$ , it is true for all  $k = 1, \dots, N$ . In which case, equation (23) can be expressed in matrix form:

$$(24) \quad \underline{P} \underline{v} = \underline{v}$$

where  $\underline{P}$  is the mixing matrix and  $\underline{v}$  is a column vector with elements  $v_k$ . Hence the theorem is proved.

In order to interpret this result it is necessary to realize that multiplication by the mixing matrix  $\underline{P}$  is the operation which converts the volume distribution at time  $n\tau$  to that which is expected at time  $(n + 1)\tau$ . This follows directly from the definition of  $\underline{P}$ . Hence equation (24) states that if  $\underline{v}$  is the observed volume distribution of water in the estuary at high water slack, then the mixing caused by a tidal oscillation must produce the same volume distribution at the next high water slack. This is exactly the intuitive meaning of conservation of the geometry of an estuary.

In terms of the theory of Markov chains, this conservation law specifies the equilibrium distribution of the chain  $\underline{P}$ . That an equilibrium distribution exists is guaranteed if certain structural conditions are fulfilled. [35] However a sufficient analytical condition on  $p_{ij}$  is that  $p_{ij} > 0$  for all  $i$  and  $j$ . [36] A less stringent sufficient condition, called the case of Hostinsky, guarantees the existence of an equilibrium distribution if the elements of the main diagonal and the two adjacent off diagonals are greater than zero. [37] It will be seen that these conditions suffice for the present analysis.

The law of conservation of salt will be derived from the following observation. If the fresh water flow rate  $q$  is constant over many tidal cycles, then the salinity distribution in the estuary will reach an equilibrium distribution. Under this condition there can be no net accumulation of salt above any segment boundary.

In order to express the law of conservation of salt mathematically, the mixing process which affects the salt must be considered. The water in each segment contains a certain amount of salt and this water is mixed with the water in the other segments by the tidal action. However, the water in each segment is also affected by the fresh water flow. Hence, strictly speaking, the mixing parameters  $p_{ij}$  are functions of the fresh water flow. However this dependence is unknown and it has been assumed that  $\underline{P}$  is only a result of tidal action. In order to incorporate the effect

of fresh water flow, another matrix will be used whose elements are functions of the flow rate. It will be assumed that the only effect of the fresh water flow is to translate the water downstream. Let  $q_j$  be the volumetric flow rate entering segment  $j$ . Then during a tidal period  $\tau$  the fresh water volume entering segment  $j$  is  $q_j\tau$ . Assuming that:

$$(25) \quad q_j\tau \leq v_j$$

the proportion of water left in the  $j^{\text{th}}$  segment is  $1 - q_j\tau/v_j$  and a proportion  $q_j\tau/v_j$  moves to segment  $j + 1$ . Hence let:

$$(26) \quad t_{j,j} = 1 - \frac{q_j\tau}{v_j} \quad j < N$$

and

$$(27) \quad t_{j+1,j} = \frac{q_j\tau}{v_j} .$$

For all other elements in column  $j$ :

$$(28) \quad t_{ij} = 0 .$$

For  $j = N$ , the last segment, let:

$$(26a) \quad t_{N,N} = 1$$

and

$$(28a) \quad t_{i,N} = 0 \quad i = 1, \dots, N - 1$$

Define  $\underline{T}(\underline{q})$ , the translation matrix, to be the matrix of elements  $t_{ij}$ . It is a function of the vector  $\underline{q}$  with components  $q_i$ , the fresh water flow rates.

The definition of  $\underline{T}(\underline{q})$  is easily extended for the case when equation (25) is not satisfied.

The edge condition which has been assumed at segment N, is that no fresh water leaves segment N by translation. Although this condition is not physically justifiable it is made in order that:

$$(29) \quad \sum_{k=1}^N t_{kj} = 1 \quad j = 1, \dots, N$$

This equation is also satisfied by  $p_{ij}$ . Hence there is no transfer of water, and hence, salt, either out of or into the finite length of the estuary being analyzed by either the tidal mixing or the fresh water flow. This is a simplification of the actual boundary condition which should be applied at segment N. As a practical matter, it will be shown in Appendix I that if segment N is at a point far enough downstream so that the salinity concentration is at the ocean value, then this simplification is of no moment. The inclusion of a realistic boundary condition would complicate the analysis considerably since the matrices would no longer be transition probability matrices.

The composite effect of tidal mixing and fresh water flow will be given by the mixing process:

$$(30) \quad \underline{\underline{PT}}(\underline{q}) \quad .$$

That is, the water is translated first by  $\underline{T}(\underline{q})$  and then mixed by tidal action. These processes occur simultaneously and, as remarked earlier,  $\underline{P}$  should be given as a function

of  $\underline{q}$ , that is  $\underline{P} = \underline{P}(\underline{q})$ . It is hypothesized therefore that the dependence of  $\underline{P}$  on  $\underline{q}$  is given by:

$$(31) \quad \underline{P}(\underline{q}) = \underline{\underline{PT}}(\underline{q}) .$$

The expression for the law of conservation of salt at steady state follows from an argument which is identical to that of the preceding theorem. The result is that for an equilibrium salinity distribution,  $\underline{sv}$ , at a fresh water flow  $\underline{q}$ , it must be true that :

$$(32) \quad \underline{\underline{PT}}(\underline{q}) \underline{sv} = \underline{sv}$$

where  $\underline{sv}$  is the vector with components  $(s_i, v_i)$  and  $s_i$  is the concentration of salt in segment  $i$ . Since  $\underline{\underline{PT}}(\underline{q})$  is the transition probability matrix which affects the salinity distribution, this theorem states that the distribution  $\underline{sv}$  is the equilibrium distribution of the Markov chain  $\underline{\underline{PT}}(\underline{q})$ . Furthermore, for any given salinity distribution  $\underline{sv}(n\tau)$  at high water slack time  $n\tau$ , and for a fresh water flow rate  $\underline{q}_n$  during the subsequent tidal cycle, the resulting salinity distribution  $\underline{sv}[(n+1)\tau]$  is given by

$$(33) \quad \underline{\underline{PT}}(\underline{q}_n) \underline{sv}(n\tau) = \underline{sv}[(n+1)\tau] .$$

Hence the three conservation laws require that the following three equations be satisfied:

$$(34) \quad \sum_{k=1}^N p_{kj} = 1 \quad j = 1, \dots, N$$

$$(35) \quad \underline{\underline{Pv}} = \underline{v}$$

$$(36a) \quad \underline{\underline{PT}}(\underline{q}) \underline{sv} = \underline{sv}$$

or

$$(36b) \quad \underline{PT}(q_n) \underline{sv}(n\tau) = \underline{sv}[(n+1)\tau] .$$

Equation (36a) applies if the available salinity distribution is an equilibrium distribution whereas equation (36b) applies for any two successive distributions. Equation (36b) can be averaged over all the available salinity data in order to give more statistically reliable coefficients in the equation.

Together equations (34), (35) and (36) specify  $3N$  linear constraints which the  $N$  by  $N$  matrix  $\underline{P}$  must satisfy. However for  $N > 3$ ,  $N^2 > 3N$ . Therefore there are more unknown  $p_{ij}$ 's than there are linear equations (34), (35) and (36). Hence there are an infinite number of possible solutions  $\underline{P}$  which satisfy the conservation laws. The problem which must be solved is which of these matrices is most appropriate.

In this context, Preddy's solution of this problem involved limiting the number of unknowns to  $3N$  by specifying a priori the structure of the matrix  $\underline{P}$ . He assumes that the  $j^{\text{th}}$  column of  $\underline{P}$  is given in terms of two unknowns  $\alpha_j$  and  $\beta_j$ :

$$\text{for } i > j, \quad |i - j| < L_j$$

$$p_{ij} = \alpha_j$$

$$\text{for } j > i, \quad |i - j| < L_j$$

$$p_{ij} = \beta_j$$

for  $i = j$

$$p_{ij} = 1 - \alpha_j - \beta_j$$

and for  $|i - j| > L_j$

$$p_{ij} = 0$$

where  $L_j$  is also specified a priori.  $L_j$  is the distance over which the one cycle distribution extends upstream and downstream from segment  $j$ . Equation (34) has been satisfied by the definition of  $p_{ij}$ . Equations (35) and (36) are then used to solve for the  $\alpha_j$ 's and  $\beta_j$ 's.

The difficulty with this solution is the arbitrary nature of the assumptions used to make the problem solvable.

In the next chapter a more rational approach will be presented which bases the choice of  $\underline{P}$  from the infinity of possibilities on the maximum entropy principle of information theory and statistical mechanics.

## CHAPTER IV

### THE MAXIMUM ENTROPY ESTIMATE OF THE MIXING MATRIX

In this chapter the maximum entropy estimate of the mixing matrix will be discussed. The use of the maximum entropy estimate of a discrete probability distribution has been justified by Jaynes<sup>[38]</sup> in terms of a subjective interpretation of the concept of probability. He has applied these ideas to problems in statistical estimation theory<sup>[39]</sup> and to conceptual problems in statistical mechanics.<sup>[40]</sup> The application of the maximum entropy estimate to Markov chains in general and to mixing processes has not previously been attempted.

#### IV-A Statistical Mechanics and the Maximum Entropy Estimate

In order to arrive at a more acceptable theory of tidal mixing in estuaries, the choice of the mixing matrix  $\underline{P}$  must be based on some rational estimation procedure. The estimation problem may be stated in geometrical terms. There are  $N^2$  unknown  $p_{ij}$  to be determined. A particular value of these  $N^2$  unknowns can be thought of as a point in  $N^2$  dimensional Euclidean space. This point is within a unit hypercube since each  $p_{ij}$  is bounded between 0 and 1. However, the conservation laws specify  $3N$  independent linear constraints. Hence the admissible  $p_{ij}$ 's are restricted to a  $N^2 - 3N$  dimensional linear manifold of the  $N^2$  dimensional space. Within

this manifold there is no a priori reason to prefer one point over another. Hence a reasonable procedure is to take the average value of all the P's in the linear manifold giving equal weight to each point. The resulting average P satisfies the requirements of the conservation laws and it does not have any property that is not shared by the majority of all the possible processes.

Another reasonable way to estimate P is to find the most probable P in the linear manifold.

Both these calculations can be carried out for a simplified situation and the results of the analysis of the simplified situation will point to a technique which can be generalized to include situations where the probabilities are specified by a Markov chain.

Consider the following situation: at high water slack  $M$  particles of dye are introduced into a particular segment  $k$  of the estuary. This segment is thought of as fixed. At the next high water slack only the first moment about the midpoint of segment  $k$  of the resulting distribution is measured. The problem is to find the most probable distribution consistent with this data. For the sake of simplicity, only the tidal action will be considered. As before, let  $p_{ik}$  be the proportion of water and hence dye that is transferred to segment  $i$  from segment  $k$ . Let  $m_i$  be the number of dye particles that are transferred segment  $i$  from segment  $k$ . Hence  $p_{ik}$  is given by:

$$(36) \quad p_{ik} = \frac{m_i}{M} .$$

The law of conservation of water and therefore of dye requires that:

$$(37) \quad \sum_{i=1}^N m_i = M \quad .$$

The available data is the first moment of the one cycle distribution function about the midpoint of segment  $k$  and is denoted by  $m_k$ , that is:

$$(38) \quad m_k = \frac{\sum_{i=1}^N |i-k| m_i}{\sum_{i=1}^N m_i} \quad .$$

The other conservation laws (equations (35) and (36)) are also neglected for the sake of simplicity.

Following Jaynes,<sup>[41]</sup> the method of Boltzmann will be applied to find the most probable distribution  $\{m_i\}$ . Consider all the possible arrangements of the  $M$  particles of dye in the  $N$  segments. That is, particle number 1 can go to any segment 1 through  $N$ , particle number 2 can go to any segment 1 through  $N$ , and so on up to particle  $M$ . There are  $N^M$  such possible arrangements of  $M$  particles in the  $N$  states. Each of these  $N^M$  arrangements corresponds to a final distribution of particles,  $m_1, m_2, \dots, m_N$ . However, many different arrangements of the  $M$  particles can result in the same final distribution since the particles are indistinguishable. For example, the distribution  $m_1 = M/2, m_2 = M/2, m_3, m_4, \dots, m_N = 0$  can occur if particles 1 through  $M/2$  go to segment 1, and particles  $M/2 + 1$  through  $M$  go to segment 2; or vice versa; or if odd numbered

particles go to segment 1 and even numbered particles go to segment 2, and so on. In fact, if  $W\{m_i\}$  is the number of ways  $M$  particles can be distributed in  $N$  segments, with  $m_i$  particles in segment  $i$ ,  $i = 1, \dots, N$ , then by a combinatorial argument:

$$(39) \quad W\{m_i\} = \frac{M!}{m_1! \dots m_N!} .$$

There is no a priori reason for favoring any one of the  $N^M$  possible arrangements of particles except that the resulting distribution must satisfy equations (37) and (38). Within this subset of arrangements, then, each possible arrangement has an equal probability of occurring. The most probable distribution  $\{m_i\}$  is that which is realized by the greatest number of possible arrangements. This is true because the probability of a particular distribution is the sum of the probabilities of all the arrangements which result in this distribution. Hence, in order to find the most probable distribution  $W\{m_i\}$  must be maximized subject to two constraints specified by equations (37) and (38). Maximizing  $1/M \log W$  is equivalent to maximizing  $W$  since the logarithm is a monotonic function. The logarithm is chosen in order that Sterling's approximation for the factorial can be employed. Using Sterling's approximation and equation (37) yields:

$$\begin{aligned}
 \frac{1}{M} \log W\{m_i\} &\cong \log M - 1 - \frac{1}{M} \sum_{i=1}^N (m_i \log m_i - m_i) \\
 (40) \quad &\cong \frac{1}{M} \sum_{i=1}^N m_i \log M - \frac{1}{M} \sum_{i=1}^N m_i \log m_i \\
 &\cong - \sum_{i=1}^N \frac{m_i}{M} \log \frac{m_i}{M}
 \end{aligned}$$

as  $M \rightarrow \infty$ . So the most probable distribution is found by maximizing

$$(41) \quad - \sum_{i=1}^N p_i \log p_i$$

subject to the constraints:

$$(42) \quad \sum_{i=1}^N p_i = 1$$

$$(43) \quad \sum_{i=1}^N |i - k| p_i = m_k$$

where the  $m_i$  have been replaced by  $p_i$  according to equation (36), and the subscript  $k$  has been dropped since  $k$  is fixed in this illustration.

The constraints can be incorporated into the theory of maxima and minima using the Lagrange multipliers  $\lambda$  and  $\mu$ . Hence finding the unconstrained maximum of:

$$(44) \quad - \sum_{i=1}^N p_i \log p_i - \lambda \sum_{i=1}^N p_i - \mu \sum_{i=1}^N |i - k| p_i$$

by taking the partial derivative with respect to  $p_i$  gives:

$$(45) \quad p_i = e^{-\lambda' - \mu|i-k|}$$

when  $\lambda' = \lambda - 1$ . The Lagrange multipliers are evaluated using the constraint equations (42) and (43).

$$(46) \quad \sum_{i=1}^N e^{-\lambda' - \mu|i-k|} = 1$$

$$(47) \quad \sum_{i=1}^N |i - k| e^{-\lambda' - \mu|i-k|} = m_k$$

The resulting  $p_i$ 's are the most probable mixing parameters which account for the first moment which was observed.

It is interesting to note that the most probable distribution can be found by maximizing the function  $-\sum_{i=1}^n p_i \log p_i$  subject to the appropriate constraints. This function is known as the entropy of the distribution  $\{p_i\}$ . Therefore, within the context of this illustration, the constrained maximum entropy estimate is the most probable distribution consistent with the available data.

Another technique which is used in statistical mechanics and which leads to the same results as the method of Boltzmann is the Darwin-Fowler method.<sup>[42]</sup> Instead of finding the most probable distribution as  $M \rightarrow \infty$ , the Darwin-Fowler method can be used to compute the average distribution over all the possible distributions. For a particular distribution  $\{m_i\}$ ,  $W\{m_i\}$  is the number of ways it can occur. Hence the average value of  $m_j$ , denoted by  $\langle m_j \rangle$ , is given by:

$$(48) \quad \langle m_j \rangle = \frac{\sum'_{\{m_i\}} m_j W\{m_i\}}{\sum'_{\{m_i\}} W\{m_i\}}$$

where the primed summation indicates that the sum is taken over only those distributions  $\{m_i\}$  which satisfy the constraints, equations (42) and (43). The results of the calculation show that the average over all possible distributions is the same as the most probable distribution, equation (45), as  $M \rightarrow \infty$ . The variance of  $m_j$  can also be calculated. The result, in terms of  $p_j$  is that:

$$(49) \quad \langle p_j^2 \rangle - \langle p_j \rangle^2 = \frac{p_j}{M} \left[ 1 + O\left(\frac{1}{M}\right) \right].$$

Hence the variance of  $p_j$  also goes to zero as  $M \rightarrow \infty$ . That the variance of  $p_j$  is small for the large  $M$  indicates that the probability of the observed results having been produced by a distribution other than  $\{p_j\}$  is very small.

Hence, for this illustration, a second interpretation of the maximum entropy estimate is given by the Darwin-Fowler method: the maximum entropy estimate of the distribution  $\{p_i\}$  is the average of all the distributions consistent with the constraints.

Therefore for this simple case, when only a simple discrete probability distribution  $\{p_i\}$  is considered, the average distribution over the manifold and the most probable distribution in the manifold can be calculated. The resulting distribution is the maximum entropy estimate consistent with

the constraints. In order to generalize this result to Markov chains, the meaning of the entropy of a probability distribution must be understood. This is considered in the next section.

#### IV-B Information Theory and the Maximum Entropy Estimate

Another point of view may be adopted as to which  $\underline{P}$  to choose from the infinity of possibilities. The most desirable  $\underline{P}$  would be the one which is "maximally non-committal with regard to missing information."<sup>[43]</sup> In other words the  $\underline{P}$  with the most random character possible, within the limitations of what information is available, is the most acceptable choice. What is needed, therefore, is a measure of the "amount of uncertainty" or the randomness represented by a mixing process.

Suppose a mixing process tends to smooth out irregularities and does not favor any particular segment too strongly over any other segment. Then it tends to randomize whatever initial distribution it mixes. In the extreme case, if after one tidal cycle the water from any segment is distributed uniformly over all the segments, then this process has no preference for any segments. Hence it behaves in an entirely random fashion and the measure of its randomness should be the highest possible.

On the other hand, suppose a mixing process leaves everything unchanged. The distribution is the same after the tidal cycle as before. Then this process does not randomize the initial distribution at all. In fact, this process

behaves in a deterministic way. There is no randomness associated with the process since the outcome is certain: all the water will end up in the same segments from which it started. In fact, any mixing process which just transfers the entire contents of a state into another state is a deterministic process. Therefore the measure of the randomness of such processes should be zero.

The problem of the appropriate measure for the amount of uncertainty represented by a discrete probability distribution  $\{p_i\}$  was solved by Shannon (1948).<sup>[44]</sup> The result is the entropy  $H$  of the probability distribution where:

$$(50) \quad H = - \sum_{i=1}^N p_i \log p_i$$

In the case of a Markov chain, the entropy can be defined in the following way:<sup>[45]</sup> for a particle initially in segment  $j$ , the probability distribution that is expected after one tidal cycle is  $\{p_{ij}\}$ . Hence the entropy of this distribution is:

$$(51) \quad H_j = - \sum_{i=1}^N p_{ij} \log p_{ij} .$$

However, the probability of a particle being in segment  $j$  to begin with is the equilibrium probability of the  $j^{\text{th}}$  segment. This supposes that the chain is in equilibrium, i.e., that the mixing process has been operating for a long time. Fortunately, one of the conservation laws, equation (24), specifies that the equilibrium distribution of the Markov

chain must be proportional to the volume distribution  $v$ . Therefore the entropy of the entire process can be defined as the average of  $H_j$  taken over the segment  $j$  from which the particle starts, weighted by the probability that a particle is there to begin with. That is the entropy  $H$  of the Markov chain  $\underline{P}$  is defined as:

$$(52) \quad H = - \sum_{j=1}^N v'_j \sum_{i=1}^N p_{ij} \log p_{ij}$$

$$(53) \quad v'_j = v_j / \left( \sum_{k=1}^N v_k \right)$$

where  $v'_j$  is the normalized volume.

Since the entropy of a Markov chain is a measure of the randomness of the mixing process it describes, a reasonable choice for the mixing process in an estuary is the process which has the maximum entropy among all those that satisfy the conservation laws. The maximum entropy estimate tends to favor the most random mixing possible consistent with the available data. None of the mixing parameters are set to zero a priori. In fact the maximum entropy estimate tends to choose the broadest one cycle distribution functions consistent with the data.

As an illustration of the type of processes which a maximum entropy estimate produces, consider the maximum entropy mixing process which satisfies only the law of conservation of water, equation (12):

$$(54) \quad \sum_{i=1}^N p_{ij} = 1, \quad j = 1, \dots, N.$$

The maximum entropy estimate constrained by this requirement is found by introducing the Lagrange multipliers  $\lambda_j$  and maximizing

$$(55) \quad - \sum_{j=1}^N v_j \sum_{i=1}^N p_{ij} \log p_{ij} - \sum_{j=1}^N \lambda_j \sum_{i=1}^N p_{ij}$$

with respect to  $p_{ij}$ . The result is:

$$(56) \quad p_{ij} = e^{-\frac{\lambda_j}{v_j} - 1}.$$

Evaluating the  $\lambda_j$  using equation (54), the resulting maximum entropy mixing process is:

$$(57) \quad p_{ij} = \frac{1}{N} \quad i = 1, \dots, N; \quad j = 1, \dots, N.$$

This process distributes the water from segment  $j$  uniformly over all the segments in one tidal cycle. It is the process which was suggested previously as the most random process possible. However this process violates the law of conservation of estuary geometry since the equilibrium distribution of this chain is a uniform volume distribution and the estuary may not be of such a form.

As a second illustration of a maximum entropy process, consider the maximum entropy process which satisfies both conservation of water and the conservation of the estuary geometry, equations (12) and (23):

$$(58) \quad \sum_{i=1}^N p_{ij} = 1 \quad j = 1, \dots, N$$

$$(59) \quad \sum_{j=1}^N p_{ij} v_j = v_i \quad i = 1, \dots, N$$

Introducing the Lagrange multipliers  $\lambda_j$  and  $\mu_i$  and maximizing:

$$(60) \quad - \sum_{j=1}^N v_j \sum_{i=1}^N p_{ij} \log p_{ij} - \sum_{j=1}^N \lambda_j \sum_{i=1}^N p_{ij} - \sum_{i=1}^N \mu_i \sum_{j=1}^N p_{ij} v_j$$

with respect to  $p_{ij}$ , yields the maximum entropy estimate:

$$(61) \quad p_{ij} = e^{-\lambda'_j - \mu_i}$$

where  $\lambda'_j = \lambda_j / v_j + 1$ . Evaluating the  $\lambda'_j$  and  $\mu_i$  using the constraining equations (58) and (59) yields:

$$(62) \quad p_{ij} = \frac{v_i}{\sum_{k=1}^N v_k}$$

It is interesting to note that the  $p_{ij}$ 's which result are not functions of  $j$ . Hence no matter what segment  $j$  is considered after one tidal cycle the water in segment  $j$  is distributed in proportion to the volume of the estuary. This mixing process tends to favor those segments with larger volumes, but only to the extent required to maintain the geometry of the water in the estuary.

Now consider the composite process  $\underline{\underline{PT}}(q)$ . For any initial distribution of salt mass  $sv$ , the resulting mass of salt in segment  $i$  after one tidal cycle is found by multiplying  $sv$  by  $\underline{\underline{PT}}(q)$ . Using  $p_{ij}$  as given by equation (62), and denoting the  $k^{\text{th}}$  component of  $sv$  by  $(sv)_k$ , the mass of salt in segment  $i$  is found to be:

$$(63) \quad \sum_{j=1}^N \frac{v_i}{\sum_{l=1}^N v_l} \sum_{k=1}^N t_{jk} (sv)_k = v_i \frac{\sum_{k=1}^N (sv)_k}{\sum_{l=1}^N v_l}$$

since by equation (29):

$$(64) \quad \sum_{j=1}^N t_{jk} = 1.$$

The concentration of salt in segment  $i$  is found by dividing the mass of salt in segment  $i$ , equation (63), by the volume of segment  $i$ . Hence, the concentration of salt in segment  $i$  after one tidal cycle is:

$$(65) \quad \frac{\sum_{k=1}^N (sv)_k}{\sum_{l=1}^N v_l} = \frac{\text{total salt mass in the estuary}}{\text{total volume of water in the estuary}}.$$

That is, the concentration of salt in any segment  $i$ ,  $i = 1, \dots, N$ , of the estuary is uniform. So the maximum entropy process predicts a uniform concentration of salt in the estuary regardless of fresh water flow or initial distribution.

This is a perfectly plausible result since any other distribution tends to favor some segment over the other seg-

ments and there is no information in the two conservation laws which would lead one to believe that a more specific, i.e., less uncertain, mixing process is actually the case.

The situation of practical interest is, of course, the maximum entropy estimate constrained by the three conservation laws in which the information contained in the actually observed salinity distributions is incorporated into the maximum entropy estimate. This situation will be considered in the next section.

#### IV-C The Maximum Entropy Estimation Equations and Their Solution.

The three conservation laws, equations (34), (35) and (36), represent the information which will be incorporated into the maximum entropy estimate. For the sake of definiteness the equilibrium version of the law of conservation of salt, equation (36a), will be used. It will be more convenient to consider these equations in component form:

$$(66) \quad \sum_{i=1}^N p_{ij} = 1 \quad j = 1, \dots, N$$

$$(67) \quad \sum_{j=1}^N p_{ij} v_j = v_i \quad i = 1, \dots, N$$

$$(68) \quad \sum_{j=1}^N p_{ij} \sum_{k=1}^N t_{jk} (sv)_k = (sv)_i, \quad i = 1, \dots, N$$

Define the vector  $\underline{r}$  with components  $r_j$  as:

$$(69) \quad r_j = \sum_{k=1}^N t_{jk}(sv)_k .$$

Then  $\underline{r}$  is the result of translating the salinity vector  $\underline{sv}$  by  $\underline{T}(\underline{q})$ . Hence equation (68) becomes:

$$(70) \quad \sum_{j=1}^N p_{ij} r_j = (sv)_i \quad i = 1, \dots, N .$$

The maximum entropy estimate of  $\underline{P}$  subject to the  $3N$  linear constraints, equations (66), (67) and (70) is found by maximizing the following expression with respect to  $p_{ij}$  :

$$(71) \quad - \sum_{j=1}^N v_j \sum_{i=1}^N p_{ij} \log p_{ij} - \sum_{j=1}^N \lambda_j \sum_{i=1}^N p_{ij} - \sum_{i=1}^N \mu_i \sum_{j=1}^N p_{ij} v_j - \sum_{i=1}^N v_i \sum_{j=1}^N p_{ij} r_j .$$

The  $\lambda_j$ 's,  $\mu_i$ 's and  $v_i$ 's are the Lagrange Multipliers for the constraining equations (66), (67) and (70) respectively. Hence the maximum entropy estimate of  $p_{ij}$  is:

$$(72) \quad p_{ij} = \lambda_j' \mu_i' e^{-v_i r_j'}$$

where

$$(73) \quad \lambda_j' = e^{-\frac{\lambda_j}{v_j} - 1}$$

$$(74) \quad \mu_i' = e^{-\mu_i}$$

$$(75) \quad r_j' = r_j / v_j .$$

The Lagrange multipliers are determined from the constraint equations (66), (67) and (70).

$$(76) \quad \sum_{i=1}^N \lambda_j \mu_i e^{-v_i r_j'} = 1 \quad j = 1, \dots, N$$

$$(77) \quad \sum_{j=1}^N \lambda_j \mu_i v_j e^{-v_i r_j'} = v_i \quad i = 1, \dots, N$$

$$(78) \quad \sum_{j=1}^N \lambda_j \mu_i r_j e^{-v_i r_j'} = (sv)_i \quad i = 1, \dots, N$$

where the prime has been dropped from  $\lambda_j'$  and  $\mu_j'$  for convenience. These equations are a set of simultaneous transcendental equations in the  $3N$  unknowns  $\lambda_j$ ,  $\mu_i$  and  $v_i$  for which no general solution is known. Hence a numerical technique must be devised in order to proceed to a solution.

Fortunately these equations are not as strongly coupled as they appear to be. Equation (77) can be solved for  $\mu_i$  directly.

$$(79) \quad \mu_i = \frac{v_i}{\sum_{k=1}^N \lambda_k v_k e^{-v_i r_k'}}$$

Substituting this expression into equation (78) yields:

$$(80) \quad \sum_{j=1}^N \lambda_j \frac{v_i}{\sum_{k=1}^N \lambda_k v_k e^{-v_i r_k'}} r_j e^{-v_i r_j'} = (sv)_i$$

However the denominator of each factor is not a function of the index  $j$  so it may be factored out, giving:

$$(81) \quad \sum_{j=1}^N \lambda_j v_i r_j e^{-v_i r_j} = (sv)_i \sum_{k=1}^N \lambda_k v_k e^{-v_i r_k}$$

or collecting like terms:

$$(82) \quad \sum_{j=1}^N \lambda_j [v_i r_j - (sv)_i v_j] e^{-v_i r_j} = 0, \quad i = 1, \dots, N.$$

The second set of  $N$  equations is obtained by substituting equation (79) into equation (76).

$$(83) \quad \sum_{i=1}^N \lambda_j \frac{v_i}{\sum_{k=1}^N \lambda_k v_k e^{-v_i r_k}} e^{-v_i r_j} = 1, \quad j = 1, \dots, N.$$

Unfortunately the denominator of each term in the summation is a function of the summation index  $i$  so no further simplification is possible. Therefore the problem has been reduced to solving  $2N$  simultaneous transcendental equations, equations (82) and (83), which is still a very difficult numerical problem.

However, the set of  $N$  equations, equation (82), has an important property: although each equation involves all the  $\lambda_j$ 's since the summation is on  $j$ , each equation involves only one  $v_i$ . If the  $\lambda_j$ 's were known, then each  $v_i$  could be computed by a simple numerical technique. This observation leads to the following algorithm for solving the  $3N$  simultaneous equations. For an assumed set of values for

the  $\lambda_j$ 's, first solve for each  $v_i$ , one at a time, using equation (82). This most easily is accomplished using the Newton root-finding algorithm.<sup>[46]</sup> Second, compute the  $\mu_i$ 's using equation (79); and third, compute the  $\lambda_j$ 's implied by these value of  $v_i$  and  $\mu_i$  using equation (76), i.e.,

$$(84) \quad \lambda_j = \frac{1}{\sum_{i=1}^N \mu_i e^{-v_i r_{ij}}}$$

With these values of  $\lambda_j$  start the process again.

This iteration technique has been tried with several sets of data and the convergence is surprisingly rapid. For example, the average computation time needed on an IBM 7094 Computer to achieve six place accuracy for an eleven segment problem ( $N = 11$ ) was less than one minute.

It is interesting to note that the numerical solution of these  $3N$  simultaneous transcendental equations on a computer is not much more difficult than solving the  $3N$  simultaneous linear equations which result from Preddy's approach. Also the problem of satisfying the constraint  $p_{ij} \geq 0$  does not arise in a maximum entropy estimate,<sup>[47]</sup> whereas it can be a problem in Preddy's method.<sup>[29]</sup> Therefore the maximum entropy estimate of the mixing matrix is a more satisfactory solution to the problem of estimating  $\underline{P}$  from both theoretical and computational points of view.

In the next chapter a set of salinity intrusion data from the Delaware River Model is analyzed. The most interesting aspect of this analysis is the resulting maximum entropy mixing process.

## CHAPTER V

### NUMERICAL RESULTS OF A MAXIMUM ENTROPY ESTIMATE

In this chapter the maximum entropy estimate of the mixing process in the Delaware River Model is calculated. The resulting mixing process is used to calculate equilibrium salinity distributions. These theoretical predictions are compared with the observed data.

#### V-A The Data Requirements

The maximum entropy estimate of the mixing process in an estuary can be calculated using various types of salinity or dye dispersion data. So long as enough data is available to establish reliable coefficients for equations (34), (35) and (36), the maximum entropy estimate can be successfully calculated. In this respect the data requirements of the theory of maximum entropy mixing are much more lenient than the requirements of the theories of salinity intrusion based on the convective diffusion equation. The theories based on the convective diffusion equation require an equilibrium distribution in order to calculate the longitudinal dispersion coefficient, [48] whereas the maximum entropy estimate can accommodate distributions collected during a period of varying fresh water flow. This is an important practical advantage since field data is usually collected during a period when

the hydrograph of the estuary is varying and the salinity distributions which are observed are not the equilibrium distributions. [49]

The maximum entropy mixing theory, as well as the theories which are based on the convective diffusion equations, can be used to analyze a limited region of the estuary if the available salinity data is limited to this region. The mixing matrix for this region can be determined; however, it is only possible to predict the salinity distribution in this region up to a normalization constant. If, however, a spatial distribution of salinity is available which extends from the upstream limit of the region of interest to a point downstream where the ocean value of salinity is maintained, then the arbitrary constant is accounted for by using the ocean value of salinity as a boundary condition. This is the preferable procedure since the value of the salinity concentration in each segment can be calculated.

#### V-B The Delaware River Model Data

For the first trial of the maximum entropy mixing theory, a set of salinity intrusion data which was obtained from a hydraulic model of an estuary is used. Since this data is obtained under controlled conditions, a more critical evaluation of the maximum entropy hypothesis is possible. The results are not obscured by the difficulties which are often encountered in the use of field data. For example, if the specific conductance is used as a measure of the salinity of the estuary water, a relationship between specific conductance and salinity

must be used to obtain the salinity distributions. However, industrial pollutants can contribute dissolved solids which affect the specific conductance and their presence must be taken into account by specifying their origin and rate of entry into the estuary. This complicates the analysis since it has been assumed in the preceding formulations that the only significant source of salinity or, in this case, dissolved solids, is the ocean water at the mouth of the estuary. In the case of the data from a hydraulic model of an estuary, this condition can be satisfied exactly.

The data chosen for this analysis was collected during salinity tests of the Delaware River Model at the Waterways Experiment Station.<sup>[50]</sup> Six tests were conducted at sustained fresh water discharges of 5000, 7000, 9000, 10,600, 13,000 and 16,475 cubic feet per second (cfs). The equilibrium salinity distributions at the high water slack time were recorded at surface, mid-depth and bottom. The data reported represented the averaged results of not less than two identical tests.

For each value of the fresh water discharge, the spatial distribution was obtained along the center-line of the channel. The seaward limit of the sampling was channel station 350, the Ship John Shoal Light. The channel stations are 1000 feet apart starting from channel station 0 at Allegheny Avenue, Philadelphia and increasing downstream. Unfortunately in the downstream portion of the estuary the salinity was recorded only at channel stations 350, 275 and 250. Hence if

this region is included in the analysis, the missing data must be accounted for by interpolation since the salinity variation is considerable in this region. This is a worthwhile procedure to follow if channel station 350 can be used as the ocean end of the estuary. However, the value of salinity at channel station 350 does not remain a constant throughout the six tests but decreases from a high of about 12,000 parts per million of chloride (ppm) at  $q = 5000$  cfs to a low of about 11,000 ppm at  $q = 16,475$  cfs. Hence the data does not extend far enough seaward to permit the use of a fixed value of salinity as a boundary condition at the furthest downstream segment. In view of the fact that most of the data between channel stations 250 and 350 would have to be interpolated using only the data available at three stations, the downstream limit of the analysis is chosen to be channel station 250.

At the upstream end of the estuary, the data for the six tests is complete only to channel station 150 although bottom sampling was continued further upstream. For a one-dimensional analysis, the average salinity over the vertical direction is usually taken as the value of the salinity at a particular station. Hence the upstream limit of the analysis is chosen to be channel station 150. The salinity data presented in Figures 4 through 9 for surface, mid-depth, and bottom samples indicate that the Delaware River Model is only slightly stratified.

The only other data required is the cross sectional area of the estuary. This data was obtained from the U.S.

Army Corps of Engineers survey of the Delaware River Estuary, [51]

The estuary is divided into eleven segments each of which is 10,000 feet in length. The centers of the segments are at the channel stations 150 to 250 inclusive. The volume of each segment is computed from a graphical integration of the cross sectional area data. An eighth order polynomial is fit by a least mean square technique and the integral of this polynomial is used to determine the volume of the segments. This technique serves to smooth the data as well as providing the integral in a straightforward way. A plot of the cross sectional area and the fitted polynomial is given in Figure 1. The calculated volumes are listed in Table 1. The volumes of these segments are large enough so that the condition on the elements of the translation matrix,  $\underline{T(q)}$ , given by equation (25) is satisfied for  $q = 16,475$  cfs. The values of  $t_{j,j}$  and  $t_{j+1,j}$  for  $q = 5000$  cfs and  $q = 16,475$  cfs are listed in Table 2. The tidal period is 12 hours and 25 minutes, [52] and this is the time interval between successive high water slack samples.

The equilibrium concentration of salinity in each segment is necessary if the distribution is to be used in the maximum entropy estimate. The missing values are obtained by a least mean square polynomial interpolation of the existing data. The interpolated data is listed in Table 3.

This completes the preliminary data analysis of the cross sectional area data and the salinity distributions.

FIG. 1 CROSS SECTIONAL AREA — DELAWARE RIVER

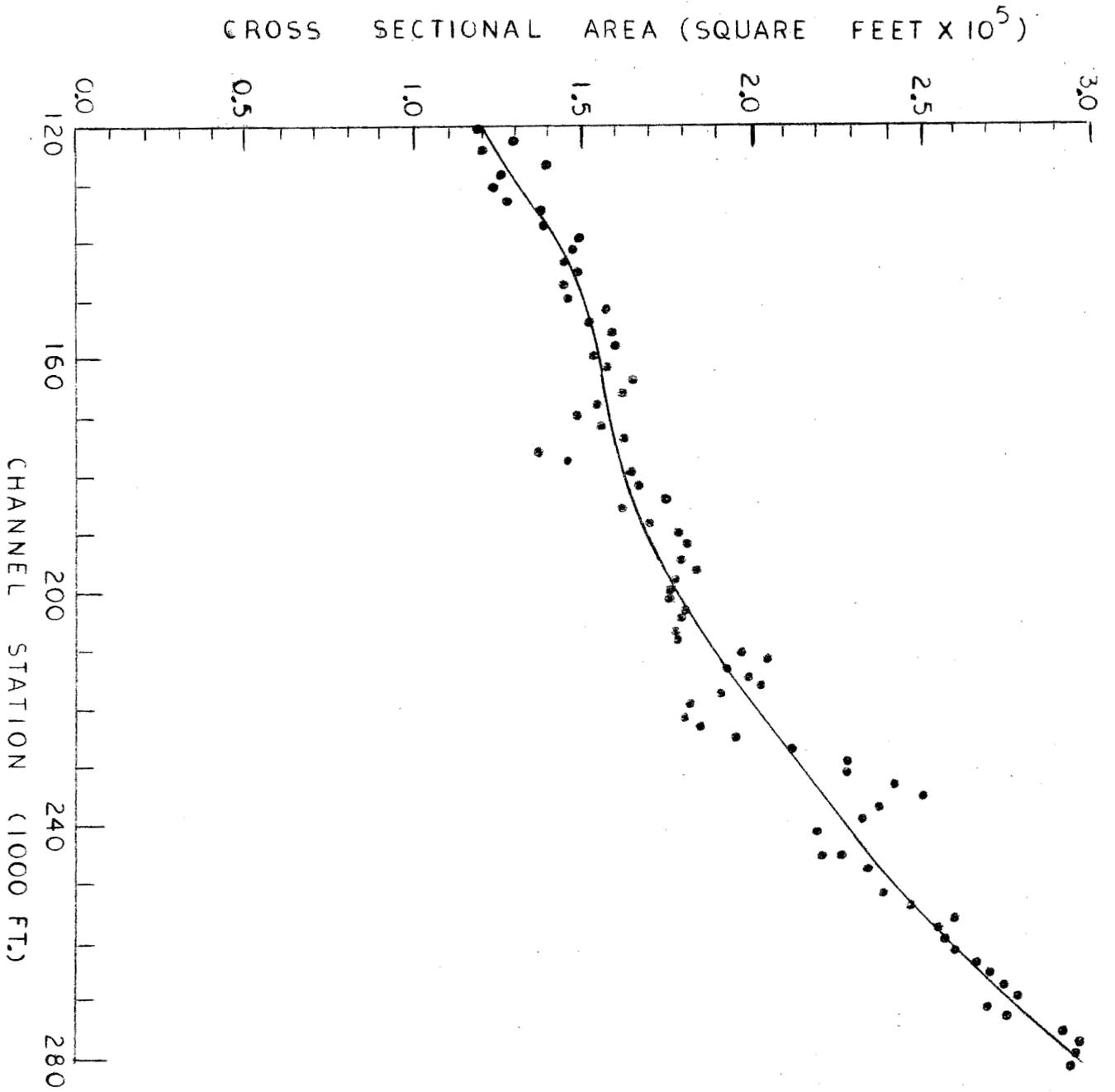


TABLE 1  
VOLUMES OF THE SEGMENTS

Segment Number	Center Channel Station (1000 ft.)	Volume (cubic ft. $\times 10^9$ )
1	150	1.501
2	160	1.548
3	170	1.575
4	180	1.612
5	190	1.680
6	200	1.765
7	210	1.877
8	220	2.006
9	230	2.130
10	240	2.260
11	250	2.414

TABLE 2  
THE ELEMENTS OF THE TRANSLATION MATRIX  $\underline{T}(q)$

Segment Number j	q = 5000 cfs		q = 16,475 cfs	
	$t_{jj}$	$t_{j+1,j}$	$t_{jj}$	$t_{j+1,j}$
1	.8511	.1489	.5093	.4907
2	.8554	.1446	.5234	.4766
3	.8581	.1419	.5321	.4678
4	.8613	.1387	.5429	.4571
5	.8669	.1331	.5615	.4385
6	.8734	.1266	.5826	.4174
7	.8809	.1191	.6076	.3924
8	.8885	.1115	.6328	.3672
9	.8951	.1049	.6542	.3458
10	.9011	.0989	.6741	.3259
11	1.0000	—	1.0000	—

TABLE 3  
INTERPOLATED SALINITY DATA  
DELAWARE RIVER MODEL - HIGH WATER SLACK

Channel Station (1000 ft.)	Salinity (ppm - chlorine)					
	q = 5000 cfs	q = 7000 cfs	q = 9000 cfs	q = 10,600 cfs	q = 13,000 cfs	q = 16,475 cfs
250	5730	5700	5180	5050	4500	4760
240	5200	5120	4770	4400	3800	3950
230	4760	4600	4300	3850	3300	3200
220	4400	4100	3820	3250	2810	2600
210	4000	3630	3300	2750	2330	2100
200	3580	3150	2800	2250	1830	1630
190	3100	2650	2300	1800	1360	1170
180	2650	2200	1830	1360	970	750
170	2250	1800	1400	1000	650	410
160	1910	1460	1080	700	380	180
150	1650	1220	830	480	230	80

In the next section the maximum entropy estimate of the mixing process which is based on this data is analyzed and the theoretical predictions of the equilibrium salinity distributions are compared to the data.

#### V-C The Maximum Entropy Estimate of the Mixing Process

The Delaware River Model data consists of six equilibrium salinity distributions which occur at six different fresh water flows. In order to test the maximum entropy mixing theory, the estimate is based on only one of these six distributions. Then the remaining five salinity distributions are compared to the theoretical predictions based on the estimated mixing process. In this way the data can be used as a test of the ability of the maximum entropy mixing theory to predict equilibrium salinity distributions at different fresh water flows.

If all the data is used in the maximum entropy estimate, then there is no independent data which can be used as a check on the theory.

Two separate calculations of this type are presented. The first estimate of the mixing process is based on the equilibrium distribution at  $q = 5000$  cfs. The maximum entropy estimate of the mixing matrix is presented in Table 4. The one cycle distribution functions, which are the columns of the mixing matrix, are plotted in Figure 2. The one cycle distribution from the segment  $j$  is the maximum entropy estimate of the proportion of the water from segment  $j$

TABLE 4  
THE MIXING MATRIX  $P = (p_{ij})$  ESTIMATED  
FROM THE DATA AT  $q = 5000$  cfs

i	P <sub>i1</sub>	P <sub>i2</sub>	P <sub>i3</sub>	P <sub>i4</sub>	P <sub>i5</sub>	P <sub>i6</sub>
1	0.5913	0.2707	0.1036	0.0185	0.0013	0.0001
2	0.2939	0.3437	0.2478	0.0955	0.0162	0.0012
3	0.0939	0.2494	0.3131	0.2366	0.0864	0.0145
4	0.0182	0.1040	0.2189	0.3099	0.2312	0.0836
5	0.0025	0.0273	0.0904	0.2207	0.3068	0.2167
6	0.0002	0.0043	0.0219	0.0889	0.2197	0.2878
7	0.0000	0.0006	0.0039	0.0248	0.1005	0.2256
8	0.0000	0.0000	0.0004	0.0044	0.0291	0.1121
9	0.0000	0.0000	0.0000	0.0007	0.0074	0.0440
10	0.0000	0.0000	0.0000	0.0000	0.0013	0.0125
11	0.0000	0.0000	0.0000	0.0000	0.0001	0.0018

i	P <sub>i7</sub>	P <sub>i8</sub>	P <sub>i9</sub>	P <sub>i,10</sub>	P <sub>i,11</sub>
1	0.0000	0.0000	0.0000	0.0000	0.0000
2	0.0000	0.0000	0.0000	0.0000	0.0000
3	0.0017	0.0001	0.0000	0.0000	0.0000
4	0.0192	0.0031	0.0005	0.0000	0.0000
5	0.0910	0.0261	0.0061	0.0008	0.0000
6	0.2110	0.1019	0.0385	0.0089	0.0001
7	0.2681	0.2040	0.1163	0.0439	0.0016
8	0.2161	0.2593	0.2231	0.1370	0.0168
9	0.1275	0.2240	0.2723	0.2516	0.0835
10	0.0536	0.1361	0.2306	0.3156	0.2722
11	0.0117	0.0454	0.1126	0.2421	0.6258

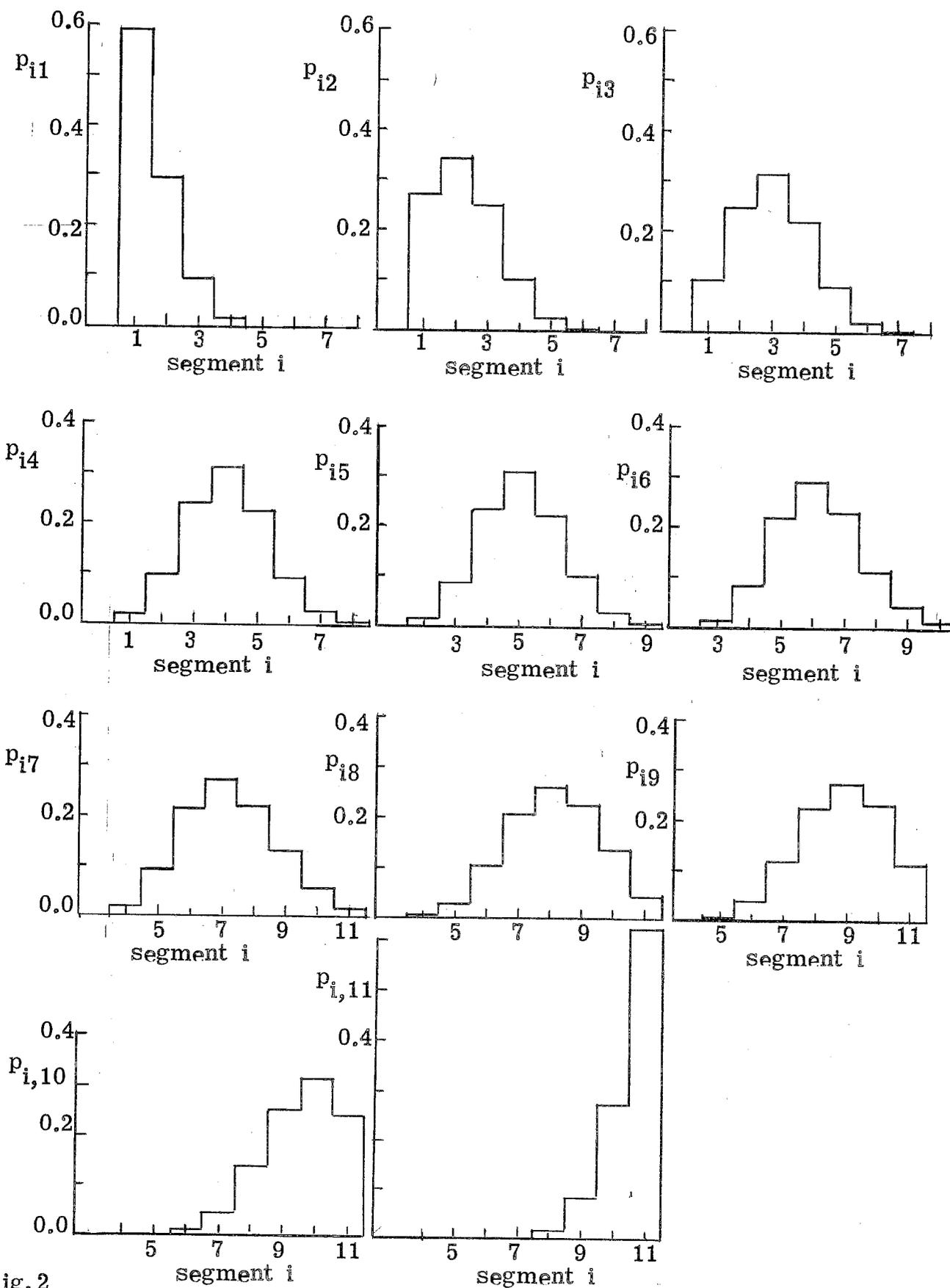


Fig. 2  
One cycle distributions of the maximum entropy estimate (est. at  $q = 5000$  cfs)

which is transferred to all the other segments during one tidal cycle. For example,  $p_{11}$ , the proportion of water which remains in segment 1 after one tidal cycle is 0.5913; and  $p_{31}$ , the proportion of water which is transferred to segment 3 from segment 1 is 0.0939.

The estimates of these distributions have two important properties in common. First, more water remains in the segment from which it originated than is transferred to any other segment. And second, the transfer of water to any of the segments more than four segments away, i.e., 40,000 feet, from the origin of the water is negligibly small. The first property agrees with the intuitive notion that the tidal action is basically oscillatory. Disregarding the effect of the fresh water flow, the water which flows out on the ebb tide flows back on the flood tide to approximately where it began. However, it is known that dispersion also takes place and not all the water returns to exactly the point from which it originated. The question that naturally arises is how far from the point that it originated can the water be reasonably expected to disperse.

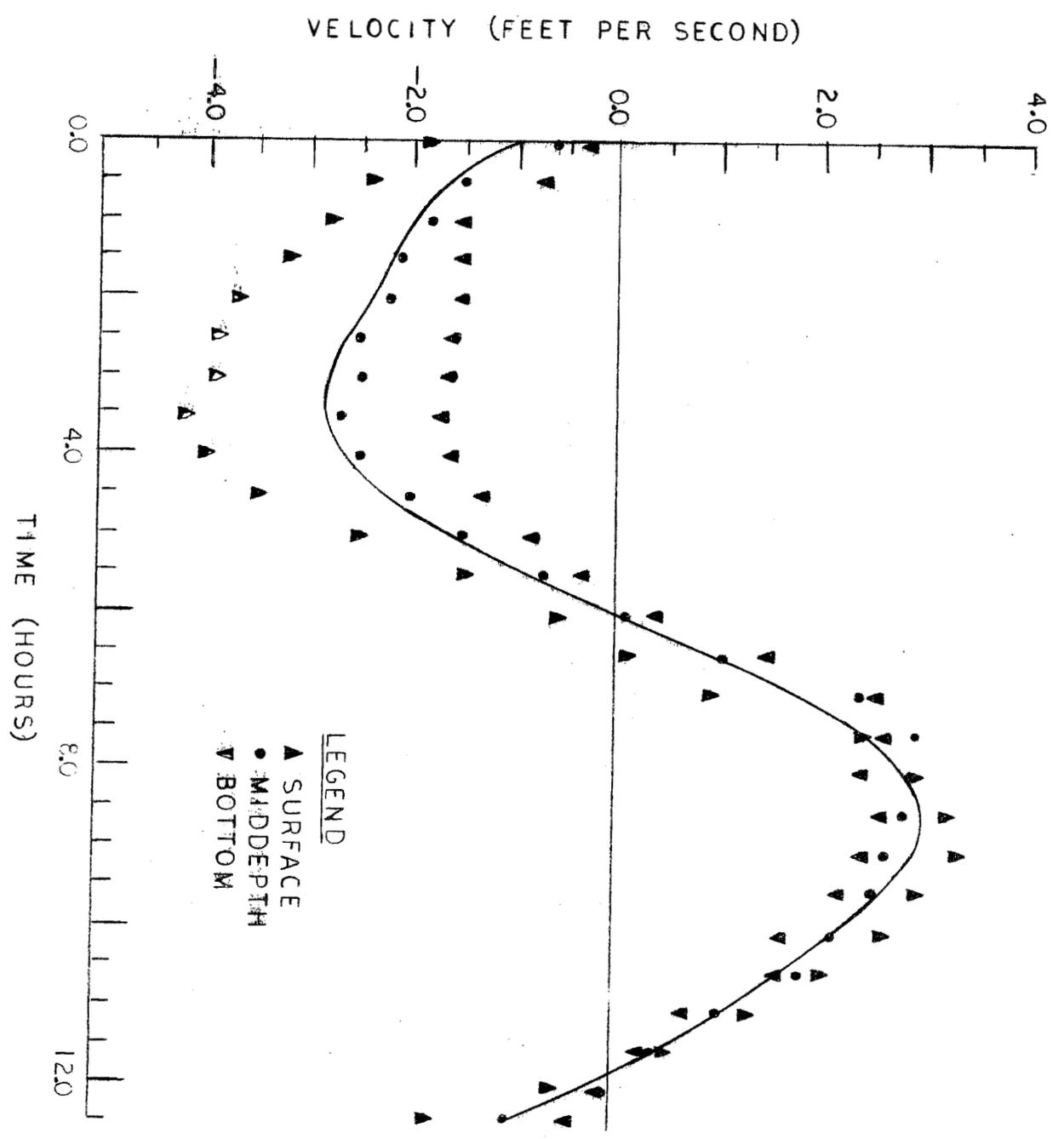
The basic mechanism of dispersion in open channel flow involves the nonuniform velocity distribution as well as the turbulent eddy diffusion.<sup>[53,54]</sup> Hence, it is conceivable that some water travels seaward during the ebb tide and remains there in a region of low velocity during the flood tide. And similarly, some water remains in a region of low velocity during the ebb tide and then may be pushed

back from its origin during the flood tide. The maximum downstream distance that water can travel in this way can be computed approximately by integrating the tidal velocity over the time of the ebb tide. The upstream distance is similarly computed by integrating the tidal velocity over the time of the flood tide. This calculation assumes that the velocity profiles are the same over the distances involved.

The tidal velocity data for channel station 315 is plotted in Figure 3.<sup>[55]</sup> An eighth order polynomial is fit to the data and the lengths are calculated by integrating the positive and negative portions of this interpolating polynomial. The results are that the maximum downstream tidal excursion length is approximately 39,000 feet and the maximum upstream tidal excursion length is approximately 45,000 feet. Some prototype data is available which gives these lengths at the Burlington-Bristol Bridge, channel station -81, as 42,200 feet downstream and 32,000 feet upstream as an average of the four days data reported. Also for the Delaware Memorial Bridge, channel station 180, the downstream length is 49,200 feet and the upstream length is 62,600 feet for one day of data.<sup>[56]</sup>

These values indicate that if any dispersion is predicted to take place upstream or downstream over distances of greater than approximately 40,000 to 50,000 feet, then the predictions are not in accord with the maximum distances which are physically possible.

FIG. 3 TIDAL VELOCITY - CHANNEL STATION 315



The maximum entropy estimates of the one cycle distributions of the water are all contained within four segments, or 40,000 feet of the origin of the water. Hence the estimated mixing process is physically realistic from this point of view.

It is interesting to note that the value of maximum length over which dispersion can occur is assumed a priori in Preddy's theory of estuarine mixing. Although as argued above this length is related to the maximum tidal excursion distance, its actual value is unknown and Preddy is forced to assume its value in order to proceed with his calculation. However, the maximum entropy estimate of the mixing process predicts the maximum length over which dispersion can occur. And, as remarked above, the predicted length is within the bounds set by the maximum tidal excursion lengths.

The equilibrium salinity distributions for fresh water flow rates of 5000, 7000, 9000, 10,600, 13,000 and 16,475 cfs which are predicted by the maximum entropy estimate and the corresponding Delaware River Model data are plotted in Figures 4 through 9. The arbitrary normalization constant is determined by using the following procedure. For each equilibrium salinity distribution, the values of the salinity in each segment are determined from the interpolating polynomial and their sum is obtained. Then the theoretical prediction is normalized so that the sum of the values of the predicted salinity agrees with the sum of the interpolated salinity data. Hence the salinity data at

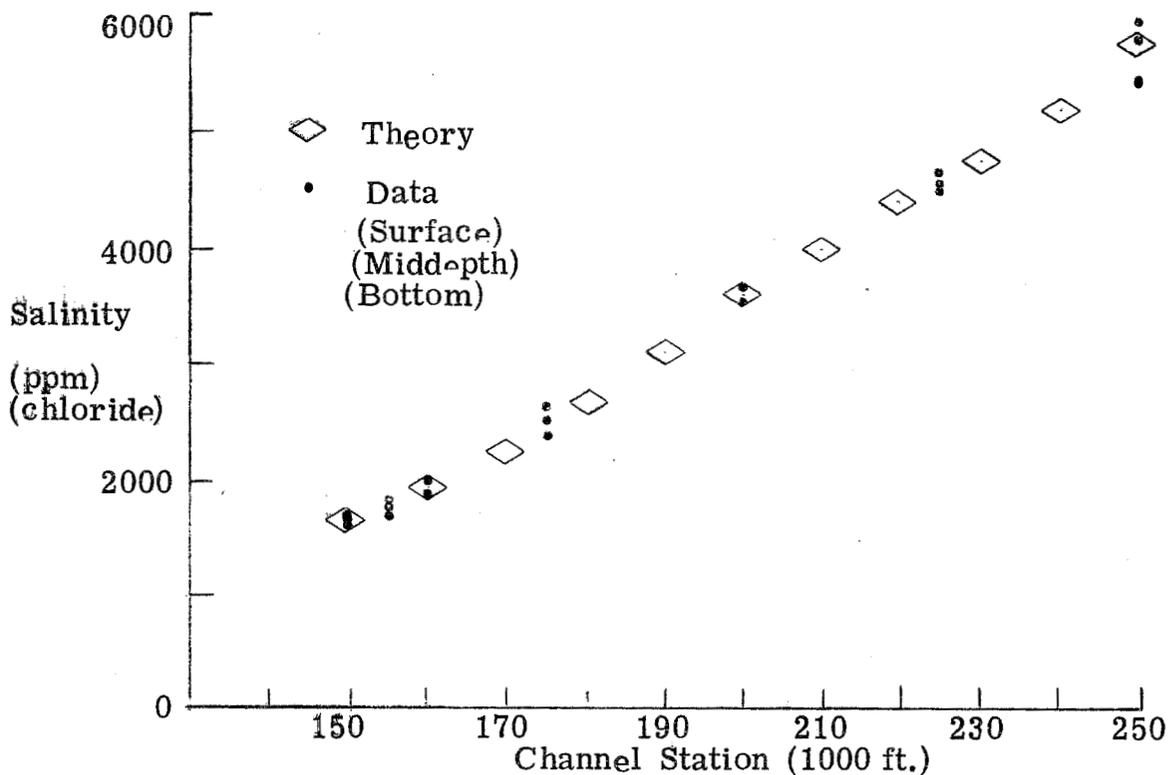


Fig. 4 Comparison of theory (est. at  $q = 5000$  cfs) and data  $q = 5000$  cfs

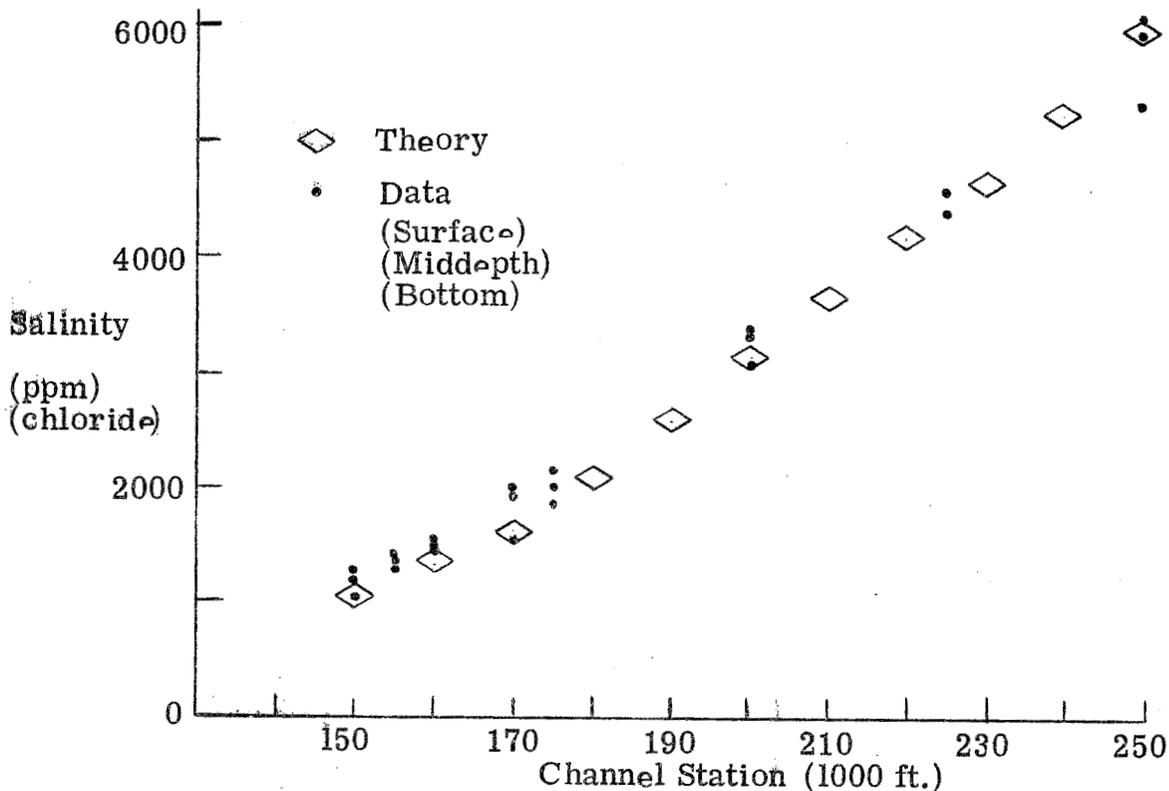


Fig. 5 Comparison of theory (est. at  $q = 5000$  cfs) and data  $q = 7000$  cfs

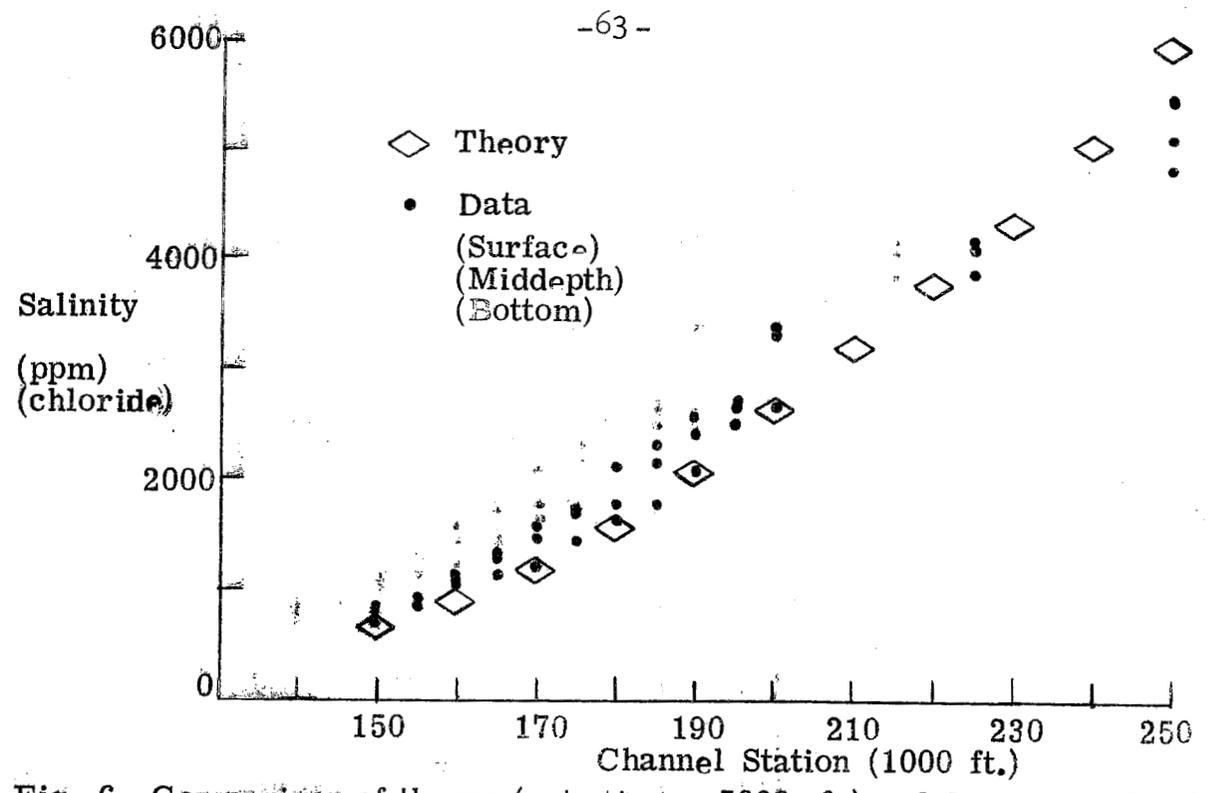


Fig. 6 Comparison of theory (est. at  $q = 5000$  cfs) and data  $q = 9000$  cfs

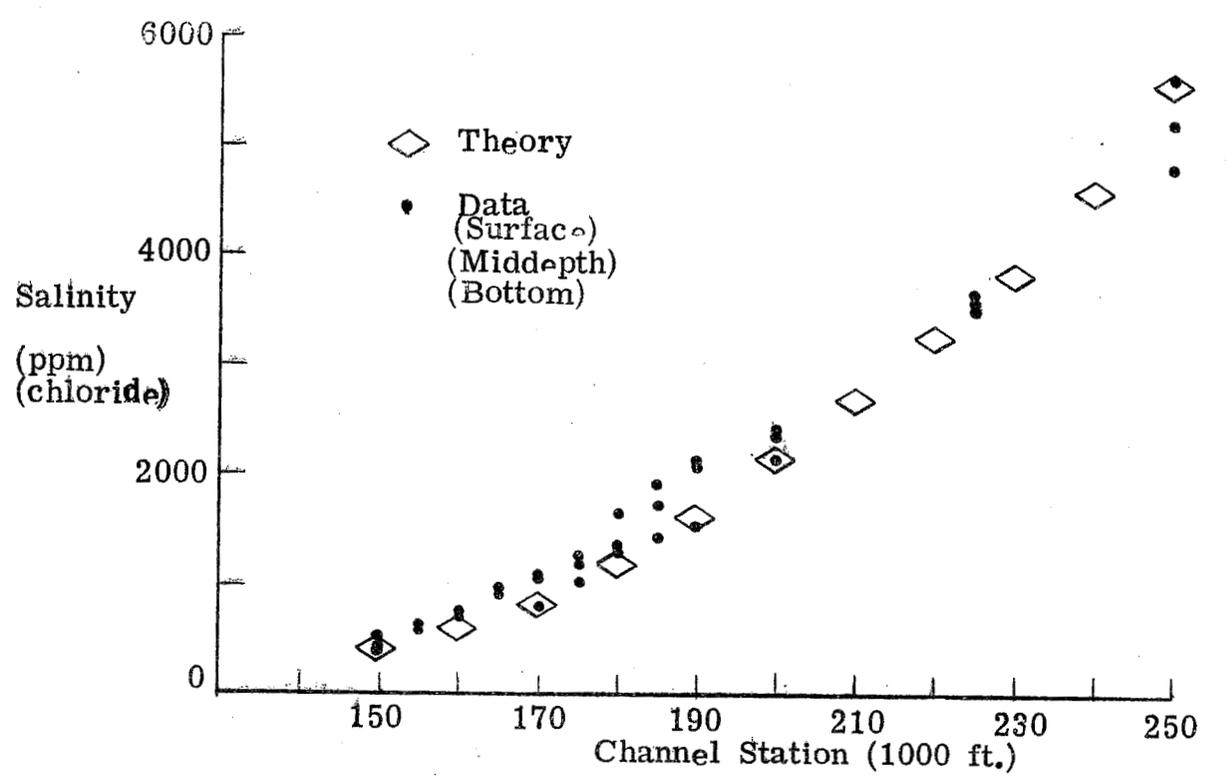


Fig. 7 Comparison of theory (est. at  $q = 5000$  cfs) and data  $q = 10,600$  cfs

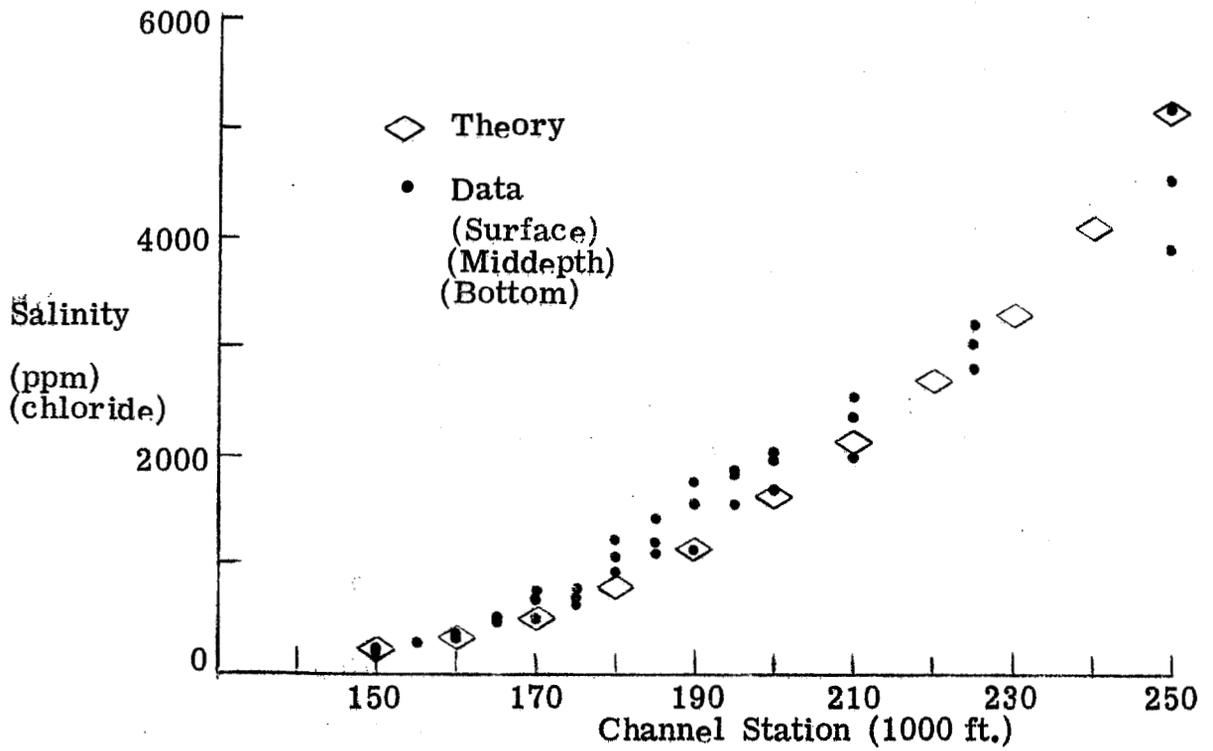


Fig. 8 Comparison of theory (est. at  $q = 5000$  cfs) and data  $q = 13,000$  cfs

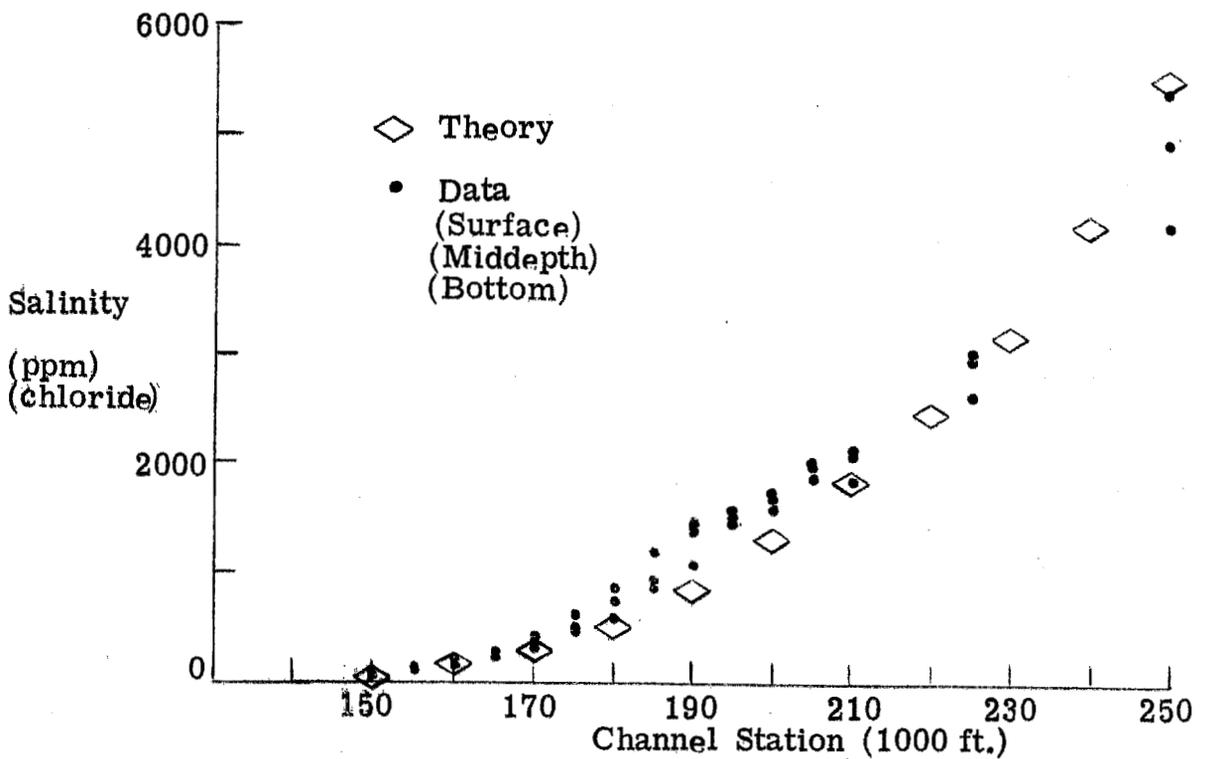


Fig. 9 Comparison of theory (est. at  $q = 5000$  cfs) and data  $q = 16,475$  cfs

each fresh water flow rate is used to obtain the normalization constant for the theoretical prediction. To this extent, therefore, the prediction at each fresh water flow rate is dependent on the actual data. This is the unfortunate consequence of not having salinity distribution which extend far enough downstream.

The resulting theoretical predictions agree rather well with the observed data, although there are some discrepancies at the higher fresh water flows. This can be attributed to the fact that the data at  $q = 5000$  cfs which is used in the estimate is mostly interpolated data. This can be seen in Figure 4. The theoretical values shown are, of course, the values of the interpolating polynomial used in the estimate. The estimate based on the lowest fresh water rate available is computed in the hope that the non-linear effects of the density current mixing would be noticeable.<sup>[57]</sup> That is, that the deviations between the theory and the data would increase at higher fresh water flow rates. This would indicate that the mixing matrix is a function of the salinity concentration. However the deviations between the theory and the data do not indicate any sizeable discrepancy which could be attributed to this phenomena.

The second calculation of the maximum entropy estimate of the mixing process is based on the equilibrium distribution at  $q = 13,000$  cfs. The estimated mixing matrix is presented in Table 5 and the one cycle distribution functions are plotted in Figure 10. The theoretical predictions

TABLE 5  
THE MIXING MATRIX  $P = (p_{ij})$  ESTIMATED  
FROM THE DATA AT  $q = 13,000$  cfs

i	$P_{i1}$	$P_{i2}$	$P_{i3}$	$P_{i4}$	$P_{i5}$	$P_{i6}$
1	0.6096	0.3080	0.0661	0.0035	0.0001	0.0000
2	0.2834	0.3746	0.2653	0.0695	0.0067	0.0002
3	0.0792	0.2939	0.3304	0.2629	0.0984	0.0163
4	0.0213	0.0793	0.2050	0.3052	0.2448	0.1024
5	0.0053	0.0255	0.0912	0.2094	0.2845	0.2257
6	0.0011	0.0067	0.0308	0.0989	0.2021	0.2632
7	0.0001	0.0015	0.0084	0.0352	0.1004	0.1959
8	0.0000	0.0003	0.0020	0.0107	0.0401	0.1090
9	0.0000	0.0000	0.0005	0.0033	0.0153	0.0537
10	0.0000	0.0000	0.0002	0.0012	0.0062	0.0261
11	0.0000	0.0000	0.0000	0.0002	0.0014	0.0075

i	$P_{i7}$	$P_{i8}$	$P_{i9}$	$P_{i,10}$	$P_{i,11}$
1	0.0000	0.0000	0.0000	0.0000	0.0000
2	0.0000	0.0000	0.0000	0.0000	0.0000
3	0.0013	0.0001	0.0000	0.0000	0.0000
4	0.0228	0.0035	0.0004	0.0000	0.0000
5	0.1024	0.0316	0.0075	0.0014	0.0000
6	0.2070	0.1104	0.0452	0.0151	0.0002
7	0.2414	0.2010	0.1283	0.0675	0.0034
8	0.1941	0.2326	0.2137	0.1630	0.0293
9	0.1266	0.2996	0.2436	0.2469	0.1186
10	0.0758	0.1479	0.2210	0.2768	0.2759
11	0.0286	0.0723	0.1403	0.2293	0.5726

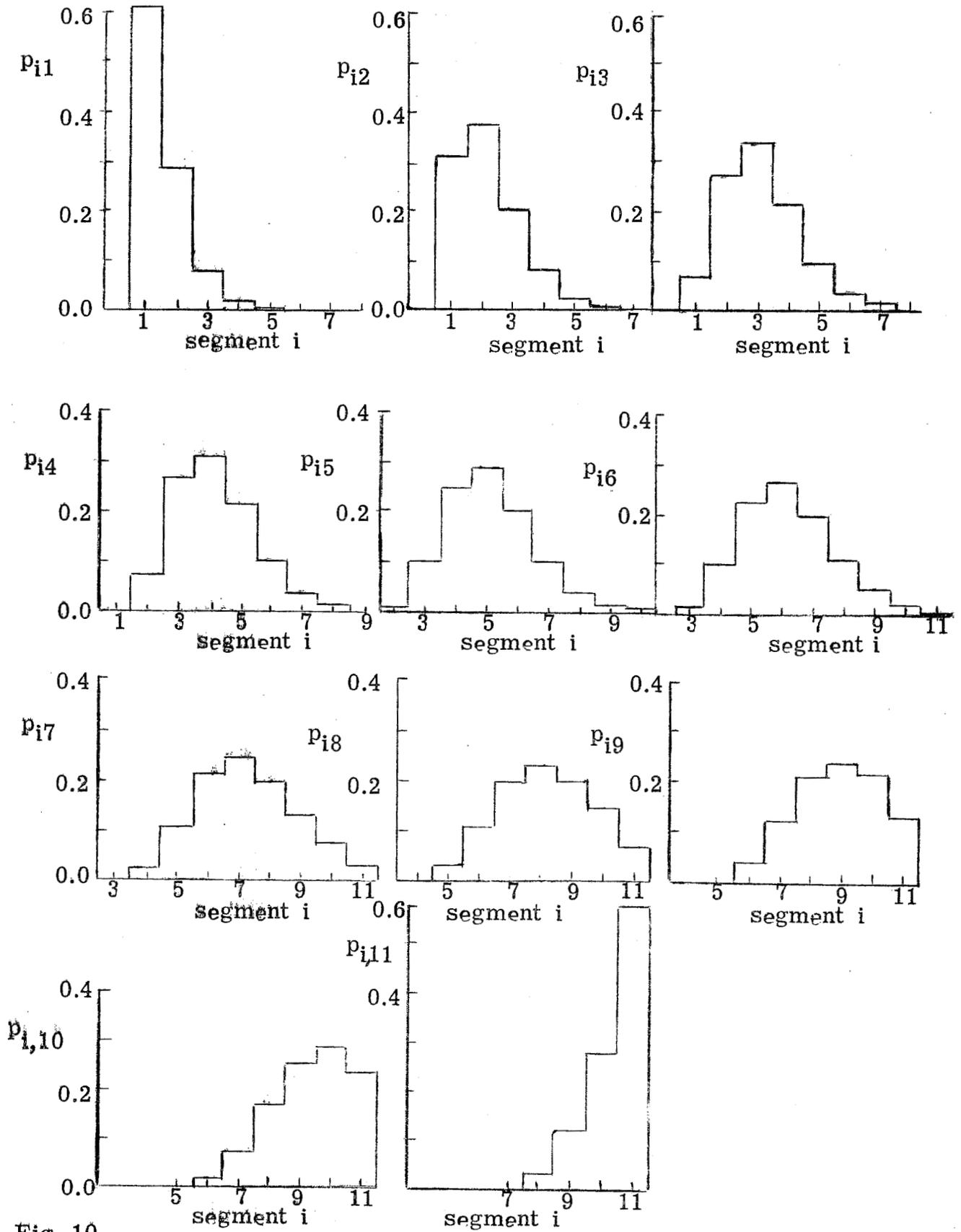


Fig. 10  
One cycle distributions of the maximum entropy estimate (est. at  $q = 13,000$  cfs)

of the equilibrium salinity distributions and the corresponding data are plotted in Figures 11 through 16. The agreement between the predictions and the data is remarkably good. Even at the low fresh water flows,  $q = 5000$  cfs and  $q = 7000$  cfs, the theoretical predictions are close to the data. It is interesting to observe the change in the shape of the theoretical prediction of the distribution at the lower fresh water flow rates. This behavior would not be observed in a theory based on the convective diffusion equation.

#### V-D Conclusions

The results of the preceding calculations clearly indicate that the maximum entropy estimate of the mixing process in an estuary is a sound theoretical and practical solution to the problem of describing the mixing process in an estuary. The predicted salinity profiles agree quite closely with the observed data. Furthermore, the resulting maximum entropy mixing process is a physically reasonable process. The one cycle distribution functions are justifiable in terms of some fundamental physical observations on how the tidal oscillations accomplish the observed dispersion. As a practical matter, the preceding calculations indicate that a linear theory of mixing in slightly stratified estuaries is acceptable if the range of prediction is not extended too far from the conditions under which the estimate of the mixing process is made. For situations within this range, the theoretical predictions of the equilibrium concentrations of salt are in close agreement with the observed data.

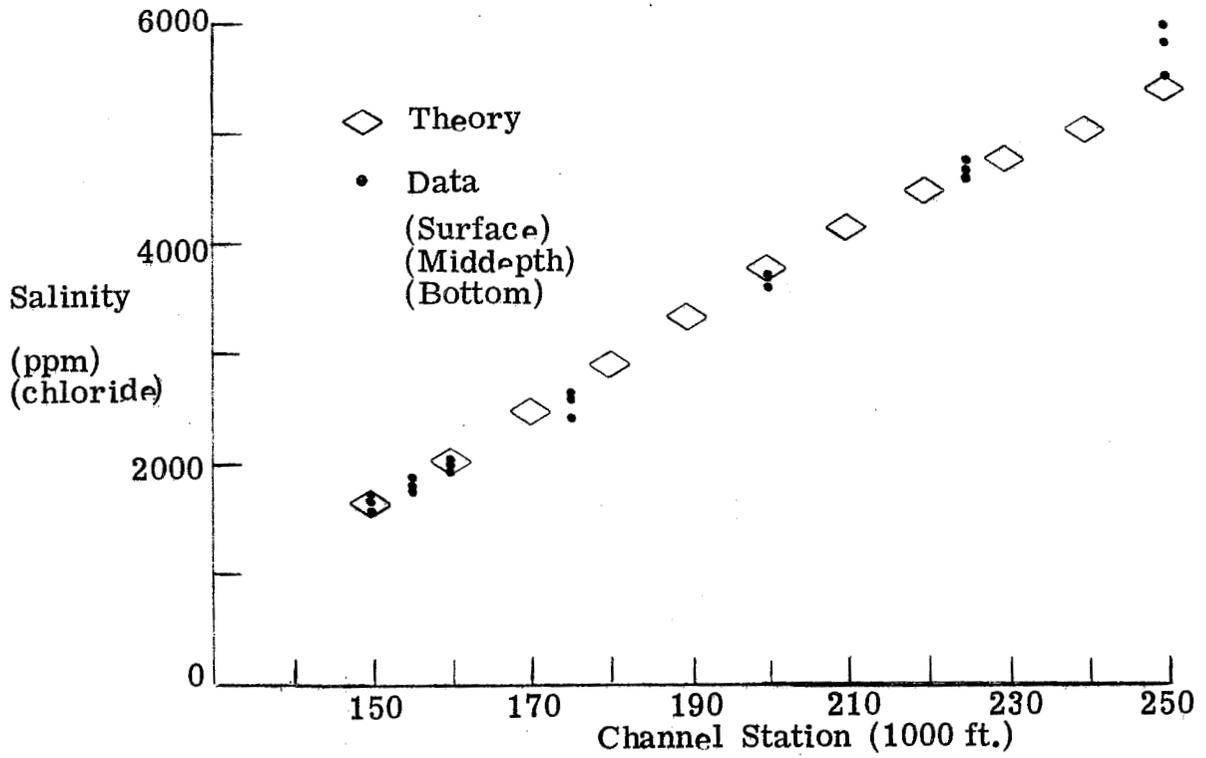


Fig. 11 Comparison of theory (est. at  $q = 13,000$  cfs) and data  $q = 5000$  cfs

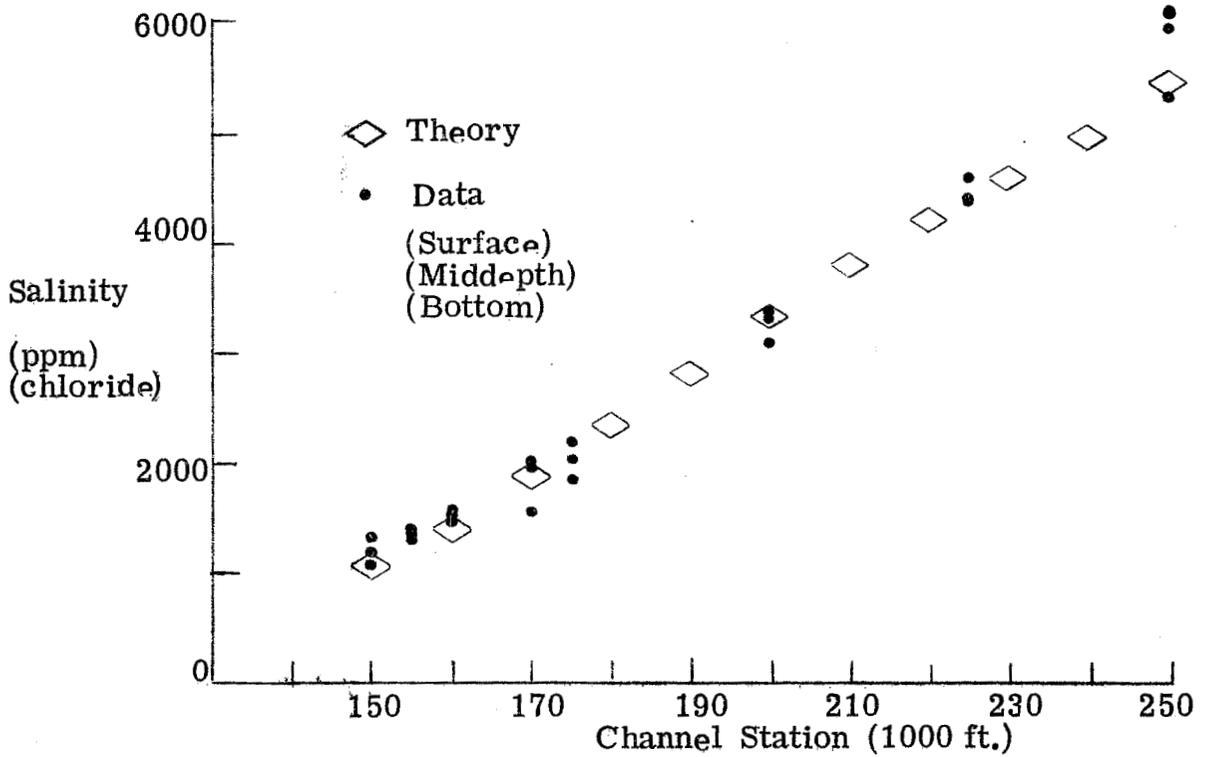


Fig. 12 Comparison of theory (est. at  $q = 13,000$  cfs) and data  $q = 7000$  cfs

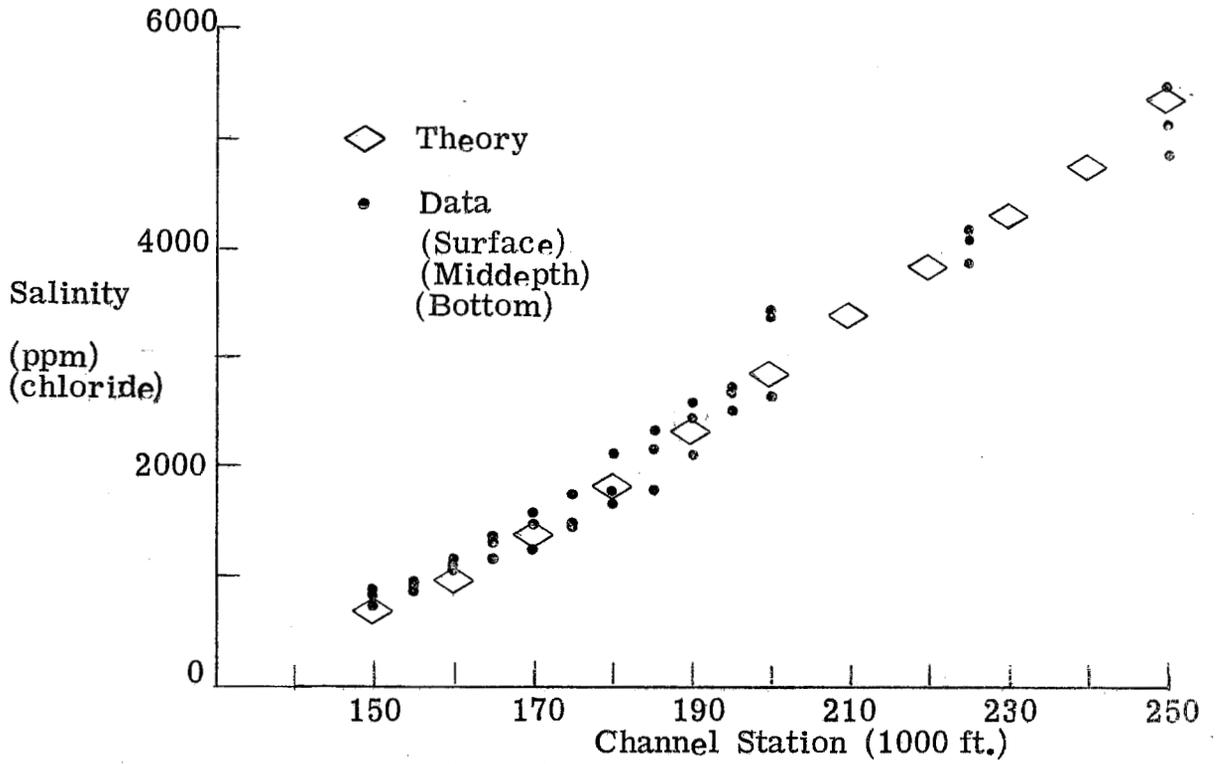


Fig. 13 Comparison of theory (est. at  $q = 13,000$  cfs) and data  $q = 9000$  cfs

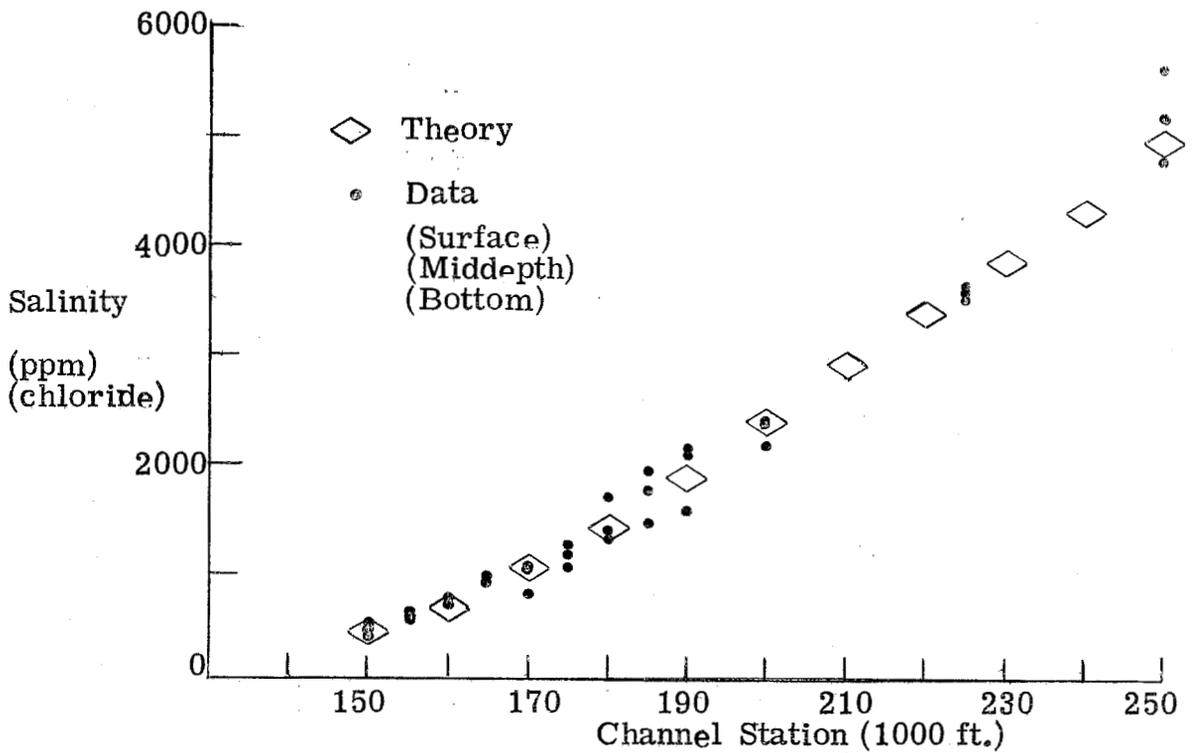


Fig. 14 Comparison of theory (est. at  $q = 13,000$  cfs) and data  $q = 10,600$  cfs

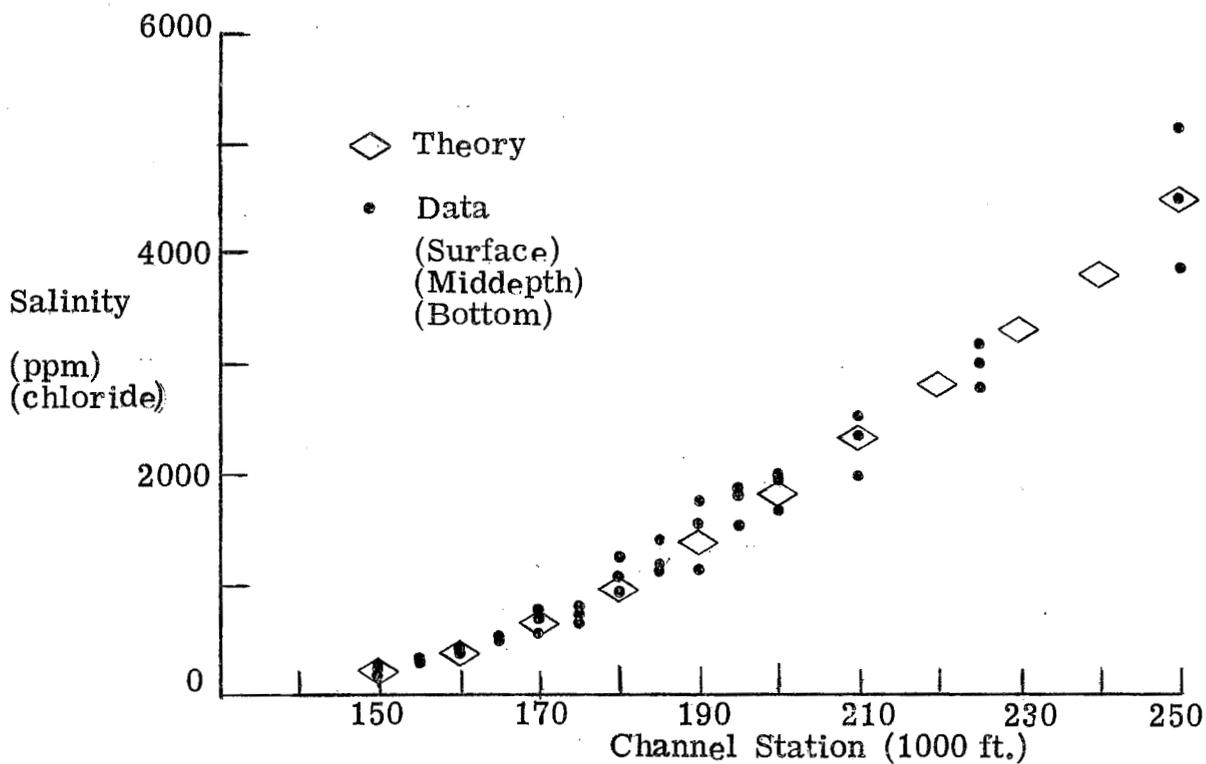


Fig. 15 Comparison of theory (est. at  $q = 13,000$  cfs) and data  $q = 13,000$  cfs

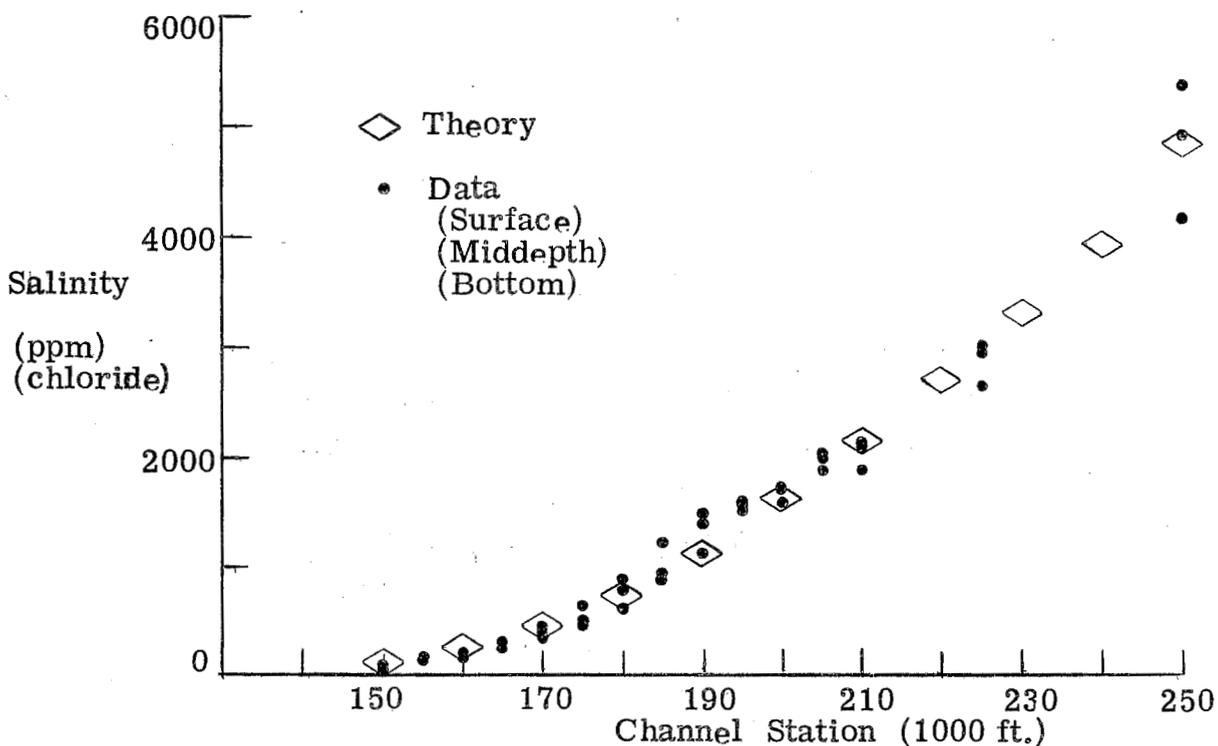


Fig. 16 Comparison of theory (est. at  $q = 13,000$  cfs) and data  $q = 16,475$  cfs

From a theoretical point of view, it is quite surprising that the results of the maximum entropy hypothesis, which at first glance appears to be a rather unspecific requirement, are physically meaningful. The resulting one cycle distribution functions are intuitively very satisfying. Also, since these one cycle distribution functions are unsteady state phenomena, and they were obtained using steady-state information, the results are still more surprising.

Much more work needs to be done using the maximum entropy estimate before its properties are better understood. In particular, the estimated one cycle distribution functions should be checked against dye dispersion data to see if, indeed, they are physically the case. Two and three dimensional analyses should be tried, since the equations and techniques which have been presented are directly applicable. Also, the maximum entropy estimate should be calculated using unsteady state data. This is the situation in an actual estuary and the results would be very interesting.

## CHAPTER VI

### RECAPITULATION

The analysis of the mixing process in an estuary is a necessary prerequisite for predicting the behavior of pollutants and dissolved oxygen in an estuary. The salinity which intrudes from the ocean end of the estuary as a result of the mixing is a convenient indicator of the nature of the mixing process. Only slightly stratified estuaries are considered since the change in the structure of the mixing process due to the changing salinity concentrations and the resulting change in density currents is small for this class of estuary.

The majority of the theories of estuarine mixing which have been proposed are based on the one-dimensional form of the convective diffusion equation. However, no solution is known for the appropriate boundary conditions which includes the harmonic tidal velocity in the convective terms of the equation. Two simplifications have been proposed to circumvent this problem. Either the average of the equation over a tidal cycle is taken or the tidal velocity terms are dropped entirely and its effects are absorbed into a longitudinal dispersion coefficient. The former simplification is justified if the fluctuations of salinity concentration, in particular, are small over a

tidal cycle. However, this is usually not the case in a tidal estuary. The latter simplification violates the assumptions under which the dispersive flux can be related to the gradient of the mean concentration. Such a relationship exists if the temporal and spatial scales of the fluctuations are small relative to the scale of the gross phenomena of interest. In an estuary, however, the tidal cycle and the tidal excursion length are usually of the same order of magnitude as the temporal and spatial scales of the gross phenomena of interest, for example, the decay time and spatial spread of a pollution cloud.

The theory of estuarine mixing devised by Preddy avoids these difficulties by adopting a more general law of mixing which relates concentrations at a distance at successive slack water times. He assumes the form of the one cycle distribution: an amount of water spreads uniformly a distance  $L$  downstream and an amount of water spreads uniformly the same distance upstream; the remainder stays where it originated. The proportions of water which spread upstream and downstream are chosen such that three conservation laws are satisfied. 1) All the water is accounted for; 2) there is no net accumulation of salt upstream from any boundary; 3) there is no net accumulation of water upstream from any boundary. These three equations are sufficient to determine the unknown proportions of water which disperse upstream and downstream. The distance  $L$  is chosen a priori as is the shape of the one cycle distribution function. These assumptions cause some doubt as to the

validity of the theory. However, the basic ideas used by Freddy form the foundation for the theory of maximum entropy mixing in estuaries.

The natural mathematical setting for a dispersion theory with discrete time intervals, e.g., a tidal cycle, between the available data is the theory of Markov chains. The estuary is segmented and each segment is a state of the chain. The concentration of salt is assumed to be uniform within each segment.

The mixing that occurs during one tidal cycle is represented by the product of two transition probability matrices,  $\underline{T}(q)$ , the translation matrix which accounts for the effect of the fresh water flow  $q$ ; and  $\underline{P}$  the mixing matrix which accounts for all the other phenomena which are responsible for the mixing during a tidal cycle. The translation matrix  $\underline{T}(q)$  is specified by assuming that the primary effect of the fresh water flow is to translate the water in each segment downstream. The mixing matrix  $\underline{P}$  must satisfy the three conservation laws: conservation of water, conservation of estuary geometry and conservation of salt. These laws are expressible as equations in terms of  $\underline{P} = (p_{ij})$ .

$$i. \quad \sum_{k=1}^N p_{kj} = 1 \quad j = 1, \dots, N;$$

$$ii. \quad \underline{Pv} = \underline{v} ;$$

$$iiia. \quad \underline{P} \underline{T}(q) \underline{sv} = \underline{sv} ,$$

where  $\underline{v}$  is the column vector of volumes and  $\underline{sv}$  is the column vector whose elements are the mass of salt in each segment at the equilibrium distribution for the fresh water flow  $q$ . Equation iiia. is a special case of the more general law which can accommodate successive unsteady state salinity distributions  $\underline{sv}(n)$  and  $\underline{sv}(n+1)$  at tidal cycle  $n$  and  $n + 1$ .

$$\text{iiib. } \underline{P} \underline{T}(q_n) \underline{sv}(n) = \underline{sv}(n+1) ,$$

where  $q_n$  is the fresh water flow rate during the time interval between the  $n^{\text{th}}$  and the  $n+1^{\text{th}}$  tidal cycle. The possibility of using unsteady-state data is an important practical advantage since most field data is collected during a period of varying fresh water flow.

The matrix  $\underline{P}$  has  $N^2$  unknown parameters  $p_{ij}$ . Equations i, ii, and iiia or iiib specify  $3N$  linear constraints on the  $p_{ij}$ . For  $N > 3$  there are more unknown  $p_{ij}$  than equations and the question is which  $\underline{P}$  shall be chosen.

The choice of  $\underline{P}$  is based on the maximum entropy principle of statistical mechanics and information theory. The  $\underline{P}$  which is chosen among all matrices which satisfy the conservation laws is the  $\underline{P}$  with the maximum entropy. A mixing process with a high entropy tends to specify broad one cycle distribution functions which do not favor any one segment over any other segment. Thus no preference is given to any segment which is not warranted by the data incorpor-

ated into the conservation laws. The entropy,  $H$ , of the Markov chain  $\underline{P}$  is given by the equation:

$$H = - \sum_{j=1}^N \frac{v_j}{\sum_{k=1}^N v_k} \sum_{i=1}^N p_{ij} \log p_{ij} .$$

This function is maximized subject to the  $3N$  equality constraints which express the conservation laws. The resulting maximum entropy estimate of  $\underline{P}$  involves  $3N$  unknown Lagrange multipliers. The  $3N$  simultaneous transcendental equations which must be solved for the Lagrange multipliers admit no general solution. A numerical procedure is found to solve these equations.

The maximum entropy estimate is applied to equilibrium salinity intrusion data collected for the Delaware River Model. The resulting one cycle distribution functions specify that mixing takes place only within the maximum tidal excursion distance. Thus the predicted one cycle distributions are physically reasonable. The estimate is based on one salinity distribution at a particular fresh water flow and the resulting mixing process is used to predict the equilibrium distributions and the other flows for which there is data. The agreement between the theoretical predictions and the data is quite close.

It is concluded, therefore, that the maximum entropy estimate provides a sound theoretical and practical solution to the problem of the characterization of the mixing process in a slightly stratified estuary.

## APPENDIX I

The difficulty with the edge condition imposed on  $P$  and  $T(q)$  by equations (12) and (29) is that no water is allowed to leave the finite region of the estuary under consideration. From the point of view of analyzing the salinity intrusion in an estuary, this condition is artificial since clearly saline water enters and leaves the mouth of the estuary during each tidal cycle. An entrance and an exit condition can be incorporated in a Markov chain but the analysis of this type of chain is more complicated. In this appendix it is shown that the equilibrium distribution of a chain with a more realistic boundary condition at the ocean end is the same as the corresponding chain which satisfies equations (12) and (29). This result justifies the use of the artificial edge condition in the analysis of equilibrium salinity distributions.

The transform technique which is employed for this analysis has been used extensively in the analysis of sampled data systems, <sup>[58]</sup> and it has been applied to the analysis of Markov chains by Sittler (1956). <sup>[59]</sup>

Consider an  $N$  state chain whose transition probability matrix is  $\underline{R} = (r_{ij})$ . This chain is represented in Figure 17 as a flow graph. The paths of flow to the states other than the adjacent states have not been drawn for the

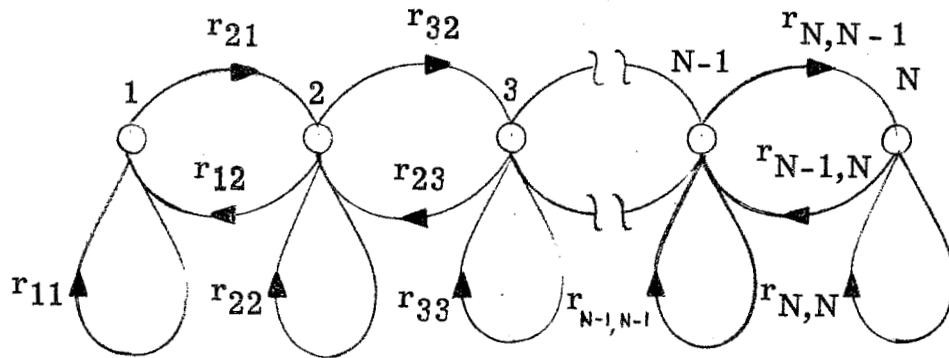


Fig. 17 Flow Graph of the Original N State Markov Chain

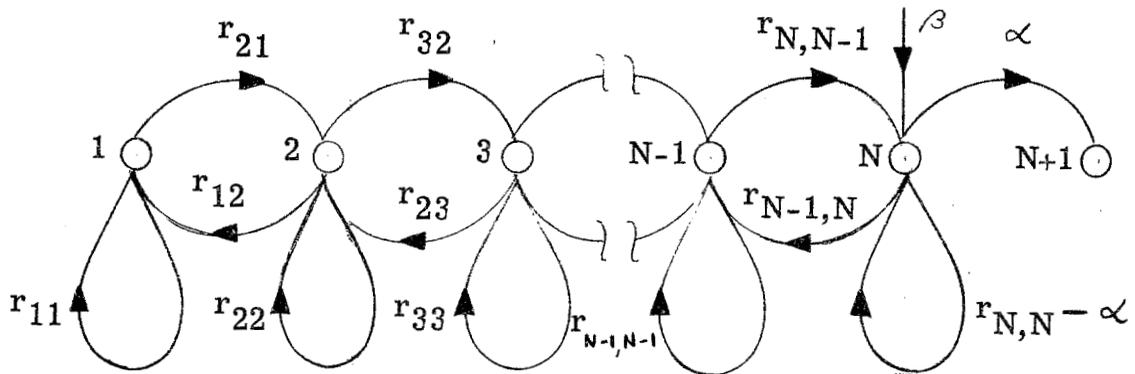


Fig. 18 Flow Graph of the Modified Markov Chain

sake of clarity, but they are assumed to exist. Let this chain satisfy the condition given by equation (12), i.e.,

$$(85) \quad \sum_{k=1}^N r_{kj} = 1 \quad j = 1, \dots, N \quad .$$

Now consider a modified version of this chain as represented in Figure 18. An additional state,  $N + 1$  has been added and it is assumed that a certain proportion,  $\alpha$ , of the water which is returned to state  $N$  in the original chain, actually goes to state  $N + 1$ . The transition probability matrix for this chain is:

$$(86) \quad \begin{matrix} r_{11} & r_{12} & \cdots & r_{1N} & 0 \\ r_{21} & r_{22} & \cdots & r_{2N} & 0 \\ \vdots & \vdots & & \vdots & \\ r_{N1} & r_{N2} & \cdots & r_{NN} - \alpha & 0 \\ 0 & 0 & \cdots & \alpha & 0 \end{matrix}$$

In terms of a mixing process in an estuary, the  $N + 1$  state is thought of as the ocean. The  $N^{\text{th}}$  state is assumed to be far enough seaward so that it is at the ocean value of salinity. Since water can leave states 1 through  $N$  and go into the ocean, a mechanism for introducing salt water at state  $N$  must also be specified. Otherwise the end result of such a mixing process is that the estuary eventually empties out into the ocean. Hence, assume that a mass,  $\beta$ , of salt water enters state  $N$  from the ocean at each high water slack. The constant  $\beta$  will be adjusted so that the equilibrium salinity in state  $N$  is at the ocean value. These conditions are

sufficient to specify the behavior of the modified chain.

For the original chain, an equilibrium distribution exists under certain conditions (Section III-B) which are assumed to hold in this case. Since the equilibrium distribution  $\underline{\sigma}$  is computed from the equations

$$(87) \quad \underline{R}\underline{\sigma} = \underline{\sigma}$$

which are a set of homogeneous linear equations in  $\underline{\sigma}$ , the solution is determined up to an arbitrary constant. This constant is evaluated by setting the value of the equilibrium salinity concentration at segment N equal to the ocean salinity.

The following theorem will now be proved:

Theorem: The equilibrium distributions of the original chain (Figure 17) and the modified chain (Figure 18) are identical.

Proof: The theorem is demonstrated by directly calculating the equilibrium distribution of the modified chain in terms of the elements of the original chain. Let the time at which the initial condition of each chain is specified be zero. Let  $s_k(n)$  be the concentration of salt in the  $k^{\text{th}}$  segment of the original chain at the  $n^{\text{th}}$  high water slack, and let  $\underline{s}(n)$  be the vector with components  $s_k(n)$ . The corresponding concentrations in the modified chain are denoted by  $s_k^{(m)}(n)$  and  $\underline{s}^{(m)}(n)$ , respectively. In the following equations the superscript (m) will refer to the modified chain.

Define the Z-transform of the sequence  $\{\underline{s}(n); n = 0, 1, \dots\}$  as:

$$(88) \quad \underline{s}^*(z) = \sum_{n=0}^{\infty} \underline{s}(n) z^n .$$

The Z-transform of the sequence  $\{\underline{s}(n)\}$  is also called the generating function of the sequence. The distribution  $\underline{s}(n+1)$  is found from the distribution  $\underline{s}(n)$  using the transition probability matrix  $\underline{R}$ :

$$(89) \quad \underline{s}(n+1) = \underline{R} \underline{s}(n) .$$

Therefore  $\underline{s}(n+1)$  is related to the initial condition  $\underline{s}(0)$  by the equation:

$$(90) \quad \underline{s}(n+1) = \underline{R}^{n+1} \underline{s}(0) .$$

Multiplying both sides of this equation by  $z^{n+1}$  and summing from  $n+1 = 0$  to  $\infty$  yields:

$$(91) \quad \underline{s}^*(z) = \left( \sum_{n=0}^{\infty} \underline{R}^n z^n \right) \underline{s}(0) .$$

Define the Z-transform of the sequence  $\{\underline{R}^n\}$  as:

$$(92) \quad \underline{R}^*(z) = \sum_{n=0}^{\infty} \underline{R}^n z^n .$$

As in the case of an ordinary geometric series, this series can be summed giving: [60]

$$\underline{R}^*(z) = (\underline{I} - z\underline{R})^{-1}$$

where  $\underline{I}$  is the identity matrix. Therefore, equation (91) becomes:

$$(94) \quad \underline{s}^*(z) = (\underline{I} - z\underline{R})^{-1} \underline{s}(0)$$

This equation gives the Z-transform of the sequence  $\{s(n)\}$  in terms of the initial condition  $\underline{s}(0)$  and the transition probability matrix  $\underline{R}$ . Hence equation (94) is the formal solution for  $\underline{s}^*(z)$ .

The equilibrium distribution  $\underline{s}(\infty)$  can be found from  $\underline{s}^*(z)$  using the final value theorem for Z-transforms: [61]

$$(95) \quad \underline{s}(\infty) = \lim_{z \rightarrow 1^+} (1 - z)\underline{s}^*(z) .$$

In order to calculate the equilibrium distributions of the original and modified chains two matrix inversions must be performed:  $(\underline{I} - z\underline{R})^{-1}$  and  $(\underline{I} - z\underline{R}^{(m)})^{-1}$ . Using Cramer's rule, [62] the inverse of a matrix  $\underline{A}$  is:

$$(96) \quad \underline{A}^{-1} = \frac{1}{\Delta} (A_{ji})$$

where  $\Delta$  is the determinant of  $\underline{A}$  and  $A_{ij}$  is the  $ij^{\text{th}}$  cofactor of  $\underline{A}$ , i.e.,  $(-1)^{i+j}$  times the determinant formed by striking out the  $i^{\text{th}}$  row and  $j^{\text{th}}$  column of the determinant of  $\underline{A}$ . In this case let

$$(97) \quad \begin{aligned} \underline{A} &= \underline{I} - z\underline{R} \\ \underline{A}^{(m)} &= \underline{I} - z\underline{R}^{(m)} . \end{aligned}$$

Only one column of each inverse is needed because of the simple initial conditions and sources which have been adopted.

For the modified chain an amount  $\beta$  of salt water is introduced into state  $N$  at every high water slack. Hence the source terms is a sequence  $\{\beta, \beta, \dots\}$  at state  $N$  and zero for all other states. The Z-transform of this sequence is  $\beta/(1-z)$ . By the convolution theorem for Z-transforms, [63] the response of the modified chain to this source term is given by multiplying  $\underline{R}^{*(m)}(z)$  by the column vector,  $(0, 0, \dots, \beta/(1-z), 0)^T$  where  $T$  denotes the transpose. Since the only non-zero term of this vector is in the  $N^{\text{th}}$  row, only the  $N^{\text{th}}$  column of the inverse is required as all the other elements of the inverse are multiplied by zero. Therefore the  $k^{\text{th}}$  component of  $\underline{s}^{*(m)}(z)$  is given by:

$$(98) \quad s_k^{*(m)}(z) = \frac{A_{Nk}^{(m)}}{\Delta^{(m)}} \frac{\beta}{1-z} .$$

For the original chain  $\underline{R}$ , the equilibrium distribution of the chain is independent of the initial condition since there are no source terms and equation (85) as well as the positivity condition discussed in Section III-B are satisfied. Hence any initial condition is allowable. For simplicity assume that at time zero a quantity  $\gamma$  of salt is in state  $N$ , and all other states are empty. Therefore the initial condition vector  $\underline{s}(0)$  is given by:

$$(99) \quad \underline{s}(0) = (0, 0, \dots, \gamma)^T$$

and the  $k^{\text{th}}$  component of  $\underline{s}^*(z)$  is given by:

$$(100) \quad s_k^*(z) = \frac{A_{Nk}}{\Delta} r \quad .$$

Because of the similarity of the modified chain and the original chain, it is possible to express  $s_k^{*(m)}(z)$  in terms of  $s_k^*(z)$ .

Consider the determinant  $\Delta^{(m)} = |\underline{I} - z\underline{R}^{(m)}|$  in equation (98). The elements of the last column of  $\underline{R}^{(m)}$  are all zeros since no water goes from the  $(N + 1)^{th}$  state to any other state (equation (86)). Hence the elements of the last column of the matrix  $\underline{I} - z\underline{R}^{(m)}$  are all zero except for element  $N + 1, N + 1$  which is equal to 1. Therefore the determinant  $|\underline{I} - z\underline{R}^{(m)}|$  can be immediately expanded in terms of the cofactors of the last column<sup>[64]</sup> to give:

$$(101) \quad \Delta^{(m)} = \begin{vmatrix} 1-z r_{11} & -z r_{12} & \dots & \cdot \\ -z r_{21} & 1-z r_{22} & \dots & \cdot \\ \cdot & & & -z r_{N-1,N} \\ \cdot & & & 1-z z(r_{N,N-\alpha}) \end{vmatrix}$$

Expanding this determinant once again in the cofactors of the last column gives:

$$(102) \quad \Delta^{(m)} = \Delta + \alpha z A_{NN}$$

Hence  $\Delta^{(m)}$  can be written in terms of the determinant and a cofactor of the original chain.

Consider the cofactor  $A_{Nk}^{(m)}$  in equation (98). When calculating this cofactor the  $N^{th}$  row of  $\underline{A}^{(m)} = \underline{I} - z\underline{R}^{(m)}$

is always deleted for any  $k = 1, \dots, N$ . For example, for  $k = 2$ , the cofactor is given by:

$$(103) \quad A_{N2}^{(m)} = (-1)^{N+2} \begin{vmatrix} 1-z r_{11} & -z r_{13} & \dots & \cdot & 0 \\ -z r_{21} & -z r_{23} & \dots & \cdot & 0 \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & -z r_{N-1,N} & 0 \\ \cdot & \cdot & & -z \alpha & 1 \end{vmatrix}$$

Expanding this determinant in the cofactors of the last column gives:

$$(104) \quad A_{Nk}^{(m)} = A_{Nk}$$

Rewriting equation (98) for  $s_k^{*(m)}(z)$  in terms of equations (102) and (104) yields:

$$(105) \quad s_k^{*(m)}(z) = \frac{A_{Nk}}{(\Delta + \alpha z A_{NN})} \cdot \frac{\beta}{1-z}$$

$$(106) \quad = \frac{s_k^*(z)/r}{(1 + z\alpha s_N^*(z)/r)} \cdot \frac{\beta}{1-z}$$

where equation (100) has been used to obtain equation (106). Using the final value theorem (equation (95)):

$$(107) \quad s_k^{(m)}(\infty) = \lim_{z \rightarrow 1^+} \frac{s_k^*(z)/r}{1 + \alpha z s_N^*(z)/r} \cdot \beta z^{-1}$$

However, for  $z$  close to 1,

$$(108) \quad s_k(\infty) = (1-z) s_k^*(z) \Big|_{z \rightarrow 1^+}$$

again by the final value theorem. Hence:

$$(109) \quad s_k^{(m)}(\infty) = \lim_{z \rightarrow 1^+} \frac{\frac{s_k(\infty)}{\gamma(1-z)}}{1 + \frac{z\alpha s_N(\infty)}{\gamma(1-z)}} \beta z^{-1}$$

$$(110) \quad = \frac{s_k(\infty)}{s_N(\infty)} \frac{\beta}{\alpha}$$

Applying the boundary condition at the  $N^{\text{th}}$  segment, that  $s_N^{(m)}(\infty) = s_N(\infty) =$  ocean salinity value, determines  $\beta$ . Hence

$$s_k^{(m)}(\infty) = s_k(\infty)$$

which proves the theorem.

## BIBLIOGRAPHY

- [1] O'Connor, Donald, J., "Oxygen Balance of an Estuary"  
J. Sanitary Engr. Div., Am. Soc. of Civil Engr. Vol. 86,  
No. SA3 (May 1960).
- [2] Effects of Polluting Discharges on the Thames Estuary,  
Water Pollution Research Technical Paper No. 11, Her  
Majesty's Stationary Office, London (1964), Ch. 14,  
"Tidal Mixing," p. 390.
- [3] Keulegan, G. H., "The Mechanism of an Arrested Saline  
Wedge," Ch. 11, Estuary and Coastline Hydrodynamics,  
ed. Arthur T. Ippen, McGraw-Hill Book Co., New York,  
(1966).
- [4] Harleman, Donald R. F., "The Significance of Longitudi-  
nal Dispersion in the Analysis of Pollution in Estuaries,"  
Second International Conference on Water Pollution Re-  
search, Tokyo, Japan (August 1964), paper No. 13, section  
IV, pp. 4-6.
- [5] Pritchard, Donald W., "The Equations of Mass Continuity  
and Salt Continuity in Estuaries," J. Marine Research,  
Vol. 17 (1958), pp. 412-423.
- [6] Batchelor, G. K.; Townsend, A. A., "Turbulent Diffusion,"  
Surveys in Mechanics, Cambridge Univ. Press (1956), p. 360.
- [7] Pritchard, (1958) p. 417.
- [8] Arons, Arnold B.; Stommel, H., "A Mixing Length Theory  
of Tidal Fusing," Trans. Am. Geophysical Union, Vol. 32,

- No. 3 (June 1951), pp. 419-421.
- [9] Stommel, H., "Computation of Pollution in a Vertically Mixed Estuary," *Sewage and Industrial Wastes*, Vol. 25, No. 9 (Sept. 1953), pp. 1065-1071.
- [10] Pritchard, D. W., "A Study of the Salt Balance in a Coastal Plain Estuary," *J. Marine Research*, Vol. 13, No. 1 (1954), p. 135.
- [11] Pritchard, D. W., "The Dynamic Structure of Coastal Plain Estuary," *J. Marine Research*, Vol. 13, No. 1 (1956).
- [12] Okubo, Akira, "Equations Describing the Diffusion of an Introduced Pollutant in a One-dimensional Estuary," Studies on Oceanography, ed. Koji Hidaka, Univ. of Washington Press (1965), pp. 217-226.
- [13] Okubo, (1965) p. 217.
- [14] Okubo, (1965) pp. 221-222.
- [15] Cohen, Bernard; McCarthy, Leo. T., Jr., Salinity of the Delaware Estuary, Geological Survey Water-Supply Paper 1586-B, U.S. Government Printing Office, Washington. (1962), p. B23, Fig. 12.
- [16] Thames, (1964), Ch. 14, "Tidal Mixing," p. 391, Fig. 219.
- [17] Harris, E. K., "A New Statistical Approach to the One-dimensional Diffusion Model," *International J. of Air and Water Pollution*, Vol. 7 (1963), pp. 799-812.

- [18] Glover, R. E., "A New Method of Predicting Transient States of Salinity Intrusion into the Sacramento-San Joaquin Delta," Trans. Am. Geophysical Union, Vol. 36, No. 4 (August 1955), pp. 641-648.
- [19] Harleman, (1964).
- [20] Harleman, (1964), p. 3.
- [21] Stommel, H. (1953), pp. 1066-1067
- [22] Lightfoot, E. N.; Cussler, E. L., Jr., Selected Topics in Transport Phenomena, Chemical Engr. Progress Symposium Series, Am. Institute of Chemical Engr., Vol. 61, No. 58 (1965).
- [23] Thames (1964), Ch. 14, "Tidal Mixing", pp. 392-395.
- [24] Miller, E. G., Observations of Tidal Flow in the Delaware River, Geological Survey Water-Supply Paper 1586-C, U.S. Government Printing Office, Washington D.C., (1962), p. C24, Table 1.
- [25] Ketchum, Bostwick, H., "The Exchange of Fresh and Salt Waters in Tidal Estuaries," J. Marine Research, Vol. 10, No. 1 (1951), pp. 18-38.
- [26] Thames (1964), Ch. 14, "Tidal Mixing", p. 392.
- [27] Preddy, W. S., "The Mixing and Movement of Water in the Estuary of the Thames," J. Marine Biology of the United Kingdom, Vol. 33 (1954), pp. 645-662.
- [28] Preddy (1954), p. 650.
- [29] Preddy (1954), p. 650.

- [30] Thames (1964), Ch. 15, "Methods of Calculating Distributions of Dissolved Substances," pp. 407-426.
- [31] Pyatt, E., On Determining Pollutant Distribution in Tidal Estuaries, Geological Survey Water-Supply Paper 1586-F, U.S. Government Printing Office, Washington D.C., p. F25.
- [32] Preddy (1954), p. 647.
- [33] Matalas, N.C.; Conover, W.J., "Derivation of the Velocity Profile from a Statistical Model of Turbulence," Water Resources Research, Vol. 1, No. 2 (1965).
- [34] Bugliarello, G.; Jackson, E.D., III, "Random Walk Study of Convective Diffusion," J. Engr. Mech. Div., Proc. Am. Soc. of Civil Engr., Vol. EM4 (August 1964), pp. 49-77.
- [35] Bharucha-Reid, A. T., Elements of the Thoery of Markov Processes and Their Applications, McGraw-Hill Book Co., New York (1960), Ch. 1.
- [36] Bellman, Richard, Introduction to Matrix Analysis, McGraw-Hill Book Co., New York (1960), Ch. 14, pp. 256-258.
- [37] Frechet, Maurice, Traite du Calcul des Probabilites et de ses Applications, Gauthier-Villars, Paris (1952), Tome 1, Fasc. III, pp. 33-35.
- [38] Jaynes, E. T., Probability Theory in Science and Engineering, Colloquium Lectures in Pure and Applied Science, No. 4, Field Reserach Laboratories, Socony Mobil Oil Co., Inc. (Feb. 1958).

- [39] Jaynes, E. T., "New Engineering Applications of Information Theory," Proceedings First Symposium on Engineering Applications of Random Function Theory and Probability, ed. Bogdanoff and Kozin, J. Wiley and Sons (1963).
- [40] Jaynes, E. T., "Information Theory and Statistical Mechanics," Physical Review, Vol. 106, No. 4 (May 1957) pp. 620-630.
- [41] Jaynes, E. T. (1958), pp. 132-134.
- [42] Huang, Kerson, Statistical Mechanics, J. Wiley and Sons (1964), pp. 206-212.
- [43] Jaynes, E. T. (1957), p. 620.
- [44] Shannon, C. E., "A Mathematical Theory of Communication," Bell System Technical J., Vol. 27(1948), pp. 379-423, pp. 623-656.
- [45] Khinchin, A. I., Mathematical Foundations of Information Theory, Dover (1957), pp. 13-16.
- [46] Stiefel, Eduard, L., An Introduction to Numerical Mathematics, Academic Press, New York (1963), pp. 79-83.
- [47] Jaynes, E. T., (1958), p. 125.
- [48] Stommel (1953), p. 1065.
- [49] Eggink, H. J., "Predicted Effects of Future Discharges of Industrial Wastes into the Eems Estuary," Third International Conference on Water Pollution Research, Munich (1966), p. 6.

- [50] Salinity Tests of Existing Channel, Delaware River Model Study Report No. 2. Waterways Experiment Station, Vicksburg, Mississippi (June 1954).
- [51] Moody, David, W., Personal Communication, U.S. Geological Survey, Branch of Quality of Water, Philadelphia, Pa (March 1966).
- [52] Salinity Tests (1954), p. 6.
- [53] Taylor, G. I., "The Dispersion of Matter in Turbulent Flow Through a Pipe," Proc. Royal Soc. of London, Vol. 223 (1954), pp. 446-468.
- [54] Fischer, Hugo, B., "Longitudinal Dispersion in Laboratory and Natural Streams," Report No. KH-R-12, W. M. Leck Laboratory of Hydraulics and Water Resources, California Institute of Technology, Pasadena, Calif. (June 1966).
- [55] Salinity Tests (1954), "Velocity Observations," plate 5.
- [56] Miller (1962), pp. C24-C26.
- [57] Harleman, D.R.F; Ippen, A. T.; Lin, J.D., "Turbulent Diffusion and Gravitational Convection in an Idealized Estuary," Tech. Report No. 38, Hydrodynamics Laboratory, Massachusetts Institute of Technology (March 1960).
- [58] Ragazzini, J. R.; Zadeh, L. A., "Analysis of Sampled Data Systems," Trans. Am. Inst. Electrical Engrs., Vol. 71, part 2 (Nov. 1952), p. 225.

- [59] Sittler, R. W., "Systems Analysis of Discrete Markov Processes," *Inst. Radio Engrs., Trans. on Circuit Theory* (Dec. 1956), pp. 257-266.
- [60] Sittler (1956), p. 261.
- [61] Ragazzini, J. R.; Franklin, G. F., Sampled-Data Control Systems, McGraw-Hill, Book Co., N. Y. (1958), pp. 61-63.
- [62] Pease, Marshall C, III, Methods of Matrix Analysis, Academic Press, N. Y. (1965), p. 30.
- [63] Ragazzini (1958), pp. 66-68.
- [64] Bellman (1960), Appendix A, p. 304.

## ABSTRACT

A review and a critique of the theories of estuarine mixing which have been proposed is presented and it is concluded that the simplifications usually employed in the mixing theories based on the convective diffusion equation are not applicable to the mixing process in an estuary. The theory of tidal mixing which has been proposed by Preddy is discussed and his approach forms the basis for the theory of maximum entropy mixing which is developed. The analysis of the mixing process in an estuary is formulated in terms of the theory of Markov chains. Three conservation laws which any physically reasonable mixing process must satisfy are formulated and related to the properties of a Markov chain. The estimate of the appropriate mixing matrix is based on the maximum entropy principle of statistical mechanics and information theory. A numerical technique is presented for the solution of the resulting simultaneous transcendental equations. The equilibrium salinity intrusion data from the Delaware River Model is analyzed and compared with the theoretical predictions based on the maximum entropy estimate of the mixing process. The resulting agreement is noted and it is concluded that the theory of maximum entropy mixing is a sound theoretical and practical solution to the problem of characterizing the mixing process in an estuary.