AN ABSTRACT OF THE THESIS OF

KEMAL GÜRER for the degree of Master of Science in Atmospheric Science presented on April 21, 1987.

Title: Concentration Distributions of Non-buoyant, Weakly Buoyant and Buoyant Effluents From a Continuous Point Source Within a Convectively Mixed Layer.

Abstract approved: Redacted for Privacy

This study examines the concentration data which were obtained with the Deardorff-Willis boundary layer tank, with emphasis on the near surface concentrations. Previous studies of air pollutant concentration distribution are also reviewed. The degree of plume buoyancies examined here are F_*=0, 0.03, 0.116 and 0.26, where the last value of F_* corresponds to a highly buoyant effluent. A double-log normal distribution function is introduced and is compared with the gamma and the log-normal distributions as candidates to fit the data. The exponential distribution is also considered. The gamma distribution provides a better fit than the double-log-normal distribution for the F_*=0 (non-buoyant) and F_*=0.03 (weakly buoyant) cases. On the other hand, the double-log-normal distribution provides a better fit than the

gamma distribution for the $F_{\star}$=0.116 (buoyant) and $F_{\star}$=0.26 (highly buoyant) cases since the double-log-normal has an additional free parameter.

CONCENTRATION DISTRIBUTIONS OF NON-BUOYANT, WEAKLY BUOYANT AND BUOYANT EFFLUENTS FROM A CONTINUOUS POINT SOURCE WITHIN A CONVECTIVELY MIXED LAYER

by

Kemal Gürer

A THESIS

submitted to

Oregon State University

in partial fulfillment of

the requirements for the

degree of

Master of Science

Completed April 21, 1987

Commencement June 1987

APPROVED:

Redacted for Privacy

Professor of Atmospheric Sciences in charge of major

Redacted for Privacy

Chairman of the Department of Atmospheric Sciences

Redacted for Privacy

Dean of Graduate School

Date thesis is presented April 21, 1987

Typed by Kemal Gürer

ACKNOWLEDGEMENTS

I am deeply indebted to my major professor James W. Deardorff for his guidance, support and understanding throughout this study. My thanks are due to Professors Allan H. Murphy, Larry Mahrt and Steven K. Esbensen for their valuable comments in the completion of my thesis. Appreciation is also due to Nimal Gamage, Paul Ruscher, Martin Ehrendorfer and Tom Sabin for their assistance in forming and writing the thesis.

I also would like to thank my wife Sennur for her encouragement and understanding.

TABLE OF CONTENTS

SECTION		PAGE
1.	Introduction	1
2.	Previous studies on the concentration distribution of pollutants	3
3.	The method of experiments and the description of the boundary-layer tank	14
4.	The gamma probability density function	20
5.	The exponential probability density function	24
6.	Log-normal probability density function	26
7.	The double-log-normal probability density function	28
8.	The intermittent nature of the plume	35
9.	Results and discussion	38
10.	Conclusions	50
11.	Suggested future studies	51
12.	References	104

LIST OF FIGURES

FIGURE		PAGE
1	Schematic diagram of the boundary layer model tank and its sections from the centerline in the y-direction with the use of its standard deviation σ_y and its sections in the z-direction from 0 to 1.2 h, mixed layer height.	52
2.a	The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $\|y\|<\sigma_y$ and $0<h<0.4$ at $X=1.02$ for $F_*=0$ case.	53
2.b	The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $\|y\|<\sigma_y$ and $0<h<0.4$ at $X=1.46$ for $F_*=0$ case.	54
2.c	The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $\|y\|<\sigma_y$ and $0<h<0.4$ at $X=1.90$ for $F_*=0$ case.	55
2.d	The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $\|y\|<\sigma_y$ and $0<h<0.4$ at $X=2.413$ for $F_*=0$ case.	56

2.e The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=2.92$ for $F_*=0$ case. 57

2.f The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=3.34$ for $F_*=0$ case. 58

2.g The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=3.79$ for $F_*=0$ case. 59

3.a The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=1.15$ for $F_*=0.03$ case. 60

3.b The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=1.96$ for $F_*=0.03$ case. 61

3.c The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=2.6$ for $F_*=0.03$ case. 62

3.d The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=3.56$ for $F_*=0.03$ case. 63

3.e The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=4.28$ for $F_*=0.03$ case. 64

3.f The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=4.96$ for $F_*=0.03$ case. 65

4.a The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=1.14$ for $F_*=0.116$ case. 66

4.b The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=2.05$ for $F_*=0.116$ case. 67

4.c The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=2.87$ for $F_*=0.116$ case. 68

4.d The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=3.57$ for $F_*=0.116$ case. 69

4.e The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=4.15$ for $F_*=0.116$ case. 70

5.a The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=1.19$ for $F_*=0.26$ case. 71

5.b The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=1.62$ for $F_*=0.26$ case. 72

5.c The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=2.26$ for $F_*=0.26$ case. 73

5.d The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=3.03$ for $F_*=0.26$ case. 74

5.e The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=3.91$ for $F_*=0.26$ case. 75

5.f The probability of occurrences of log-concentration, $p_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=4.96$ for $F_*=0.26$ case. 76

6.a The probability of occurrences of concentration, $p_i(C_p)$, versus the concentration for $|y|<\sigma_y$ and $0<h<0.4$ for $F_*=0$ case at $X=1.02$. 77

6.b The probability of occurrences of concentration, $p_i(C_p)$, versus the concentration for $|y|<\sigma_y$ and $0<h<0.4$ for $F_*=0.03$ case at $X=1.15$. 78

6.c The probability of occurrences of concentration, $p_i(C_p)$, versus the concentration for $|y|<\sigma_y$ and $0<h<0.4$ for $F_*=0.116$ case at $X=1.14$. 79

6.d The probability of occurrences of concentration, $p_i(C_p)$, versus the concentration for $|y|<\sigma_y$ and $0<h<0.4$ for $F_*=0.26$ case at $X=1.19$. 80

7.a The probability of occurrences of concentration, $p_i(C_p)$, versus the concentration for $|y|<\sigma_y$ and $0<h<0.4$ for $F_*=0$ case at $X=3.79$. 81

7.b The probability of occurrences of concentration, $p_i(C_p)$, versus the concentration for $|y|<\sigma_y$ and $0<h<0.4$ for $F_*=0.03$ case at $X=4.28$. 82

7.c The probability of occurrences of concentration, $p_i(C_p)$, versus the concentration for $|y|<\sigma_y$ and $0<h<0.4$ for $F_*=0.116$ case at $X=4.15$. 83

7.d The probability of occurrences of concentration, $p_i(C_p)$, versus the concentration for $|y|<\sigma_y$ and $0<h<0.4$ for $F_*=0.26$ case at $X=4.96$. 84

8.a Skewness and kurtosis values of the experimental data versus downwind distance X for $F_*=0$ case for $|y|<\sigma_y$ and $0<h<0.4$. 85

8.b Skewness and kurtosis values of the experimental data versus downwind distance X for $F_*=0.03$ case for $|y|<\sigma_y$ and $0<h<0.4$. 86

8.c Skewness and kurtosis values of the experimental data versus downwind distance X for $F_*=0.116$ case for $|y|<\sigma_y$ and $0<h<0.4$. 87

8.d Skewness and kurtosis values of the experimental data versus downwind distance X for $F_*=0.26$ case for $|y|<\sigma_y$ and $0<h<0.4$. 88

9.a The mode and its frequency for the experimental data versus the downwind distance X for $F_*=0$ case for $|y|<\sigma_y$ and $0<h<0.4$. 89

9.b The mode and its frequency for the experimental data versus the downwind distance X for $F_*=0.03$ case for $|y|<\sigma_y$ and $0<h<0.4$. 90

9.c The mode and its frequency for the experimental data versus the downwind distance X for F_*=0.116 case for $|y|<\sigma_y$ and 0<h<0.4. 91

9.d The mode and its frequency for the experimental data versus the downwind distance X for F_*=0.26 case for $|y|<\sigma_y$ and 0<h<0.4. 92

10.a Both left and right standard deviations of the experimental data versus distance X for F_*=0 case for $|y|<\sigma_y$ and 0<h<0.4. 93

10.b Both left and right standard deviations of the experimental data versus distance X for F_*=0.03 case for $|y|<\sigma_y$ and 0<h<0.4. 94

10.c Both left and right standard deviations of the experimental data versus distance X for F_*=0.116 case for $|y|<\sigma_y$ and 0<h<0.4. 95

10.d Both left and right standard deviations of the experimental data versus distance X for F_*=0.26 case for $|y|<\sigma_y$ and 0<h<0.4. 96

11.a The mean and the standard deviation of the experimental data versus distance X for F_*=0 case for $|y|<\sigma_y$ and 0<h<0.4. 97

11.b The mean and the standard deviation of the experimental data versus distance X for F_*=0.03 case for $|y|<\sigma_y$ and 0<h<0.4. 98

11.c The mean and the standard deviation of the experimental data versus distance X for F*=0.116 case for $|y|<\sigma_y$ and $0<h<0.4$. 99

11.d The mean and the standard deviation of the experimental data versus distance X for $F_*=0$ case for $|y|<\sigma_y$ and $0<h<0.4$. 100

LIST OF TABLES

TABLE		PAGE
1	Negative and positive differences between the expected and measured frequencies of concentrations, their sum of absolute values and total sum of square of error for the double-log-normal and the gamma distributions at $F_*=0$ (non-buoyant) case.	101
2	Negative and positive differences between the expected and measured frequencies of concentrations, their sum of absolute values and total sum of square of error for the double-log-normal and the gamma distributions at $F_*=0.03$ (weakly buoyant) case.	101
3	Negative and positive differences between the expected and measured frequencies of concentrations, their sum of absolute values and total sum of square of error for the double-log-normal and the gamma distributions at $F_*=0.116$ (buoyant) case.	102
4	Negative and positive differences between the expected and measured frequencies of concentrations, their sum of absolute values and total sum of square of error for the double-log-normal and the gamma distributions at $F_*=0.26$ (highly buoyant) case.	102

5 The chi-square values for the double-log-normal and the gamma distributions for $F_{\star}$=0 case. 103

6 The chi-square values for the double-log-normal and the gamma distributions for $F_{\star}$=0.03 case. 103

7 The chi-square values for the double-log-normal and the gamma distributions for $F_{\star}$=0.116 case. 103

8 The chi-square values for the double-log-normal and the gamma distributions for $F_{\star}$=0.26 case. 103

CONCENTRATION DISTRIBUTIONS OF NON-BUOYANT, WEAKLY BUOYANT AND BUOYANT EFFLUENTS FROM A CONTINUOUS POINT SOURCE WITHIN A CONVECTIVELY MIXED LAYER

1. INTRODUCTION

Effluent concentration values are typically measured in the atmosphere at receptors placed at various distances downwind from a stack emitting a plume. Measurements are taken at 5 - 10 min intervals, for example CAMP (Continuous Air Monitoring Program) data used by Knox and Pollack (1974), for a typical averaging time period of 30 - 60 min. When these concentration values are plotted on a diagram of time versus concentration, it is generally found that significant concentrations may occur for only a fraction of the time, and that they show extreme fluctuations above the mean value. Due to the probabilistic nature of pollutant concentration, the observed distribution of concentration is described by a sample probability density function which gives the relative frequencies of concentration values over finite intervals of the sample space. If the true probability distribution or density function of the concentration can be determined, then by using this function the probability of occurrence of a particular concentration or the probability of occurrences of

some concentration values which are less or greater than some specific value may be determined. Then the degree of emission control required to meet air quality standards may be estimated. For example, after air pollution controls are introduced, it will be desirable to determine their effectiveness. In order to make a decision on this, one can use standard statistical procedures if the frequency distribution of concentrations is known. Thus, estimating a particular concentration which is expected to occur for a particular time in a certain time period, for example 0.03 ppm concentration for 8-hr duration samples once a year, can guide environmental control officials to take the necessary measures to decrease the hazardous effect of pollutants. In addition, sampling for pollution monitoring can be most effectively designed if the statistical properties of the temporal distribution of air pollution events are known.

2. PREVIOUS STUDIES ON THE CONCENTRATION DISTRIBUTION OF POLLUTANTS

Ott and Mage (1976) treated the diffusion equation as a stochastic differential equation by assuming that the concentration of pollutant C_s entering from a source with the volume is much greater than the concentration of pollutant C. Then the diffusion equation is written for a pollutant carrier medium

$$\frac{\partial C}{\partial t} + VC = U \tag{8.1}$$

where

$$V = K_1 + \frac{1}{C}\left[u \frac{\partial C}{\partial x} + v \frac{\partial C}{\partial y} + w \frac{\partial C}{\partial z} - D\nabla^2 C \right] \tag{8.2}$$

$$U = K_1 C^* + K_2 C_s - K_3 \tag{8.3}$$

Here K_1 is the homogeneous chemical reaction rate coefficient, K_2 dilution coefficient, K_3 zeroth order heterogeneous chemical rate coefficient, D isotropic diffusion coefficient, C_s initial concentration of pollutant emitted by source within the volume, C^* concentration of pollutant if in chemical equilibrium with all other species present. K_1, K_2, K_3, C_s, C^*, D, u, v, w are all treated as random variables independent of C. Then by assuming U and V as independent random variables arising from unknown distributions with finite variance, Ott and Mage approximated

the diffusion equation as a discrete time series and focused on the case where U and V are linearly related as $U_i=KV_i$ with $r_{UV}=+1$, the correlation coefficient between U and V. By simplifying the diffusion equation in this way and taking its logarithm, they obtained

$$\log(C_{n+1} - K) = \log(C_1 - K) + \sum_{i=1}^{n} \log(1 - V_i \, \Delta t) \qquad (8.4)$$

and by making use of the central limit theorem the 3-parameter log-normal distribution is given in the form

$$f(C) = \frac{1}{\sigma(C-K)\sqrt{2\pi}} \exp\left[- \frac{1}{2} \left[\frac{\ln(C-K) - \mu}{\sigma} \right]^2 \right] \qquad C>K \qquad (8.5)$$

where K is the third parameter whose magnitude is determined by a graphical trial-and-error process in plotting the data values on a logarithmic-paper. They applied the 3-parameter log-normal to 10 air quality data sets of SO_2, O_3, CO, particulate, hydrocarbons and NO_2 from the U.S.A., France, West Germany and Denmark and to 9 water quality data sets of BOD (Biochemical Oxygen Demand), coliform, chloride and sulfate from the Ohio River. In application of curve fitting to the observed data, the zero concentration values are seen as the area under the density curve which corresponds to the intermittent nature of the concentration. This portion of the curve, to the left of the origin, is separated and represented by a discrete probability value. In their model, the successive instantaneous concentrations are serially

correlated. Thus the time averages of the instantaneous concentrations are constrained to keep their log-normal characteristics if the averaging times are short, for example 1 hr to 8 hr.

Ott and Mage showed that the 3-parameter log-normal was a better fit than the 2-parameter log-normal. Since the plot of the data values on logarithmic-paper in a cumulative sense for the log-normal distribution is seen as a straight line, the third parameter K was adjusted to satisfy this condition. The 2-parameter log-normal distribution, while being the same as the 3-parameter log-normal distribution but without K, showed a downward curvature on the logarithmic paper. The error for the 3-parameter log-normal distribution was also found to be significantly lower than for the 2-parameter log-normal for the 19 air and water quality data sets they examined.

Larsen (1969) applied his mathematical model to air pollutant concentration data of CO, hydrocarbons, nitric oxide, nitrogen oxides (NO, NO_2), oxidant and sulphur dioxide in 6 U.S. states for 3 years from December-1961 to December-1964. He concludes from the observations that the air pollutant concentration follows a log-normal distribution for all pollutants in urban areas for a sample duration time of an hour and for all averaging times over which all samples were collected from 5 minutes to a year. The maximum concentration is found to be inversely proportional to the

averaging time raised to an exponent which is a function of the standard geometric deviation.

Cats and Holtslag (1979) sought the dependence of air pollution frequency distribution on wind direction in the Netherlands by using sulphur dioxide data sets from 1971 to 1974 at three stations. They divided the data sets consisting of hourly observations into a total of 18 subsets according to wind direction within a 20^{o}sector. They pointed out that the concentration was mainly determined by wind direction and assumed that the other factors such as wind speed and emission statistics would lead to similar concentration frequency distributions within each wind direction subset. The frequency distribution followed the log-normal at urban areas. The log-normal fitted the overall 98 percentiles of the observed data by deviating at most 15 % from the observed 98 percentiles. They emphasized that the data sets used in this study involved background pollution since it was collected in urban areas.

Lynn (1974) used the suspended particulate data collected in Philadelphia at three stations from 1960 to 1968. He tested the 2-parameter and 3-parameter log-normal, gamma and Pearson curves to represent the observed data by using the method of moments and its four parameters: mean, standard deviation, skewness and kurtosis. As a comparison criterion between the observed and the expected frequencies, the sum of the absolute differences was used. He showed,

based on the data he tested and on the fitting procedure, that the 2-parameter log-normal does slightly better than all the others do, even though in some instances all four of them almost equally represent the data and some even do better than the log-normal. As a reason for the 2-parameter log-normal to fit the data better than the others, he gave the insensitivity of the 2-parameter log-normal to especially extreme values of concentration.

Shoji and Tsukatani (1973) showed that a log-normal distribution usually approximates the random variation of air pollutant concentration based on the record of SO_2 concentration in the atmosphere and of the tracer (fluorescent pigments) concentration for 30 min to 1 h release time in diffusion experiments. They tried to establish some air quality standards for SO_2 in Japan and showed the reason for the mechanism of the occurrence of the log-normal distribution by giving the hypotheses which were introduced by Chow in 1955 and Kolmogorov in 1941. Chow showed that when the variate is composed of the multiplication of many factors which are independent, the variate has a log-normal distribution. Kolmogorov also proved that the log-normal distribution applies in the case of particles which undergo splitting. Shoji and Tsukatani used the geometric mean in the log-normal probability density

function of s-hour averaged time series of X(t),

$$B_s(x) = \frac{1}{\sigma(s)\sqrt{2\pi}} \exp\left[-\frac{(x-\bar{x})^2}{2\sigma^2(s)}\right] \qquad (9.1)$$

where

$$\sigma(s,t) = \frac{1}{s}\int_0^s x(t+u)\, du \qquad (9.2)$$

is the geometric standard deviation. X(t) is the logarithmic value of C(t); C(t) is the air pollutant concentration at time t.

Knox and Lange (1974) presented their work on the surface air pollutant concentration frequency distribution for area sources and continuous point sources. They particularly limited their attention to the most inert pollutant CO and gave the relationship between the concentration of pollutants emitted into a steady-state unidirectional flow and wind speed for a continuous urban area source by

$$\ln(C/Q)=\ln K_1-\ln U \qquad (10.1)$$

and for a continuous point source by

$$\ln(C/Q)=\ln K_2-\ln(\sigma_a\sigma_e U) \qquad (10.2)$$

where

C:average concentration K_1, K_2:proportionality constants

Q:source strength σ_a:lateral standard deviation

U:mean wind speed σ_e:vertical " "

They showed that in either case a log-normal distribution provided a reasonable fit over most of the frequency range with an implication of a bigger geometric standard deviation (1.9 ppm) in the urban area case than that (5.0 ppm) in the continuous point source case. The log-normal distribution, in this case, is associated with the reciprocal of the wind speed.

Bencala and Seinfeld (1976) tried to understand the observation that the frequency distributions tend to be log-normal regardless of averaging time changing from 10 min to a year for the reading of concentrations at every 5 min interval. They used the data of CAMP (Continuous Air Monitoring Program) for the years 1962-1968 and restricted their attention to carbon monoxide. An unweighted least squares criterion was selected for determining the parameters in the distributions which best fit the data. The 2-parameter and 3-parameter log-normal, Weibull and gamma distributions were tested. The 3-parameter log-normal was found to be superior to the 2-parameter log-normal by least squares approximation to the observed data. However, the 2-parameter was emphasized as a useful distribution both for describing data and for establishing an understanding of air pollutant statistics. They gave two major factors, wind speed and mixing heights, as influencing factors for air pollutant frequency distributions. With the easy use of the

2-parameter log-normal distribution it was shown that the mean and variance could easily be determined from a log probability plot of the data even though the Weibull and gamma distributions represented the data better than the log-normal. They presented an important relation between the standard geometric deviations and two averaging times, T_a and T_b, over a total period of time, T, over which data are available (usually a year) by

$$\sigma_b = \sigma_a^v \quad ; \quad v = \left[\frac{\ln(\frac{T}{T_b})}{\ln(\frac{T}{T_a})} \right]^{1/2} \tag{11}$$

This relationship was explained as a consequence of the near log-normality for short averaging times.

Berger, Melice and Demuth (1982) used the data of 24 hr averaged SO_2 concentrations recorded between January 1977 and March 1979 at 12 stations in Belgium. They were specifically interested in finding the distribution function which fits the extreme values and pointed out that the log-normal distribution did not represent the extreme values. The log-normal distribution overestimated the extreme observations while the 2-parameter exponential distribution fitted very well. The probability distribution of extreme values being of initially exponential type, the series of exceedances which are distinguished as large values selected over the 95th percentile by assuming the set of independent

values are well fitted by the 2-parameter (μ,s) exponential distribution which is of the form

$$F(x)=1-\exp[-(x-\mu)/s] \qquad (12)$$

where

$-\infty<x<\infty$ $\qquad \mu=\emptyset-s\cdot\ln(k)$

ϕ: location parameter

s: scale parameter

k: average number of exceedences

Since the gamma distribution is a generalization of the 2-parameter exponential law, they used it to represent the whole sample of concentration data in the form

$$f(x) = \frac{1}{\beta^{\alpha}\,\Gamma(\alpha)}\, x^{\alpha-1}\, e^{-\left(\frac{x}{\beta}\right)} \qquad (13.1)$$

and

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx \qquad (13.2)$$

which is the gamma function of α.

α : shape parameter

β : scale parameter

The gamma distribution provided a good fit for all stations and was accepted by a test based on the median and the largest value at the 0.2 significance level at 6 stations, at 0.05 significance level at 2 stations, at 0.01 significance level at 2 other stations while the fit was rejected at 2 stations.

Wilson and Simms (1985) tried to represent a reliable and easy-to-use prediction method for estimating ground level mean and peak concentrations. They tested the exponential, the log-normal and the gamma distribution functions to represent the actual distribution of concentration based on the data sets of the atmosphere and in a wind tunnel. They found that the gamma distribution function for the fraction of time of exceedence was a better choice than the log-normal or exponential distributions since they saw that both the log-normal and the exponential distributions were less accurate at low concentrations than the gamma distribution. According to their results, an initially gaussian distribution function is rapidly skewed with increasing time until meandering produces an exponential frequency of occurrence. The exponential distribution function is another form of the gamma distribution function when $\alpha=1$ in equation (13.1) and (13.2). Here, the word "conditional" is used to explain the data set which includes only non-zero concentration values.

Fackrell and Robins (1982a) made a series of experiments on concentration measurements in an open circuit wind tunnel by using the flame ionization detector system. They emphasized the distinction between a ground-level and an elevated source in terms of scale arguments of plume and turbulence. For a ground level source and rather far downstream on the center line near to the ground, the

distribution was found to be approximately gaussian. With increasing height the distribution took the approximate form of an exponential distribution

$$P(c)=A \exp(-Bc) \qquad (14.1)$$

At small downstream distances with increasing height the exponential form was seen, but a power law

$$P(c)=\alpha\ c(-\beta) \qquad (14.2)$$

was a better fit to the observed distribution. For ground level sources near the source the distribution was approximately gaussian at low heights and exponential at higher elevations. Farther downstream a gaussian form was most likely near the ground.

3. THE METHOD OF EXPERIMENTS AND THE DESCRIPTION OF THE BOUNDARY-LAYER TANK

The Deardorff-Willis boundary-layer tank (figure 1) has been extensively used in diffusion experiments to produce concentration data and in attempts to explain the nature of the planetary boundary layer in terms of its variables such as mixed layer height. There have been many studies on diffusion performed by J.W.Deardorff and G.E.Willis since 1974, see Deardorff and Willis (1987) and references therein.

The convective planetary boundary-layer tank is 124 cm in width and 124 cm in length. A stack 3.1 cm in height was held 1.05 cm above the bottom centerline of the tank on a horizontal axis. The stack was connected to a pump by a narrow pipe (the axis) through which effluent could pass. Although a water-ethanol mixture, colored by a blue or yellow dye can be used as effluent, in our experiments a fluorescent dye was used. It is transparent unless illuminated by a blue-green (laser) light, in which case the dye fluoresces with a green-yellow color within the plane of the light.

The effluent is steadily pumped out from the stack while traversing the stack across the bottom of the tank in the x direction (see figure 1) along the centerline, at a constant speed U, in order to simulate a uniform mean wind in the boundary layer. It should be noted that this mean wind

field is rather like a real one which occurs in a sunny daytime boundary layer since the wind shear then occurs only in a very shallow layer adjacent to the ground. In both cases, the turbulence is mainly generated by buoyant convection and the condition of $h >> -L$ is met, L is the Monin-Obukhov length and h is the boundary layer height. This condition requires the dominance of the buoyant convection over wind shear, (Deardorff and Willis, 1985).

The tank is filled up to 40 cm with filtered, degassed water, initially at 20°C in the lower 20 cm and increasing to 32.5°C above. This is done to create a nearly constant stratification of 0.63°C/cm above 19 cm. This layer is analogous to an inversion layer in the atmosphere. Then the convection is started by applying a large heat source to the bottom of the tank. After a few minutes to allow for the convection to fully fill the lower layer of the tank, the effluent is released. A high intensity planar-spread beam of laser light with a width of 0.01 h is directed downwards from above the tank. In this way, a thin-sliced cross section of the plume can be illuminated by the laser beam. While the stack is being traversed, the front view photographs of the illuminated y-z plane at the center of the tank are taken at 5 seconds intervals by a camera centered at a height of 19 cm. The travel time or effluent release time in the model tank, t, is converted to a downstream space coordinate

by using Taylor's hypothesis, that is

$$t=x/U \qquad (1)$$

where x is the downstream coordinate and U the mean wind speed (in the model tank U is the stack-towing speed). The time t is made dimensionless (Deardorff and Willis, 1987) by the eddy-diffusion time scale. Thus equation (1) becomes

$$X = \left[\frac{w_*}{h}\right] t = \left[\frac{w_*}{U}\right]\left[\frac{x}{h}\right] \qquad (2)$$

where w_* is the convection velocity scale, h the mixed layer height and X the dimensionless downstream coordinate.

The front view photographs (transparencies) of the tank are analyzed by projecting the plume image on a screen which is equipped with a traversing densitometer. The distance between each step of densitometer in y and z coordinates is about 0.04 h. A circular aperture with a radius of 0.007 h on the densitometer collects the light-rays passing through the transparency negative in different intensities corresponding to different amount of concentrations. Thus the sampling volume has a cylindrical shape with a radius of 0.007 h and a length of 0.01 h. An equivalent sampling volume in the atmosphere would be a volume with 7 m radius and 10 m length for a typical mixed layer height of 1000 m. From another point of view, the instrument collects air samples in a volume having dimensions as stated above, which is fairly large for sampling concentrations of the atmosphere. While the photographs are taken, the stack is

translated along x, giving the concentration measurements at different distances away from the stack. The concentration, c, is made dimensionless by using the mixed layer height, h, and the mean wind, U, by

$$C = \frac{cUh^2}{S} \tag{3}$$

where S is the effluent source strength (mass emitted per unit time).

The stack buoyancy parameter F_* is defined as

$$F_* = \frac{w_s r_s^2 g(\rho_o - \rho_s)}{\rho_o w_*^2 Uh} \tag{4}$$

Here w_s is the plume exit velocity, r_s the radius of the stack, g the acceleration due to gravity, ρ_o ambient air density, ρ_s plume density. For values of F_* which exceed about 0.02, the plume buoyancy effect becomes noticeable. The stack radius was 1 mm for the F_*=0.03 case and 2 mm for the other three buoyancy cases. The values of F_* used in this study are 0.0 (non-buoyant), 0.03 (weakly buoyant), 0.11 (buoyant) and 0.26 (highly buoyant). Each incremental analysis volume will be named by index I along both sides of the y-direction from the plume centerline at the bottom of the tank and by index J in the z-direction of the boundary layer.

The index I is defined by making use of the lateral standard deviation

$$\sigma_y = \left[\frac{\int_{-\infty}^{\infty} c\, y^2\, dy}{\int_{-\infty}^{\infty} c\, dy} \right]^{1/2} \qquad (5)$$

where c is the effluent concentration. By doing so, we take the concentration values into $\pm 3\sigma_y$ limits with 6 equal segments in the y-direction on either side of the plume centerline, figure 1. For example, I=1 includes the area $0<y<+1/2\sigma_y$ to the right of the centerline and the area $-1/2\sigma_y<y<0$ to the left of the centerline viewing the boundary layer tank from the front. Similarly, the height of the boundary layer is divided into six equal segments of 0.2 h starting from the bottom of the boundary layer, which is the layer J=1, to 1.2 h . The last section of the boundary layer in the z-direction is considered to cover any particulates which undergo overshooting above the mixed layer height. Since the data which were obtained in the volumes I=1 and 2 and J=1 and 2 did not show a systematic difference, these data sets were averaged and used in the graphics. The data were transformed by applying the expression

$$C\ (\text{Transformed}) = N + \ln(C-N+1) \quad (\ C>N\ ) \qquad (6)$$

to the higher concentration values, where C is the dimensionless concentration; N was set to be between 10 and

20. The average intensity plus two standard deviations of noisy small concentrations, measured in regions where no effluent existed, was also removed from the whole concentration field. The data were recalibrated by multiplying the resulting C values by a constant for each y-z cross section in order to ensure an approximate conservation of plume mass, as explicitly explained in Deardorff and Willis (1987). The resulting concentration values are plotted on histograms relative to their frequency of occurrences by satisfying

$$\sum_{i=1}^{17} p_i(C_*) = 1 \qquad ; \qquad C_* = \ln C_p \tag{7}$$

Here, the threshold level is C=0.01 and C_p stands for the concentration values above this threshold concentration level. In order to be consistent with the expressions of Deardorff and Willis (1987) and Wilson (1982), C_p will be used instead of C to distinguish the non-zero and zero concentrations. Then distribution functions such as the gamma distribution were applied to the histograms to fit the experimental data.

4. THE GAMMA PROBABILITY DENSITY FUNCTION

By using the mean $(\overline{C_p})$

$$\overline{C_p} = \int_0^\infty C_p \, p(C_p) \, dC_p \tag{15}$$

the variance $(\overline{C_p'^2})$

$$\overline{C_p'^2} = \overline{C_p^2} - (\overline{C_p})^2 \quad ; \quad \overline{C_p^2} = \int_0^\infty C_p^2 \, p(C_p) \, dC_p \tag{16}$$

and the fact that the area under the probability density curve equals 1

$$\int_0^\infty p(C_p) dC_p = 1 \tag{17}$$

the gamma probability density function becomes

$$p(C_p) = \frac{(m+1)^{m+1}}{\overline{C_p} m \Gamma(m)} \left[\frac{C_p}{\overline{C_p}}\right]^m e^{\left[-(m+1)\frac{C_p}{\overline{C_p}}\right]} \tag{18}$$

where

$$m = \frac{(\overline{C_p})^2}{\overline{C_p'^2}} - 1 \tag{19}$$

and

$$\Gamma(m) = \int_0^\infty C_p^{(m-1)} \, e^{-C_p} \, dC_p \tag{20}$$

which is the gamma function of m.

This procedure, the presentation of concentration distribution with the gamma distribution, was earlier utilized by Wilson and Simms (1985). The curves are sketched on a logarithmic scale by making a conversion from "C_p versus $p(C_p)$" scales to "$\ln C_p$ versus $p(\ln C_p)$" scales by

$$p(\ln C_p) = \left[\frac{\Delta C_p}{C_p}\right] C_p p(C_p) \quad ; \quad \frac{\Delta C_p}{C_p} = 0.4605 \qquad (21)$$

Figures 2(a-g), 3(a-f), 4(a-e) and 5(a-f) show the gamma probability density curves for cases with nondimensional buoyancy parameter, F_*, of 0.0, 0.03, 0.116 and 0.26 at different X distances for I=1,2 and J=1,2 locations within a mixed-layer plume.

The gamma probability distribution shows a very good fit to the experimental data for the non-buoyant effluent case, figures 2(a-g). This fit will be confirmed later by chi-square and mean-square-error evaluations for the non-buoyant case. However, for the other buoyancy cases examined here, F_*=0.03, 0.116 and 0.26, the chi-square goodness of fit test and the mean-square-error give different values as explicitly explained in section 9. The reason for the good fit of the gamma distribution in the non-buoyant case might be explained by the fact that such histograms do not show sharp peaks as the plume is rapidly diluted and entrained into the clear air by small eddies near the ground shortly after it is released from the stack. At small X

(X<1) the plume maximum, however, is lifted to a height of Z<0.7 by the effect of the largest eddies and in our case by the initial vertical momentum of the effluent. Convection carries the upper portion of the plume up to the mixed layer height and a little overshooting takes place depending on the stability of the inversion layer. Since the plume is not buoyant, the environmental convection is dominant throughout the entire boundary layer except at small X, where the effects of the effluent's strong initial momentum were evident. The histograms, for this reason, don't have sharp peaks although the effluent, bounded in vertical mixing, evolves towards a uniform (though decreasing) value of concentration near y=0 as X increases.

As the concentration uniformly decreases through dilution with increasing X, both tails of the distribution shorten. The gamma distribution underestimates the peak values, but not much, while it almost perfectly fits the low and high concentration values. Another apparent feature of the gamma distribution is that it produces the peak value at a reasonable value of concentration. For example, the peak value is around the concentration value of 1 ($\ln C_p=0$) in the 0.0 and 0.03 buoyancy cases, figures 2(a-g) and 3(a-f). The peak values of the experimental data, on the other hand, show a shift back and forth from the suggested peak of the gamma distribution without showing a very systematic variation with distance, figures 4(a-e) and 5(a-f). For the weakly buoyant

case, F_*=0.03, the peaks get more distinguishable and sharper with increasing X. At the same time the tails get shorter by having small frequency values at the lower and higher concentration values. In this case, the gamma distribution shows a great amount of underestimation for the peak values. On the other hand, it fits the lower and higher concentrations rather well, especially the lower ones.

In all cases made up of a large numbers of samples, the gamma distribution using the combinations of I's and J's showed a smooth increase before and after the peak although it consistently underestimated the peaks. Wilson, by using the wind tunnel data of Fackrell-Robins (1982), tested the gamma distribution and showed that it did not fit the peak values. He suggested this to be a scaling error. However, there is also an unexplained shift in the peak. Since his gamma distribution in concentration doesn't well represent the peak and the peak value can be as high as the mean, or even higher, it is not a good estimate for much of the concentration values even though it is a surprisingly good fit to the low concentrations. Section 9 discusses the fit of the gamma distribution by using some statistical significance tests such as the chi-square goodness of fit test. In that section, the distribution functions examined in this study are compared with the experimental data on the basis of the results of the test statistics.

5. THE EXPONENTIAL PROBABILITY DENSITY FUNCTION

When concentration values are plotted on a " C_p versus $p(C_p)$ " diagram rather than " $\ln C_p$ versus $p(\ln C_p)$ ", histograms suggest an exponential distribution. It is another form of the gamma distribution, equation (18), when $(\overline{C_p})^2 = \overline{C_p'^2}$ and hence m=0 :

$$p(C_p) = \frac{1}{\overline{C_p}} \exp(-\frac{C_p}{\overline{C_p}}) \qquad (22)$$

The exponential distribution is expected to deal with intermittent cases more adequately and is also shown to be most likely in cases where the plume more or less randomly passes over a monitor, for example due to meandering (Hanna, 1985). These conditions occur especially at small X where there is less time for the eddies due to convection to dilute the plume rapidly into the mixed layer. Thus the plume can show meandering, and small intermittency may be most likely compared to larger distances where the convective turbulence overcomes the plume buoyancy effect. Figures 6(a-d) show the exponential distribution at small X and figures 7(a-d) show it at large X for all four buoyancy cases. Inside windows in each figure are drawn to allow a better comparison between expected probability and experimental probability using the measured C_p at low frequencies. The probability values are plotted on the graphs by using the conversion formula

$$p(C_p)_{Rev} = \frac{p\left[C_p(n)\right]}{C_p(n+1) - C_p(n)} \quad ; \quad n=1,2,\ldots,17 \qquad (23)$$

As is seen, the exponential distribution function estimates the high concentration values well but not the low ones. In fact the exponential distribution lies well below the experimental data at low concentrations. This feature states that the intensity of the plume

$$i_p^2 = \frac{\overline{C_p'^2}}{(\overline{C_p})^2} \qquad (24)$$

is projected to be lower according to equation (22) while the experimental data suggest the opposite at low concentrations. In some situations, we may want to know the probability of being exposed to some gases or toxic materials without crossing some low concentration threshold level. The exponential distribution function is weak in answering this problem.

6. LOG-NORMAL PROBABILITY DENSITY FUNCTION

The log-normal is the first distribution which was used to fit the observed data. It is in the form of

$$p(\ln C_p) = \frac{1}{C_p \sqrt{2\pi}\ \sigma_{\ln C_p}} \exp\left[- \frac{(\ln C_p - \overline{\ln C_p})^2}{2\ \sigma^2_{\ln C_p}} \right] \qquad (25)$$

where $\ln C_p$ is the mean of the log-concentrations and $\sigma_{\ln Cp}$ is the log-standard deviation.

In the earlier studies already reviewed, the log-normal was found to be the most accurate among some other functions such as gamma or beta in wind tunnel experiments. The observation that concentration distributions for all averaging times are approximately log-normal can be explained if the short averaging time data are themselves assumed to be log-normally distributed, Bencala and Seinfeld (1976). If the log-normal distribution applies to the concentration distribution of pollutants in the atmosphere, in general, we might expect to see a probability density curve having equal (or at least tending to have equal) areas under its two tails with X for all effluent buoyancy cases on the measured frequencies of pollutants obtained by using the model tank. However, the probability density curves are skewed for all four different buoyancy cases at every X examined here, figures 8(a-d). Therefore, a simple log-normal distribution will not be considered a suitable function for the

concentration distribution of pollutants in the " unstable" planetary boundary layer and for short sampling times (5 - 10 min).

7. THE DOUBLE-LOG-NORMAL PROBABILITY DENSITY FUNCTION

In the light of the studies which have been discussed on the air pollution concentration distribution, it is clearly seen that none of the distribution functions fit sufficiently well to the experimental or observed data, except for the gamma distribution for $F_*=0$. The log-normal distribution is not the one which can be used to represent the concentration distribution since the distribution doesn't appear to follow log-normal distribution as examined in the previous section. The gamma cannot determine the average and peak concentrations accurately. Lynn (1974) also tested the Pearson curves which are different combinations of skewness and kurtosis. That is, for different values of the skewness and the kurtosis, the probability density curves take different shapes. Lynn couldn't find a distribution function among the functions he tested to fit the overall data.

Deardorff suggested using the log-normal in a modified form and giving it a slightly different name: " double-log-normal ". This function applies the normal density or Gaussian curve separately to each side of the distribution curve and taking the peak value as origin, hence ensuring that the most probable concentration or mode will be well estimated.

Again, the area under the probability density curve must be 1:

$$\sum_{i=1}^{17} p_i(\ln C_p) = 1 \tag{26}$$

By multiplying both sides of equation (26) with $\Delta C_p/C_p$, we obtain

$$\sum_{i=1}^{17} p_i(\ln C_p) \frac{\Delta C_p}{C_p} = \frac{\Delta C_p}{C_p} \tag{27}$$

When equation (27) is written in continuous form, then

$$\int_{-\infty}^{\infty} p(\ln C_p)\, d(\ln C_p) = \Delta(\ln C_p) = \Delta C_* \tag{28}$$

or

$$\frac{1}{\Delta C_*} \int_{-\infty}^{\infty} p(C_*)\, dC_* = 1 \tag{29}$$

where $\frac{\Delta C_p}{C_p}$ is in our case 0.4605, which is the interval width of the two subsequent log-concentrations. Here the function $p(C_*)$ is the double-log-normal probability density function.

The double-log-normal probability density function is defined as

$$P(C_*) = \begin{cases} P(C_{*_m}) \exp\left[- \dfrac{(C_* - C_{*_m})^2}{2\sigma^2_{Left}} \right] & \text{when } C_* \le C_{*_m} \\ P(C_{*_m}) \exp\left[- \dfrac{(C_* - C_{*_m})^2}{2\sigma^2_{Right}} \right] & \text{when } C_* > C_{*_m} \end{cases} \tag{30}$$

where $C_*=\ln C_p$ is the log-concentration, C_{*_m} is the value of the log-concentration at which the mode occurs. $P(C_{*_m})$ is the frequency of the mode and σ_{Left} and σ_{Right} are the standard deviations for the left and right side of the curve, respectively.

The standard deviations are found by evaluating the normal distribution function for each left and right area from the mode under the density curve. The normal distribution function is

$$\frac{1}{\sigma_L\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(t-m)^2}{2\sigma^2}} dt = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x-m}{\sigma\sqrt{2}}\right) \right] \tag{31}$$

with mean m and variance σ.

To perform the integration let the variables in equation (31) be as below:

$$\left.\begin{aligned} x = m = (\ln C_p)_m = C_{*_m} \\ \sigma = \sigma_{Left} \quad (\text{ and } \sigma = \sigma_{Right} \text{ in case of } \int_x^{\infty} \text{ }) \\ t = \ln C_p = C_* \end{aligned}\right\} \tag{32}$$

By making necessary notational changes with the equalities (32) in equation (31) for the left side of the curve, we find that

$$\frac{1}{\sigma_L\sqrt{2\pi}} \int_{-\infty}^{C_{*_m}} e^{-\frac{(C_* - C_{*_m})^2}{2\sigma_L^2}} dC_* = \frac{1}{2}\left[1 + \text{erf}\left(\frac{C_{*_m} - C_{*_m}}{\sigma_L\sqrt{2}}\right)\right] \tag{33}$$

where

$$\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt \tag{34}$$

which is the error function of z:

$$z = \frac{x - m}{\sigma_L\sqrt{2}} \tag{35}$$

The multiplication of both sides of equation (33) with

$$\frac{P(C_{*_m})\sigma_L\sqrt{2\pi}}{\Delta C_*} \tag{36}$$

gives

$$\frac{P(C_*)_m}{\Delta C_*} \int_{-\infty}^{C_{*m}} e^{-\frac{(C_* - C_{*m})^2}{2\sigma_L^2}} dC_* = \frac{P(C_*)_m \sigma_L \sqrt{2\pi}}{2\Delta C_*} \qquad (37)$$

However, the left hand side of equation (37) is Cs_{Left} since the error function is zero in our case due to $x=m=C_{*m}$. By evaluating the indefinite integral, we obtain the standard deviation for the left side of the density curve as

$$\sigma_L = 0.3674 \left[\frac{Cs_L}{P(C_*)_m} \right] \qquad (38)$$

Similarly, we find the standard deviation for the right side of the density curve:

$$\sigma_R = 0.3674 \left[\frac{Cs_R}{P(C_*)_m} \right] \qquad (39)$$

where, Cs_{Left} and Cs_{Right} are the areas under the probability density curve for the left and right side of the density curve from the peak, respectively. Figures 2, 3, 4 and 5 show the plots of the double-log-normal probability density curves for F_*=0, 0.03, 0.11 and 0.26, respectively. The mean can be calculated as

$$\overline{C_p} = \sum_{i=1}^{17} (C_*)_i \; p_i(C_*) \qquad (40)$$

or by using the mean which involves the parameters Cs_{Left} and Cs_{Right}

$$\left.\begin{aligned} \overline{C_p} &= Cs_L \left(\overline{C_L} - \overline{C_R} \right) + \overline{C_R} \\ \text{or} \\ \overline{C_p} &= Cs_R \left(\overline{C_R} - \overline{C_L} \right) + \overline{C_L} \end{aligned}\right\} \qquad (41)$$

where

$$\overline{C_L} = \int_{-\infty}^{C_{*_m}} C_* \, p(C_*) \, dC_*$$

and (42)

$$\overline{C_R} = \int_{C_{*_m}}^{\infty} C_* \, p(C_*) \, dC_*$$

are the means for the left and right side of the curve, respectively.

To obtain the probability values of pollutant concentration by using the double-log-normal probability density function, one needs to know the frequency of occurrence of the peak value (mode) and the area under the density curve corresponding to the sum of the frequencies less and greater than the peak for that particular x, buoyancy and the intermittency situation. Then the expected probability value of a particular concentration can be calculated by equation (30) or the expected probability value which is less than or equal to a particular concentration, say x, can be calculated by

$$F(C_* \leq x) = \int_{-\infty}^{x} p(C_*) \, dC_* \qquad (43)$$

The last information needed to obtain a complete form of the double-log-normal probability distribution function is the intermittency value which is examined in the next section.

8. THE INTERMITTENT NATURE OF THE PLUME

The data of experiments conducted in the boundary layer tank do not include concentration values which are smaller than C_p=0.01. The probability density functions examined here must be also true for intermittent plume cases. There is the fact that a receptor receives not only non-zero values of pollutant concentration, it also receives zero values or undetectably small values which are assigned as zero values because the measurement system cannot tell the difference. In order to determine the true probabilistic nature of the concentration distribution, the intermittency factor must be included in the probability density functions.

Wilson (1982) gave a nice discussion of the intermittency factor and included it in the probability density functions he described. He distinguished the conditional mean ($\overline{C_p}$) and variance ($\overline{C_p'^2}$) as the values which are calculated by using only non-zero values of concentration. By knowing the fact that the fluctuation part C' including zero values is calculated with

$$C' = C - \overline{C} \tag{44}$$

and C' excluding zero values is calculated with

$$C_p' = C_p - \overline{C_p} \tag{45}$$

where C represents the actual concentration field including zero values and C_p excluding zero values.

When the plume is carried to a receptor for a fraction of the time, γ, the mean is

$$\overline{C} = \gamma \overline{C_p} + (1-\gamma) \cdot 0 \qquad (46)$$

and the variance

$$\overline{C'^2} = \gamma \, (C_p - \overline{C_p})^2 + (1-\gamma) \, (-\overline{C})^2 \qquad (47)$$

From these

$$\overline{C_p} = \frac{\overline{C}}{\gamma} \quad ; \quad \overline{C_p'^2} = \frac{1}{\gamma} \left[\frac{1-\gamma}{\gamma} (\overline{C})^2 - \overline{C'^2} \right] \qquad (48)$$

where $\overline{C_p}$ and $\overline{C_p'^2}$ are the conditional mean and variance calculated by using only non-zero concentration values. $\overline{C_p}$ and $\overline{C_p'^2}$ must be replaced by their equivalent values in the gamma probability density function, equation (18), to fully represent the distributive nature of pollutants. In equation (30) we used σ^2_{Left} and σ^2_{Right} rather than $\overline{C_p'^2}$, so that these two variances must be replaced with their equivalences or the computation of the frequencies falling into each concentration category must be recalculated by considering both zero and non-zero concentrations.

The plume is present over a receptor with non-zero concentration for a fraction of the time, γ, and the concentration is zero for $1-\gamma$ time. For the time at which zero concentration occurs, the probability density fuction is a Dirac delta function (Wilson, 1982). For the time at which

non-zero concentrations are observed, the probability density function is in our case a double-log-normal density function. By combining these two probabilities

$$p(C_*)_{Rev} = (1-\gamma)\,\delta(C_*) + \gamma\, p(C_*) \qquad (49)$$

where the delta function has the property

$$\int_{-\infty}^{\infty} \delta(C_*)\; dC_* = 1 \quad . \qquad (50)$$

Equation (43) also needs to be replaced with the one that includes the intermittency. By redefining the cumulative distribution function, equation (43) is rewritten as

$$F(C_* \leq x)_{Rev} = (1-\gamma) + \gamma \int_{-\infty}^{\infty} p(C_*)\; dC_* \qquad (51)$$

Then we can calculate the probability of observing a concentration less than a particular concentration by accounting for the intermittency factor. We also need to know the calculation of the intermittency, but our goal is not to predict this factor here (see Wilson, 1982).

9. RESULTS AND DISCUSSION

In this section, the fit of the distribution functions reviewed here are examined by some statistical tests such as the classical chi-square goodness of fit test. Some general characteristics of the parameters of the distribution such as its mean and standard deviation versus the distance from the stack for each of the four buoyancy cases are also given. At the same time, the distribution of the parameters of the double-log-normal distribution such as its two standard deviations versus the downwind distance are shown in figures 8-10, to help the understanding of the behavior of this distribution along with the observed distribution before presenting a general discussion of the fit for each distribution function examined here.

In non-buoyant and weakly buoyant cases the mode shows a tendency towards lower concentration as downwind distance X increases, figures 8(a-d). However, this tendency is not very rapid. If the distribution were log-normal, the skewness,

$$\text{Skewness} = \frac{\mu_3}{\sigma^3} \tag{52}$$

where μ_3 is the third moment about the mean and σ^3 is the cube of the standard deviation, would tend to zero. In fact, it moves away from zero and the distribution tends to be more

negatively skewed even though this departure from zero is not big. The magnitude of the skewness changes between 0 and -1 for F_*=0 and 0.116 while this change is seen between -0.5 and -2 for F_*=0.03 and between 0.5 and -0.1 for F_*=0.26 buoyancy cases. As X increases, kurtosis which is in the form of

$$\text{Kurtosis} = \frac{\mu_4}{\sigma^4} - 3 \tag{53}$$

where μ_4 is the fourth moment and σ^4 is the square of the variance, increases as well. This shows that the the plume is getting more dilute with time as is expected for a plume in the mixed layer due to the convective turbulence. From the double-log-normal probability density curves of the concentration, figures 2-5, we observe a piling up of relative frequencies around the mode while the tails of the distribution get shorter. Thus the distribution acquires a taller peak at the mode by decreasing its amount of flatness and causes an increasing kurtosis. The increase in the kurtosis is observed especially for the weakly buoyant case, F_*=0.03, as larger (from -0.5 to 4.5) than those in the other buoyancy cases (from -1 to 1 for F_*=0, from -1 to -0.1 for F_*=0.116 and from -1 to -0.2 for F_*=0.26 buoyancy cases). As a result, the distribution gets more negatively skewed while its kurtosis gains more positiveness with X.

When we look at the mode of the distribution and its frequency values, figures 9(a-d), for non-buoyant and weakly buoyant cases, the distribution seems to be tending to a

gaussian distribution by pulling its mode towards lower concentrations. However, the skewness preserves its negativeness for all buoyancy cases. The increases in the peak relative to X for the F_*=0 and F_*=0.03 buoyancy cases are bigger than those in buoyant cases since the plumes with F_*=0 and 0.03 entrain faster, so that they get homogenized faster. The buoyant plumes can carry themselves intact to greater distances from the stack due to their higher buoyancies without getting significantly entrained into the clear air by large eddies. Therefore, the peak values are probably not built up faster in buoyant plumes than those in the less buoyant and non-buoyant plumes. The peak values are almost constant for the 0.116 and 0.26 buoyancy cases. The mode increases with X for two buoyant cases while it decreases for weakly buoyant and non-buoyant cases. In general, the plume is diluted towards a particular, well mixed concentration value. If there were a particular concentration value, the probability density curve would have a sharp peak around that concentration without left and right tails.

The left and right standard deviations of the double-log normal, figures 10(a-d), show their expected features of decrement with increasing downwind distance X. The left standard deviation decreases with X more rapidly than the right standard deviation does for all buoyancy cases except for the most buoyant case, F_*=0.26. This situation

refers to the fact that in approaching a particular mixing concentration that value dominates over the concentrations which are present in an air parcel with smaller frequencies compared to the other concentrations with higher frequency of occurrences. This approach eventually causes the concentrations to have smaller frequency of occurrences than their initial frequencies before the convection starts.

This is probably the cause of a faster decrease in the magnitude of the left standard deviation with X. The measured standard deviation shows a decrease with X for F_*=0, 0.03 and 0.116 buoyancy cases while it is almost constant with X for F_*=0.26, figures 11(a-d). This is a result of reaching a well-mixed concentration value and an appearance of a tall-peak following shorter tails in the probability density curve with X, figures 2-5. The measured mean of the concentration distribution does not change much with X for all buoyancies even though it shows a slight increase between C_*=-3 and -2 for F_*=0.26. The mean has an average value of C_*=-0.75 for the F_*=0 and F_*=0.03 cases and an average value of C_*=-1.5 for the F_*=0.116 case.

In order to see how well the functions examined here fit the measured probability distribution of concentration, we used the chi-square goodness-of-fit test which is

$$\chi^2 = \frac{\sum_{i=1}^{17} \left[N \cdot f_i(C_*) - N \cdot p_i(C_*) \right]^2}{N \cdot p_i(C_*)} \qquad (54)$$

where $p_i(C_*)$ is the calculated or expected frequency, $f_i(C_*)$ is the measured or observed frequency and N is the total number of points in the sample. The non-parametric Kolmogorov-Smirnov test was also applied to the measured frequencies of concentration, which is in the form of

$$D = \max \mid F_i(C_*) - P_i(C_*) \mid \quad \text{for all (i)s} \qquad (55)$$

where D is the maximum of the differences between the calculated frequencies, $P_i(C_*)$, and the measured or observed frequencies, $F_i(C_*)$, in a cumulative sense. The D test statistics gives very big values for the double-log-normal and the gamma distribution since it considers the frequencies in a cumulative sense rather than considering the cell frequencies individually. This causes a piling up of departures while the functions may be seen as good fits to the observed data individually at small or high concentrations. Therefore, this test will not be considered here as a good test criteria for the fit.

The valid chi-square goodness of fit test requires the sample size for each class interval be at least 5. If this requirement is not met in each interval, the chi-square test may give unrealistic values since this test is sensitive to the denominator in equation (54). In order to meet this condition, the interval which has a sample size of less than 5 is combined with the adjacent interval or intervals until this condition is satisfied. By following this procedure, the chi-square values for each of the buoyancy cases (F_*=0,

0.03, 0.116 and 0.26, tables 5 thru 8 respectively) at all X distances were calculated.

From the results of the chi-square test, the gamma distribution is a good fit to the experimental data at a 1 percent significance level in F_*=0, 0.26 at all downwind distances and in the F_*=0.116 buoyancy case at all X except at X=2.87. The degrees of freedom for the gamma distribution are 14 since we lose two degrees of freedom in estimating the mean and the variance of the distribution. The double-log-normal distribution, on the other hand, is a good fit to the observed distribution at the 1 percent significance level at X=1.02, 1.46 and 2.413 for F_*=0, at only X=1.14 for F_*=0.116 and at X=1.19, 1.62 and 2.26 in the F_*=0.26 highly buoyant case. The degrees of freedom for the double-log-normal are 13 since we lose three degrees of freedom in estimating the two variances and in finding the concentration where the mode occurs. The chi-square values at the 1 percent significance level are 27.69 with 13 degrees of freedom and 29.14 with 14 degrees of freedom. The chi-square values in the F_*=0.03 case for the gamma and the double-log-normal distributions were unrealistically big. This is probably caused by a sampling error in the weakly buoyant case.

An important point in calculating the chi-square values is that towards the end of the distribution some intervals are not taken into account since there will be no more

intervals to add in completing the requirement that the sample size in each interval be 5 or more. Therefore, the chi-square values are incomplete in giving the exact values to test the null hypothesis that the fit of the distribution function tested does not significantly differ from the observed distribution function. However, the remainder of the chi-square value will not be big since the cell frequencies get smaller towards the end of the distribution at higher concentrations and so does their square. Therefore, the contribution of the remainder to the chi-square value will not be big, so the resulting value of the chi-square test will be assumed within acceptable limits at the 1 percent significance level.

Another very important point is that the summation of the adjacent class intervals causes a big chi-square value due to the squaring and the division process in equation (54). When we examine the fit of the gamma and the double-log-normal distributions, for example in figure 5(f), the gamma density function suggests a highly skewed distribution from the observed distribution of experimental data. Even though the gamma distribution is a very good fit to the small and high concentrations, overall it is not a good fit compared to the double-log-normal distribution because the gamma density function estimates the mean at a different concentration value at the $F_{\star}$=0.116 and 0.26 cases (see figures 4 and 5). On the other hand, the chi-square

value shows a big difference in favor of the gamma distribution when it is compared for both the gamma and the double-log-normal distributions, as in the example of the F_*=0.26 highly buoyant case at X=4.96. The reason for the big chi-square values is that the double-log-normal distribution underestimates the frequencies at small concentrations. Thus, when the class intervals at small concentrations are added to meet the criteria for the expected frequency, $Np_i(C_*)$ term, the observed frequency, $Nf_i(C_*)$ term, gets bigger and so does the square of the numerator as a result of division by the numerator of the equation (54). This numerical result gets large as the computing continues but with smaller incremental amounts.

Similar computing results are seen in all buoyancy cases examined in this study. We cannot designate the double-log-normal distribution function as a bad estimator for the observed concentration distribution by looking at only the chi-square test statistics since these results are heavily dependent on the poor fit at small concentrations. On the other hand, we cannot name the gamma distribution as a good estimator to the observed distribution when it does not give good estimates of the cell frequencies around the peak concentration.

The sum of absolute differences and the mean square of error were also calculated to see the goodness of fit of the gamma and the double-log-normal distributions. There are no

tables similar to the one for the chi-square test to compare the fit of the functions to the observed distribution for the sum of absolute differences and the mean square of error. Therefore, the sum of squares of error or the sum of absolute differences can give only an idea about the discrepancies between the expected distribution of concentration according to the density function tested and the observed distribution of concentration. However, the mean square of error and the sum of absolute differences are still among the acceptable test statistics in comparing the distribution suggested by a function such as the gamma density function with the observed distribution, Lynn (1974).

The sum of the absolute differences, tables 1-4, is in the form of

$$D = \sum_{i=1}^{17} | f_i(C_*) - p_i(C_*) | \qquad (56)$$

where $p_i(C_*)$ is the frequency value calculated by using the gamma and the double-log-normal distributions, and $f_i(C_*)$ is the measured frequency. The negative and positive differences were also posted in the tables in order to see if there was a significant discrepancy in estimation of the measured frequencies through the gamma and double-log-normal distributions in terms of underestimation or overestimation of the probability density curves. The total sum of squares of error whose form is

$$\text{Error sum of squares} = \sum_{i=1}^{17} \left[f_i(C_*) - p_i(C_*) \right]^2 \qquad (57)$$

was also added to the tables.

The total absolute differences show the gamma distribution as a better fit than the double-log-normal for the non-buoyant case at all X distances (table 1), for the weakly buoyant case at all X except at X=4.28 (table 2), and for the buoyant case at X=3.57 and 4.15 (table 3). The double-log-normal distribution is seen as a better fit than the gamma distribution for the weakly buoyant case at only X=4.28 (table 2), for buoyant case at X=1.14, 2.05 and 2.87 (table 3), and for highly buoyant case at all X except X=2.26. In the buoyant case at X=3.03 both the distributions give the same total absolute difference value but the double-log-normal distribution should be observed as a better fit than the gamma distribution since the latter is highly skewed from the observed distribution of concentration. In F_*=0.03, the weakly buoyant case, we suspect a sampling error by interpreting the patterns of mean concentration which showed relatively minor differences between the contours of the mean concentration for F_*=0 and 0.03 buoyancy cases, Deardorff and Willis (1987).

When the positive and negative differences are examined in tables 1-4, the double-log-normal deviates positively from the experimental probability density by the same magnitude of the sum of the frequencies as it does negatively for all

buoyancy cases at all X. However, the negative differences are seen to be of less magnitude than the positive differences for the gamma distribution, especially for F_*=0.116 and 0.26 buoyancy cases. According to equation (56) negative differences mean an overestimation of the function tested over the observed distribution, while positive differences mean an underestimation of the observed distribution before the absolute operator is applied to that equation. Thus, the gamma density function mostly underestimates the observed distribution for F_*=0.116 and 0.26 at all downwind distances.

When we review the sum of squares of error, the double-log-normal distribution is seen to be a better fit than the gamma distribution for the two high buoyancies F_*=0.116 and 0.26 at all X except X=4.15 for F_*=0.116 and at X=2.26 for F_*=0.26 (tables 3 and 4). In the F_*=0 and F_*=0.03 cases the gamma distribution does fit the observed distribution of concentration better than the double-log-normal distribution except for F_*=0 at X=1.02. In the weakly buoyant case at X=4.28, the sum of absolute differences shows the double-log-normal distribution as a better fit than the gamma distribution, while the sum of square of error shows the latter as a better fit. This is a result of the squaring procedure in equation (57). In the highly buoyant case at X=2.26 equation (56) gives equal magnitude of absolute differences for both distributions

examined while equation (57) supports the fit of the gamma distribution.

10. CONCLUSIONS

By looking at all these results, we can say that the gamma distribution is a better fit for the observed concentration distribution than the double-log-normal distribution for the $F_*=0$ (non-buoyant) and $F_*=0.03$ (weakly buoyant) cases. On the other hand, the double-log-normal distribution is a much better fit than the gamma distribution for $F_*=0.116$ (buoyant) and $F_*=0.26$ (highly buoyant) cases. One point should be emphasized here: The gamma density function gives very good estimates of the frequencies of small concentrations for all buoyancy cases at almost all X distances from the stack. The double-log-normal, however, is seen as a better estimator function to the observed distribution of concentrations for high buoyancy cases which are likely the cases for environmental pollution due to the buoyant gases released from stacks. Another proviso here is that this study done with the Deardorff and Willis boundary layer model tank was conducted for a continuous point source at a small elevation above the surface in a convectively mixed layer.

11. SUGGESTED FUTURE STUDIES

In this study the concentration data of the laboratory model tank of Deardorff and Willis (1974) were used to determine whether the distribution of concentration of pollutants in the boundary layer could be represented by a distribution or density function. Then by using this function we can attempt to make predictions about the possible behavior of the pollutants in the atmosphere in terms of its occurrence for a time period such as a year. The gamma probability density function was shown to be a good estimator function for the distribution of concentration for non-buoyant and weakly buoyant cases. It was also shown that it was estimating the occurrences of low concentrations very well. The double-log-normal probability density function was shown to be a better fit to the distribution of concentrations than the gamma density function at buoyant and highly buoyant cases. These two functions should be compared by using the observed pollutant concentration data obtained in a convective atmospheric boundary layer in order to see if either of these two shows a better estimate of the occurrences of pollutants than the other. Then by looking at those results we may be able to make strong statements about how to best estimate the occurrences of concentrations in the actual atmosphere.

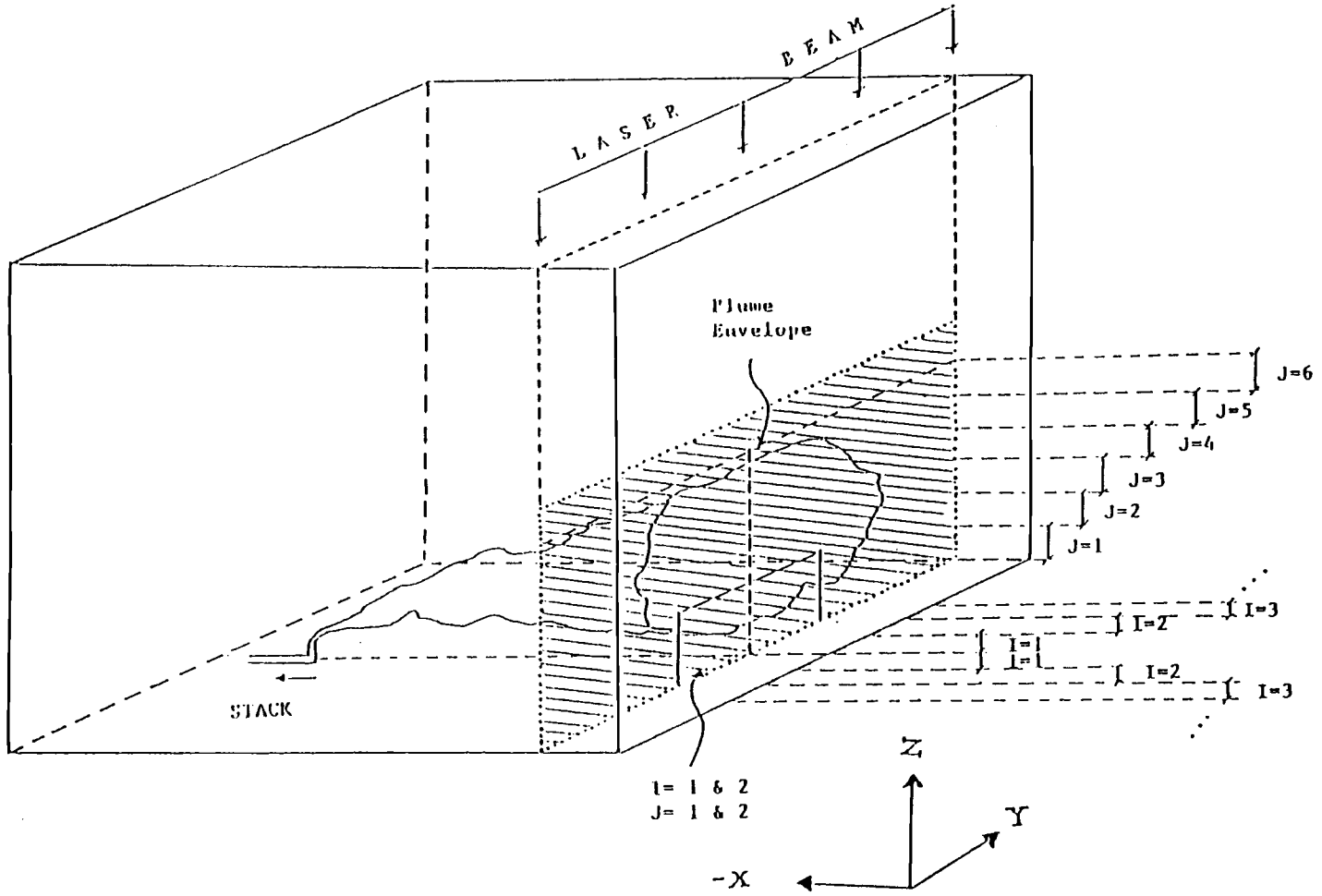

Figure 1 : Schematic diagram of the boundary layer model tank and its sections from centerline in y-direction with the use of its standard deviation, σ_y , and its sections in z-direction from 0 to 1.2 h, mixed layer height.

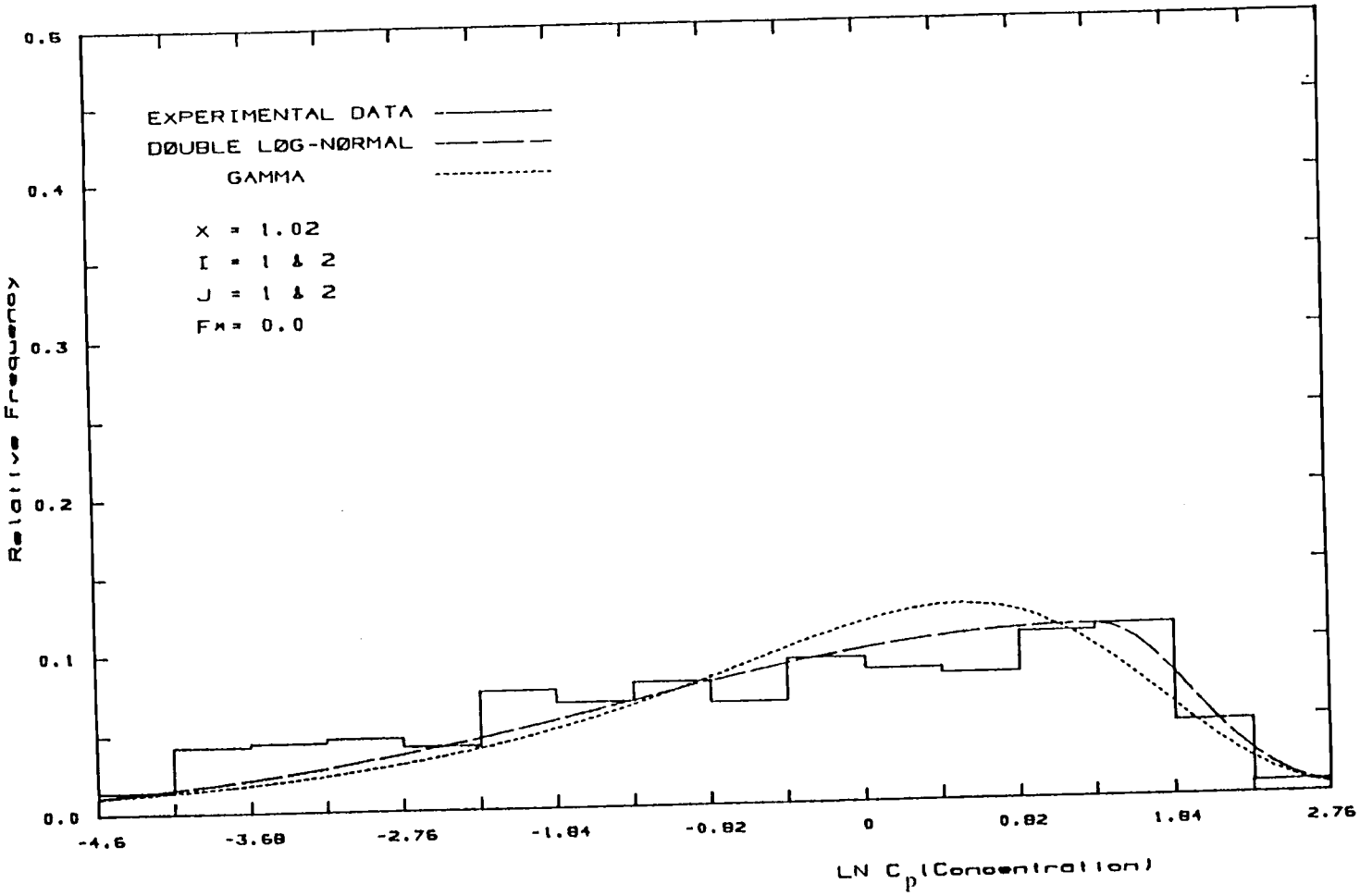

Figure 2 (a) : The probability of occurrences of log-concentration, $P_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y| \leq \sigma_y$ and $0<h<0.4$ at X=1.02 for $F_*=0$ case.

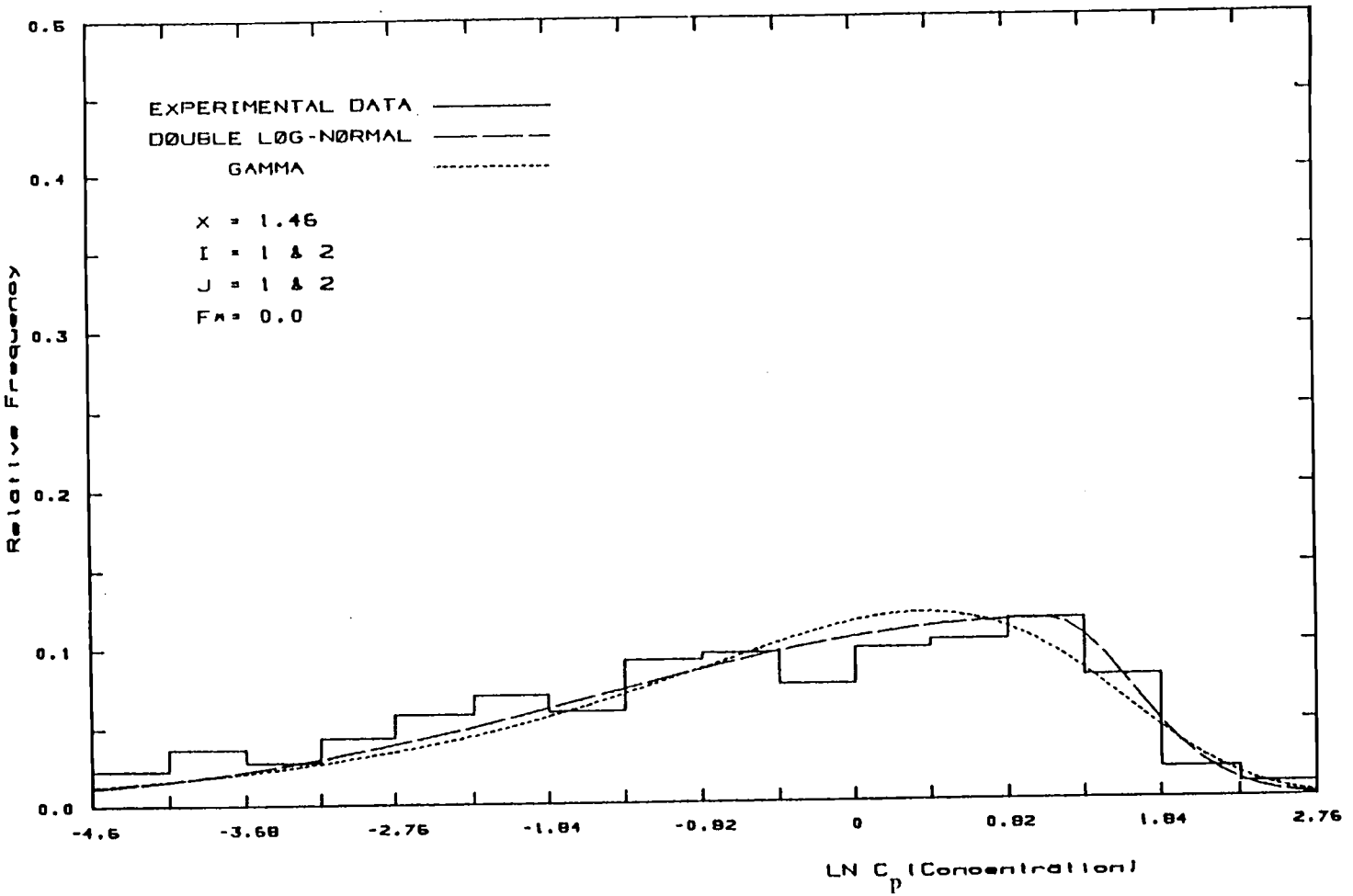

Figure 2 (b) : The probability of occurrences of log-concentration, $P_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at X=1.46 for F_x=0 case.

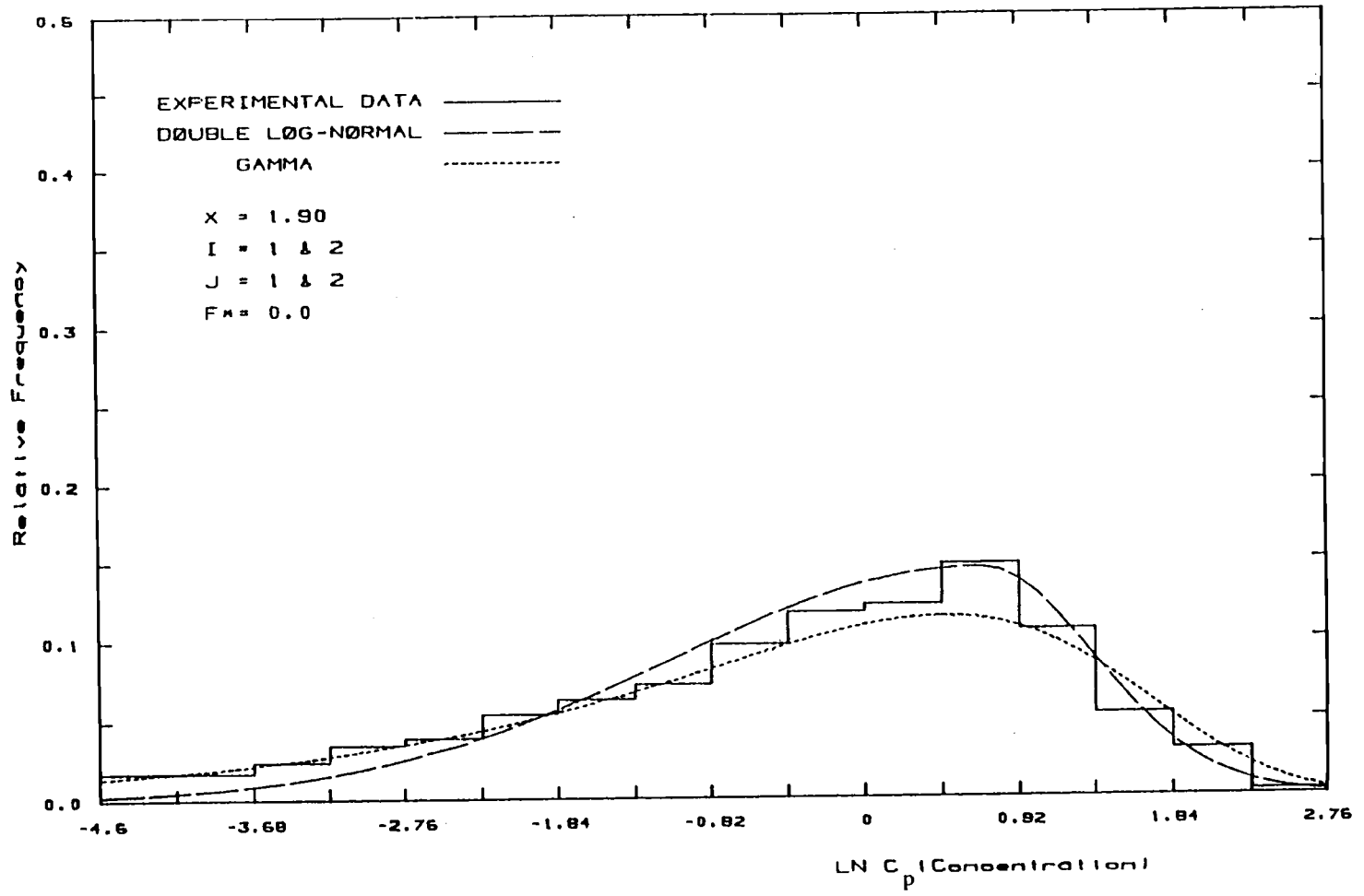

Figure 2 (c) : The probability of occurrences of log-concentration, $P_1(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $\bar{X}=1.90$ for $F_*=0$ case.

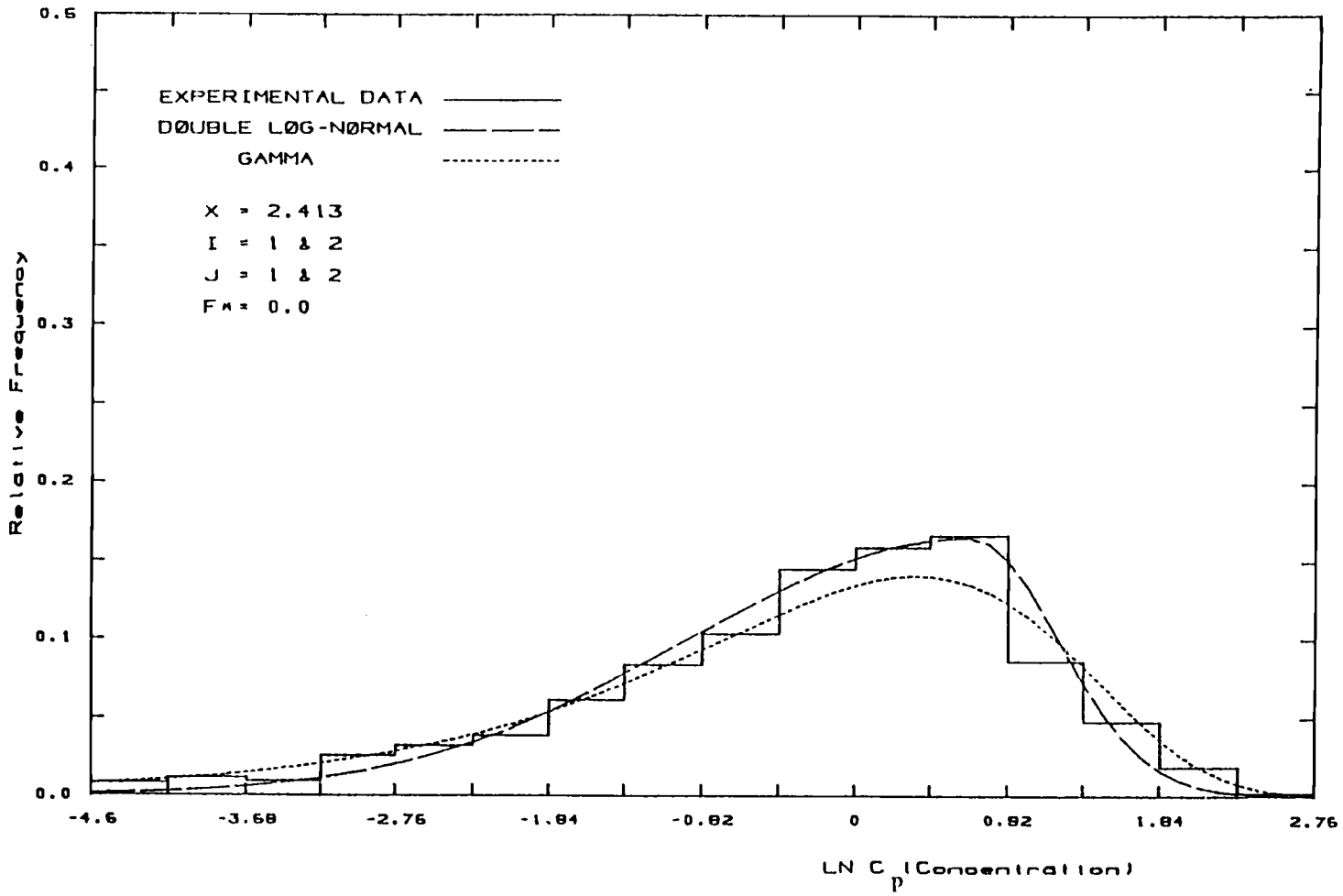

Figure 2 (d) : The probability of occurrences of log-concentration, $P_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=2.413$ for $F_*=0$ case.

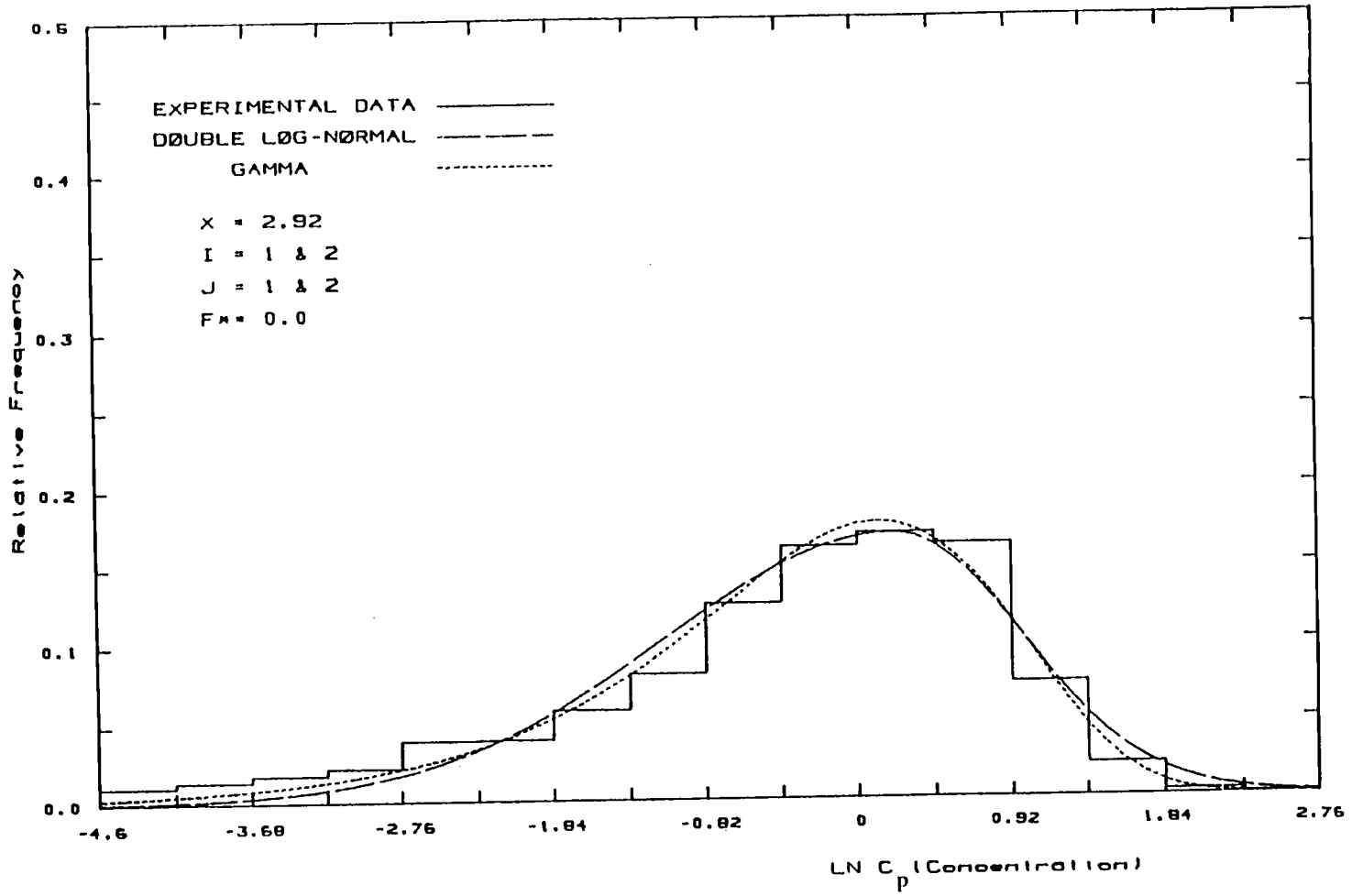

Figure 2 (e) : The probability of occurrences of log-concentration, $P_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=2.92$ for $F_*=0$ case.

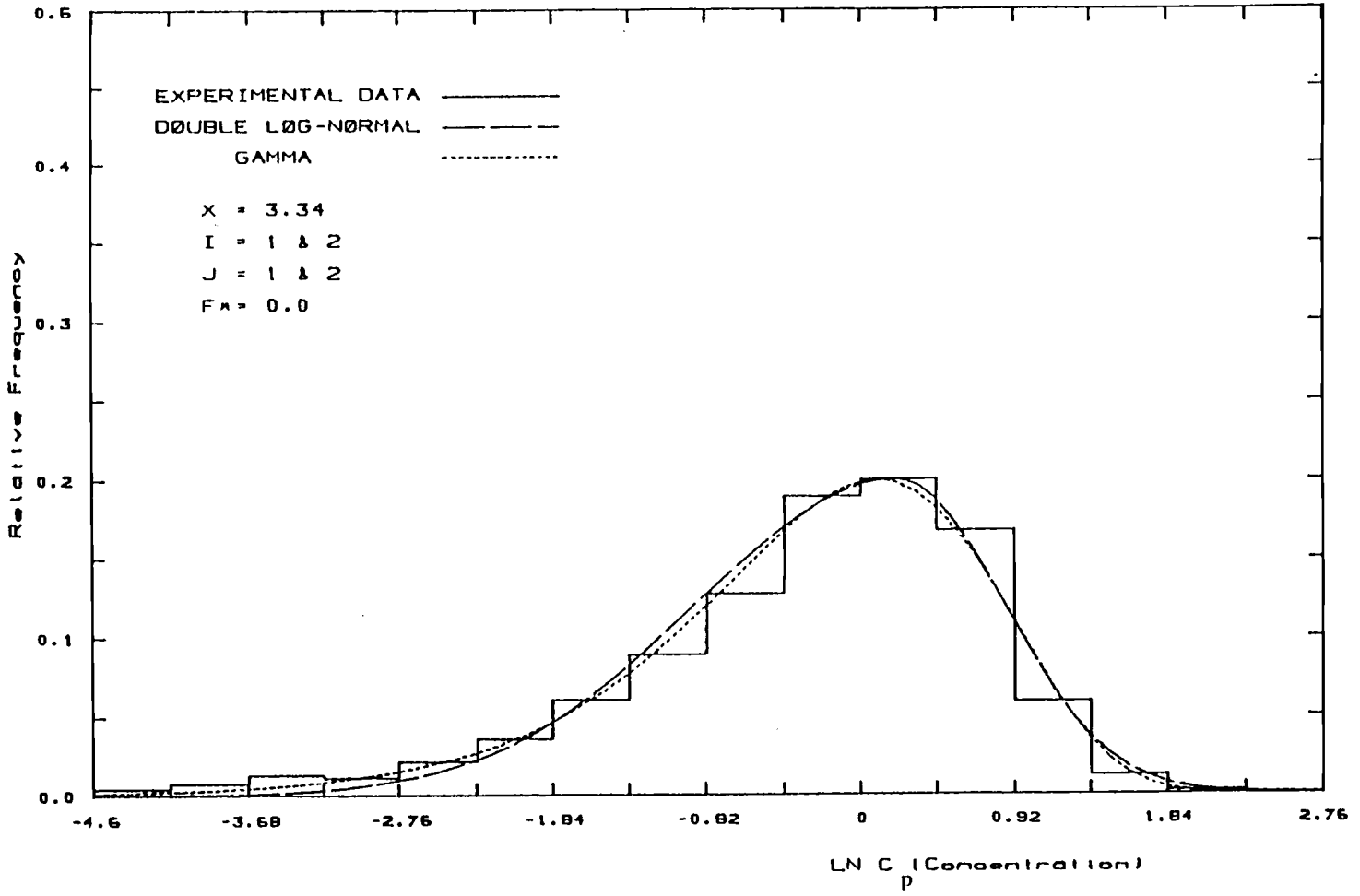

Figure 2 (f) : The probability of occurrences of log-concentration, $P_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=3.34$ for $F_x=0$ case.

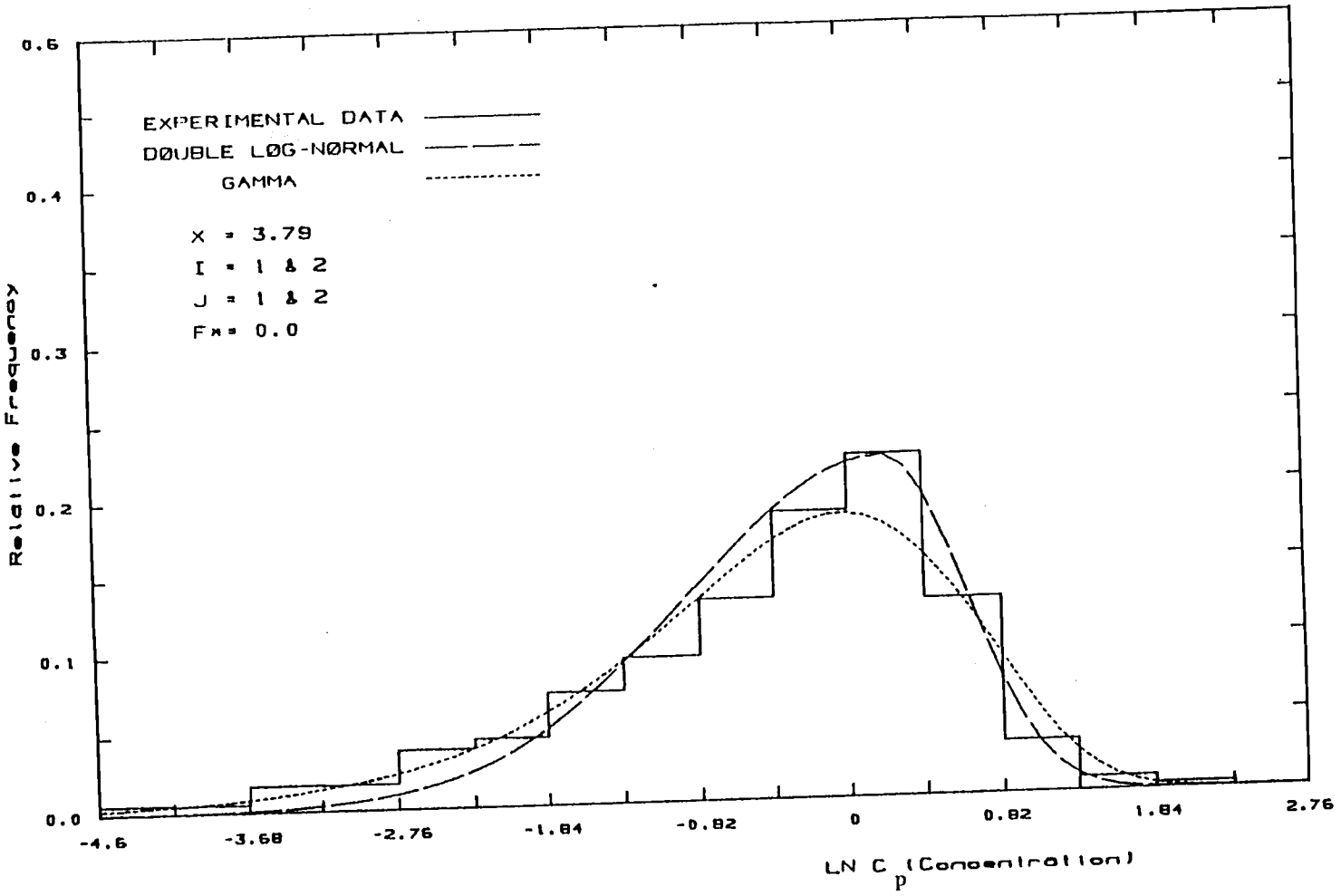

Figure 2 (g) : The probability of occurrences of log-concentration, $P_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=3.79$ for $F_x=0$ case.

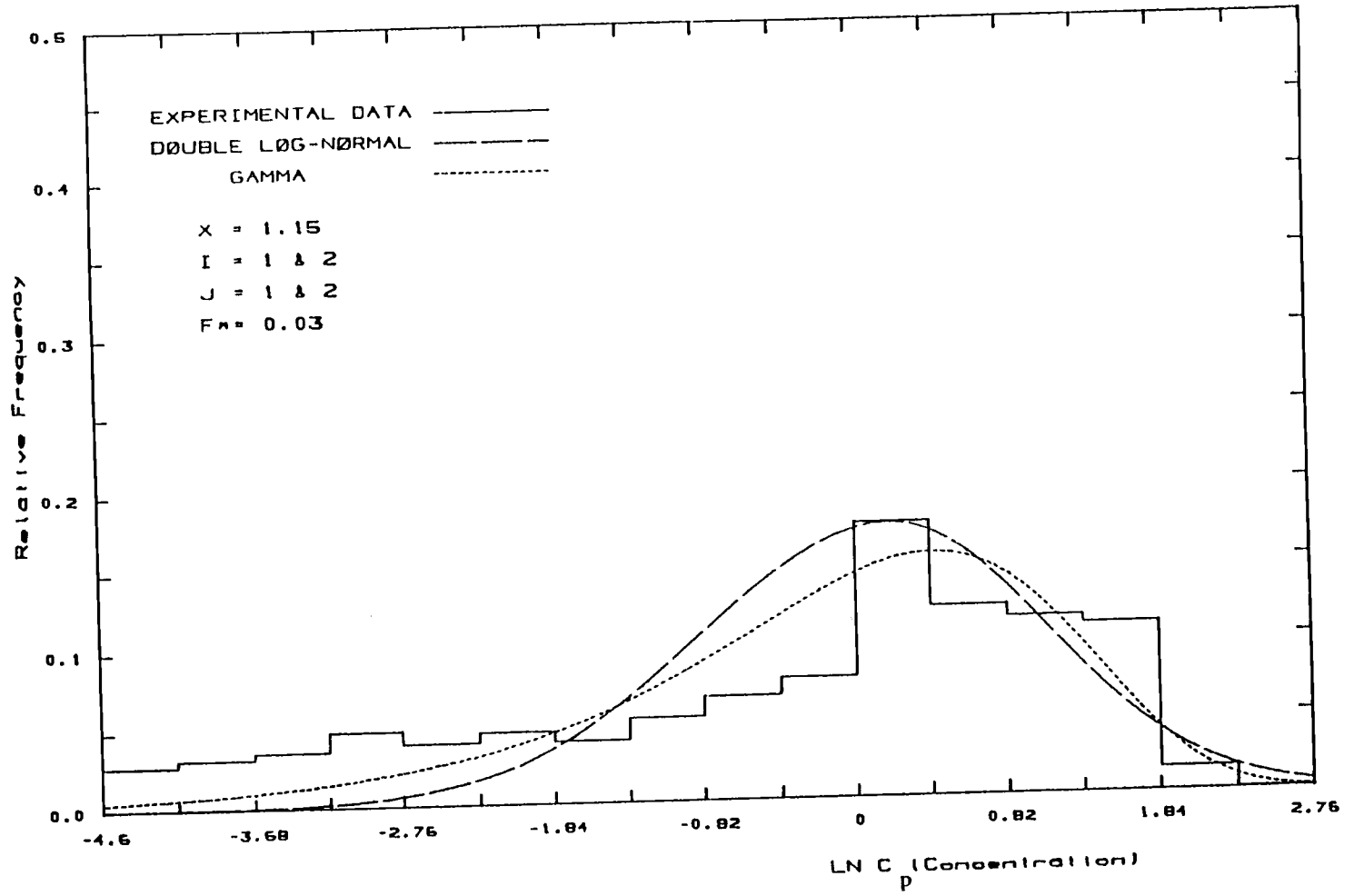

Figure 3 (a) : The probability of occurrences of log-concentration, $P_1(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=1.15$ $F_*=0.03$ case.

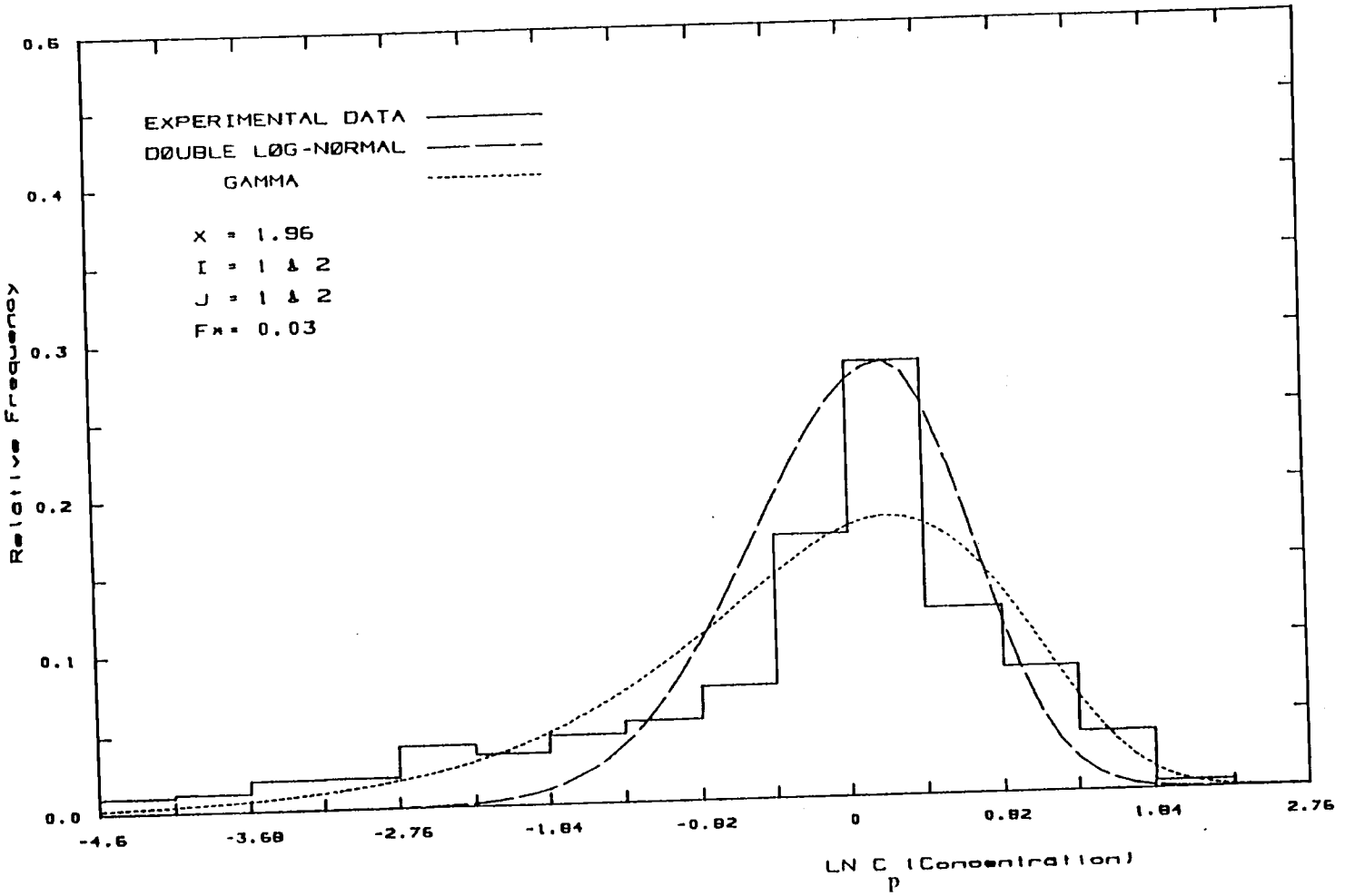

Figure 3 (b) : The probability of occurrences of log-concentration, $P_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at X=1.96 for F_*=0.03 case.

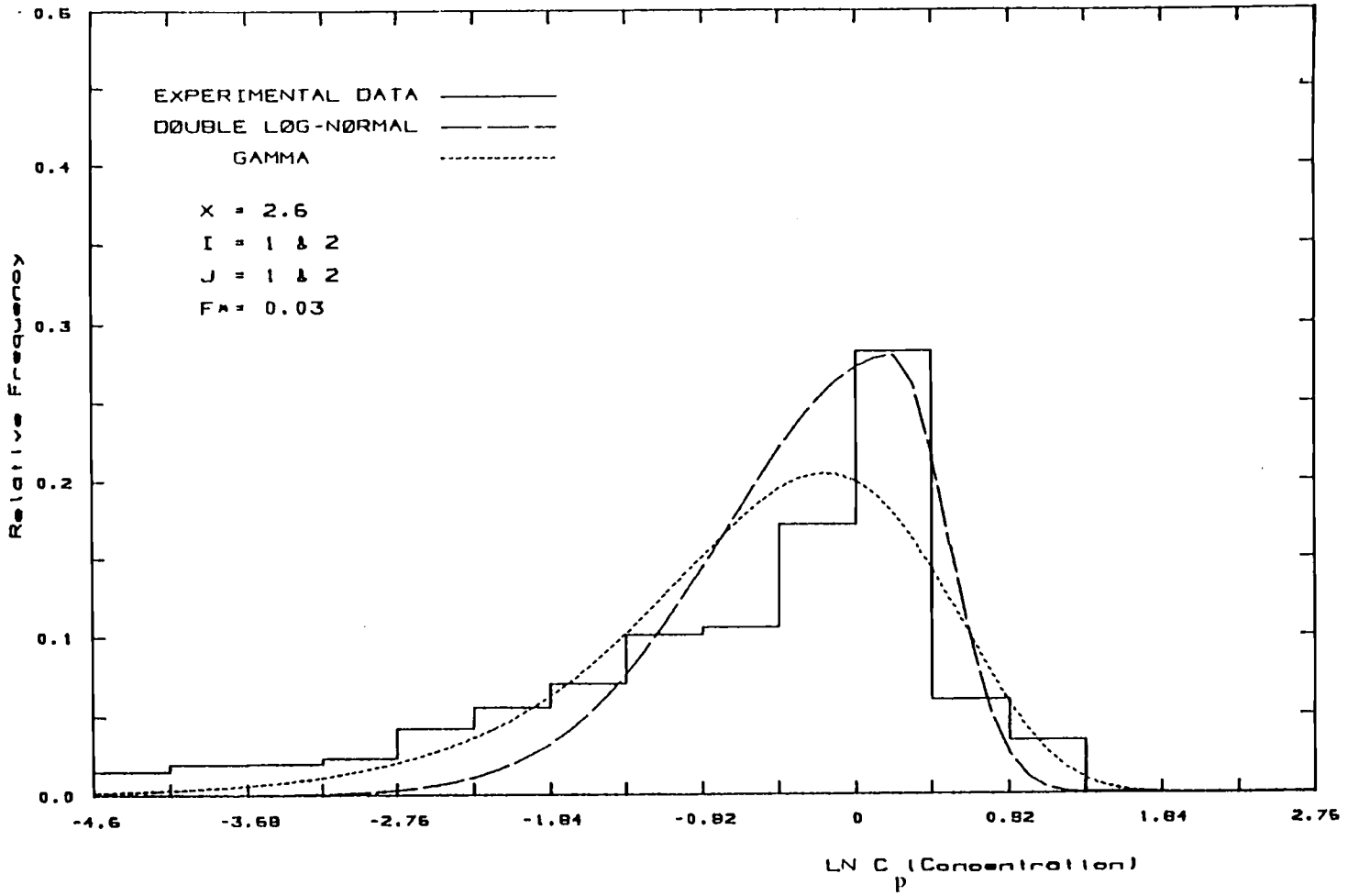

Figure 3 (c) : The probability of occurrences of log-concentration, $P_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at X=2.6 for F_*=0.03 case.

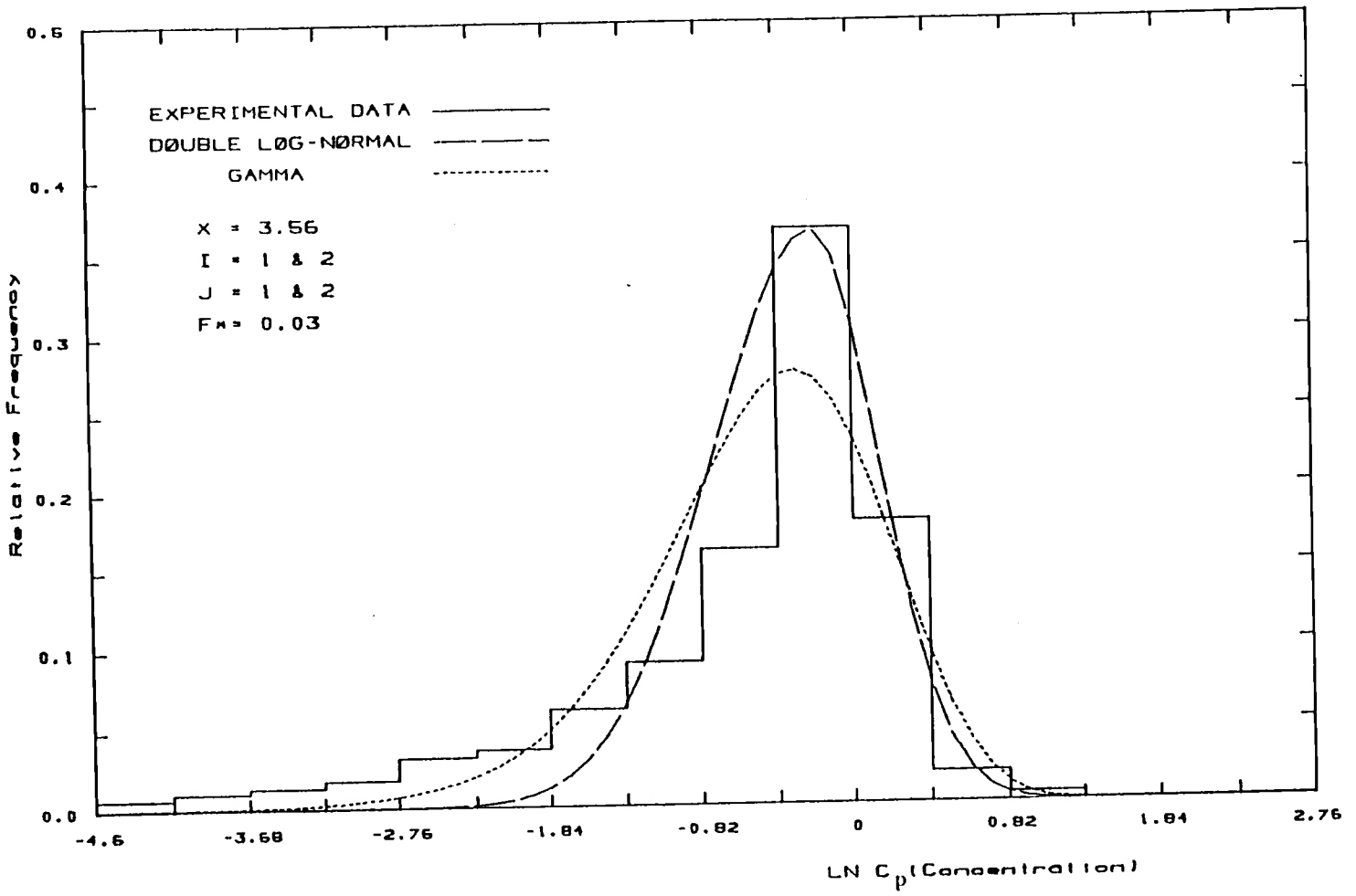

Figure 3 (d) : The probability of occurrences of log-concentration, $P_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at X=3.56 for F_*=0.03 case.

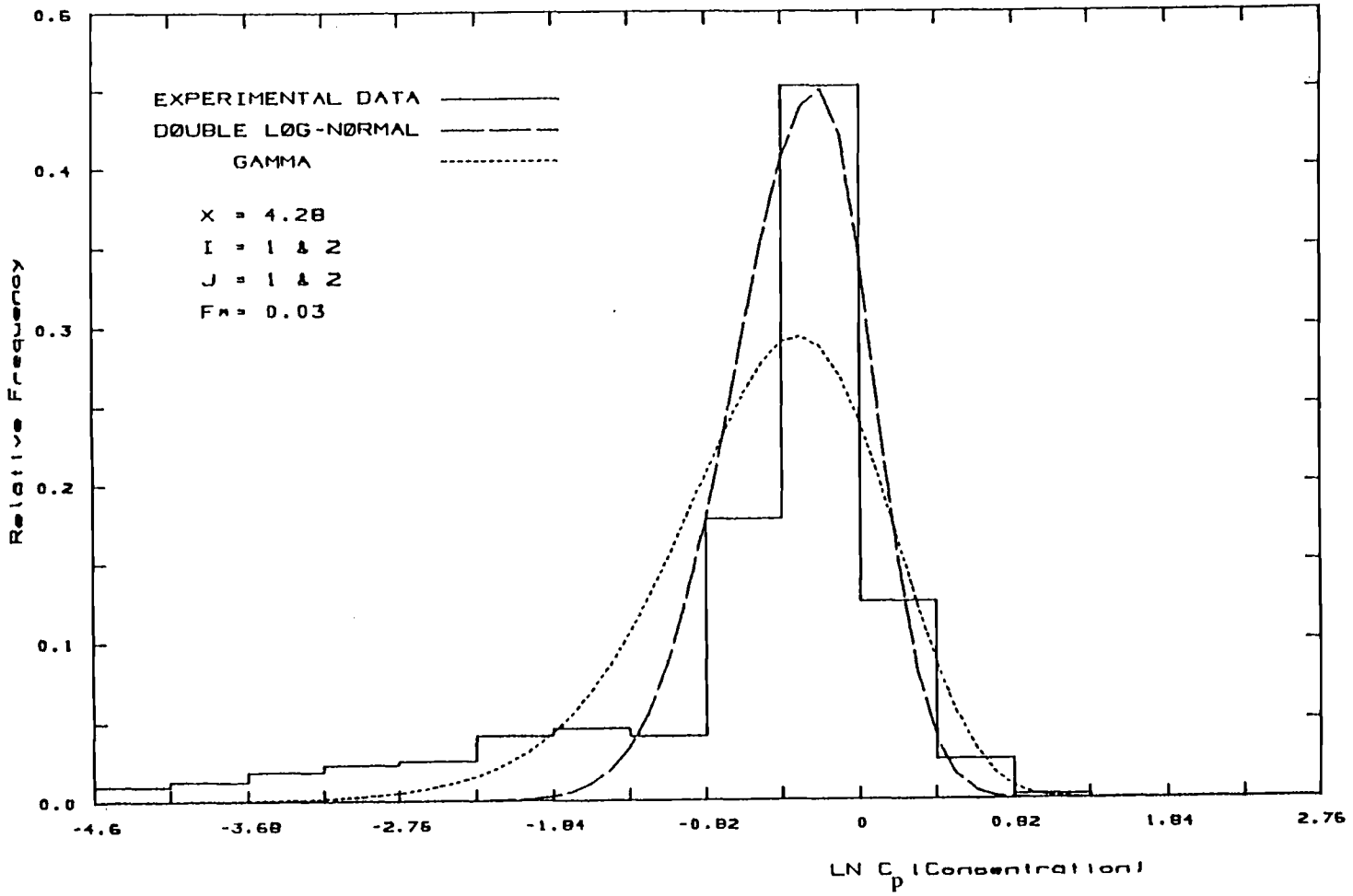

Figure.3 (e) : The probability of occurrences of log-concentration, $P_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=4.28$ for $F_*=0.03$ case.

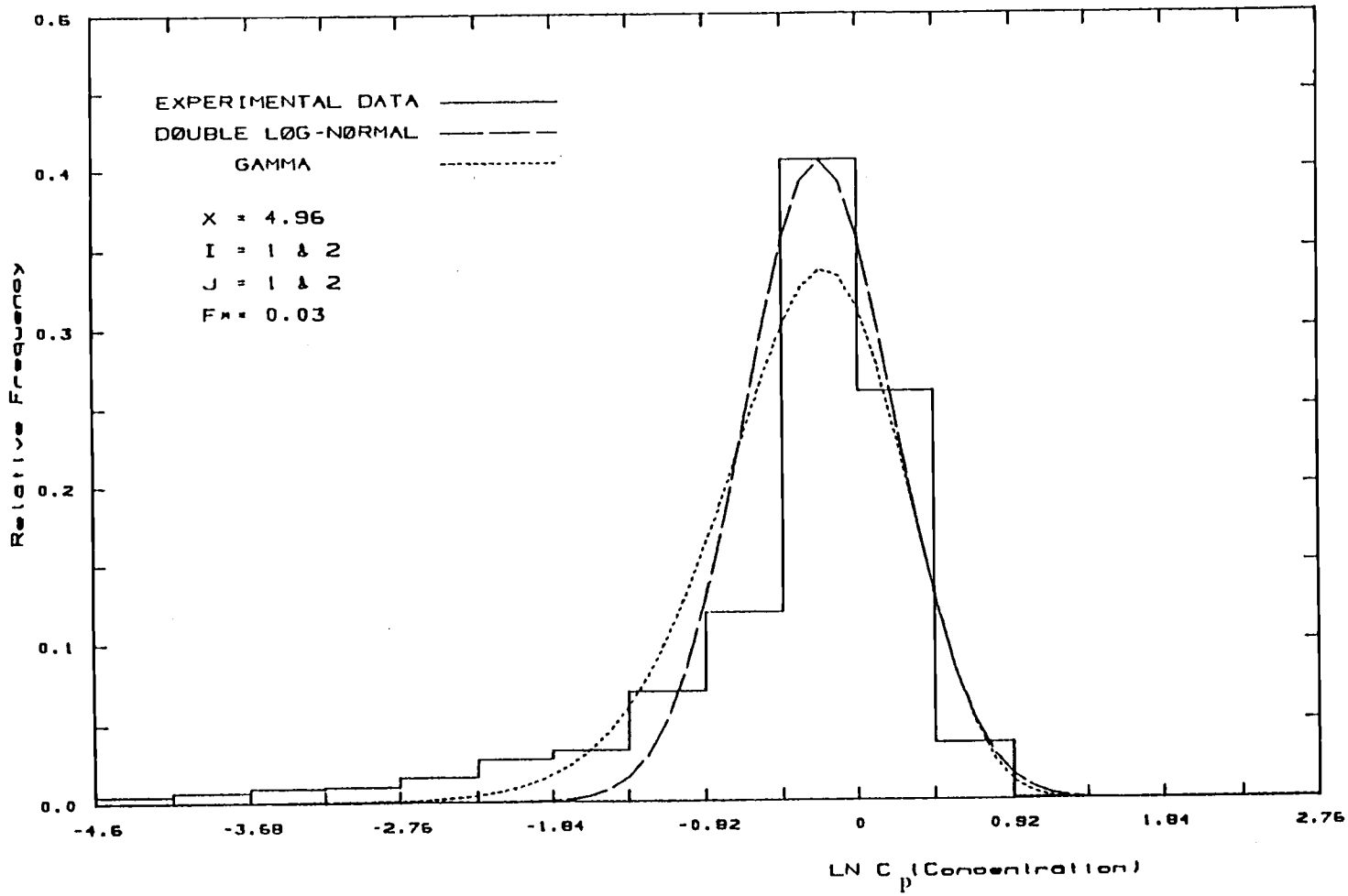

Figure 3 (f) : The probability of occurrences of log-concentration, $P_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=4.96$ for $F_*=0.03$ case.

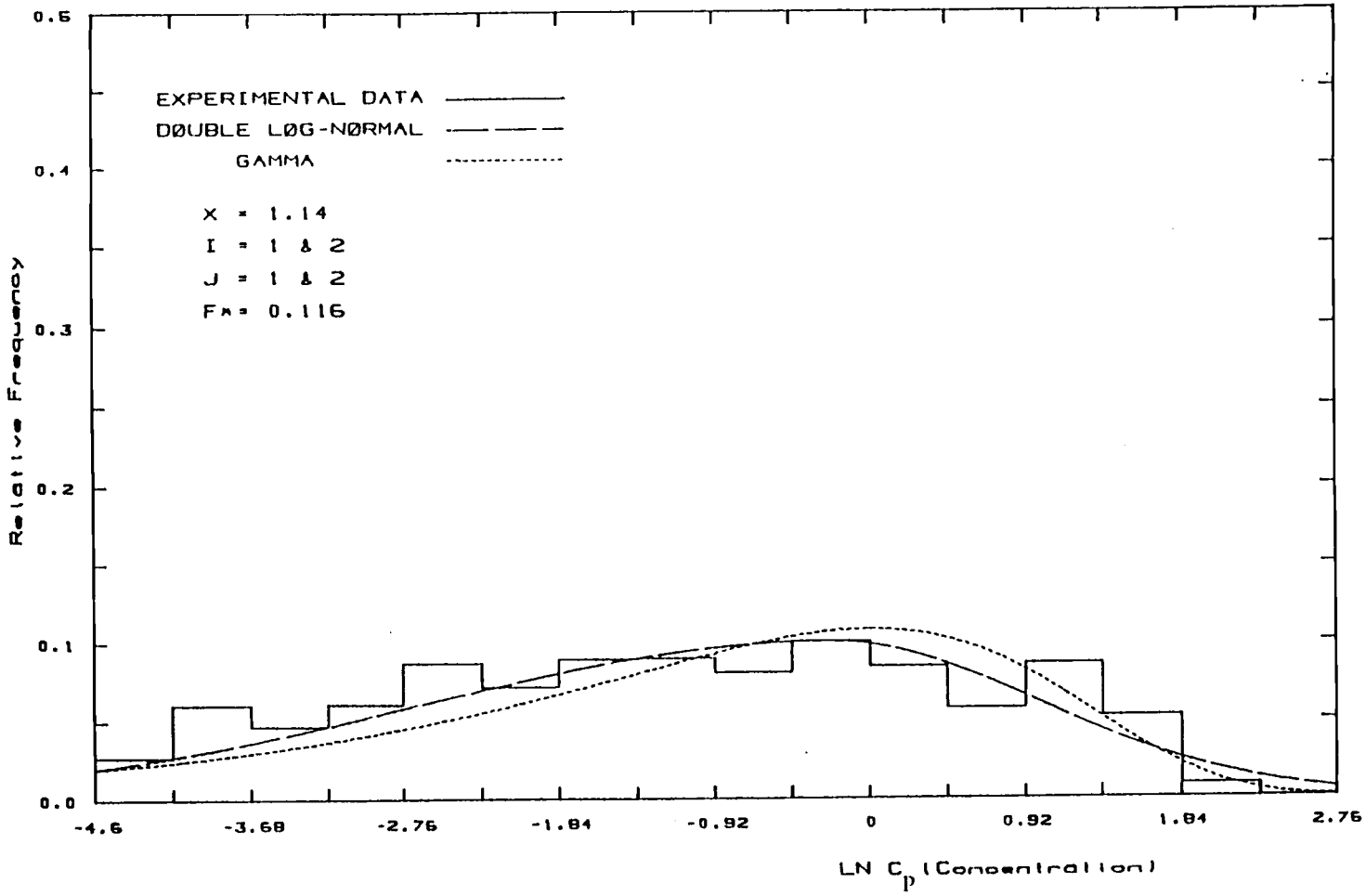

Figure 4 (a) : The probability of occurrences of log-concentration, $P_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=1.14$ for $F_*=0.116$ case.

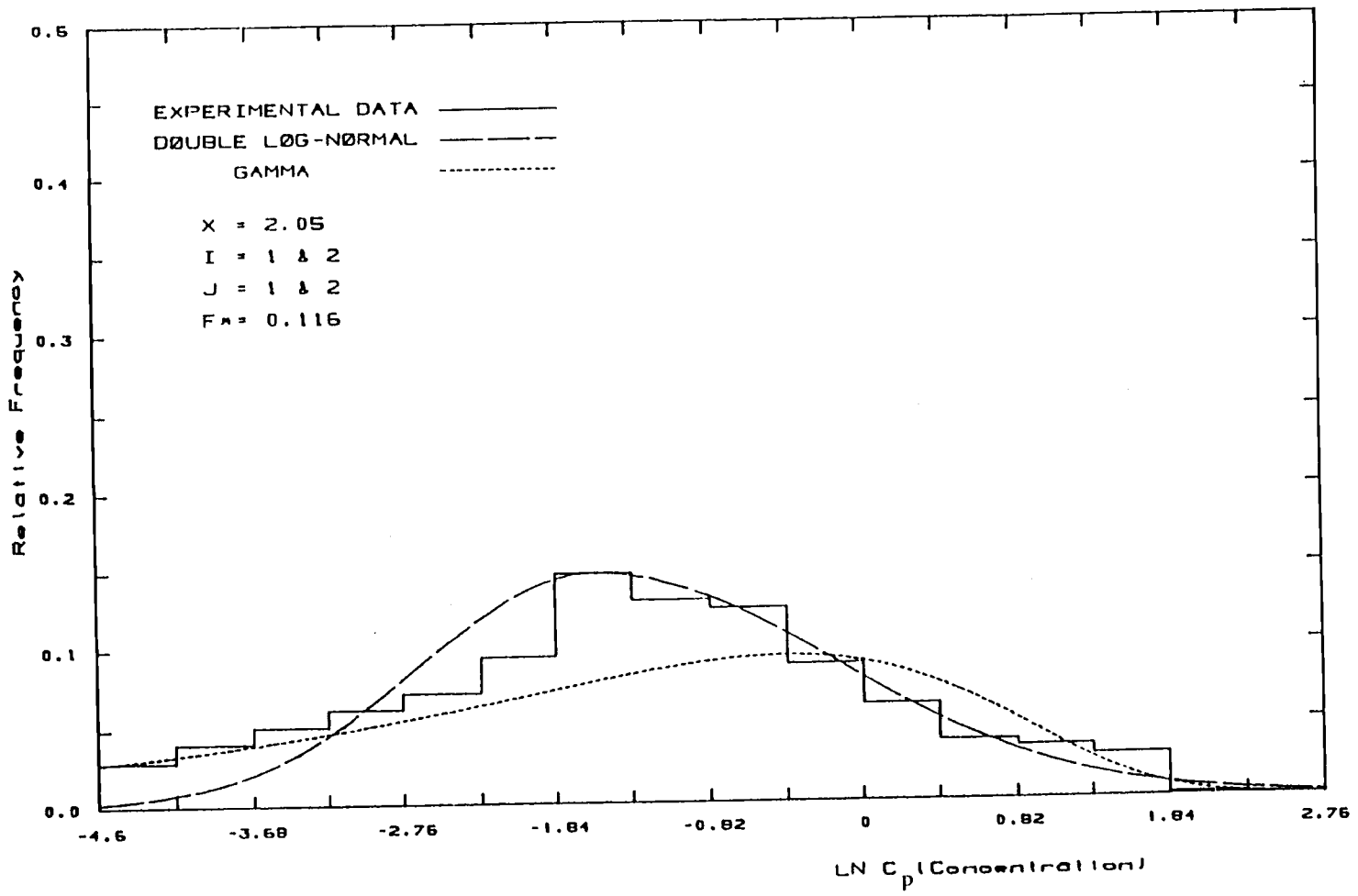

Figure 4 (b) : The probability of occurrences of log-concentration, $P_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=2.05$ for $F_*=0.116$ case.

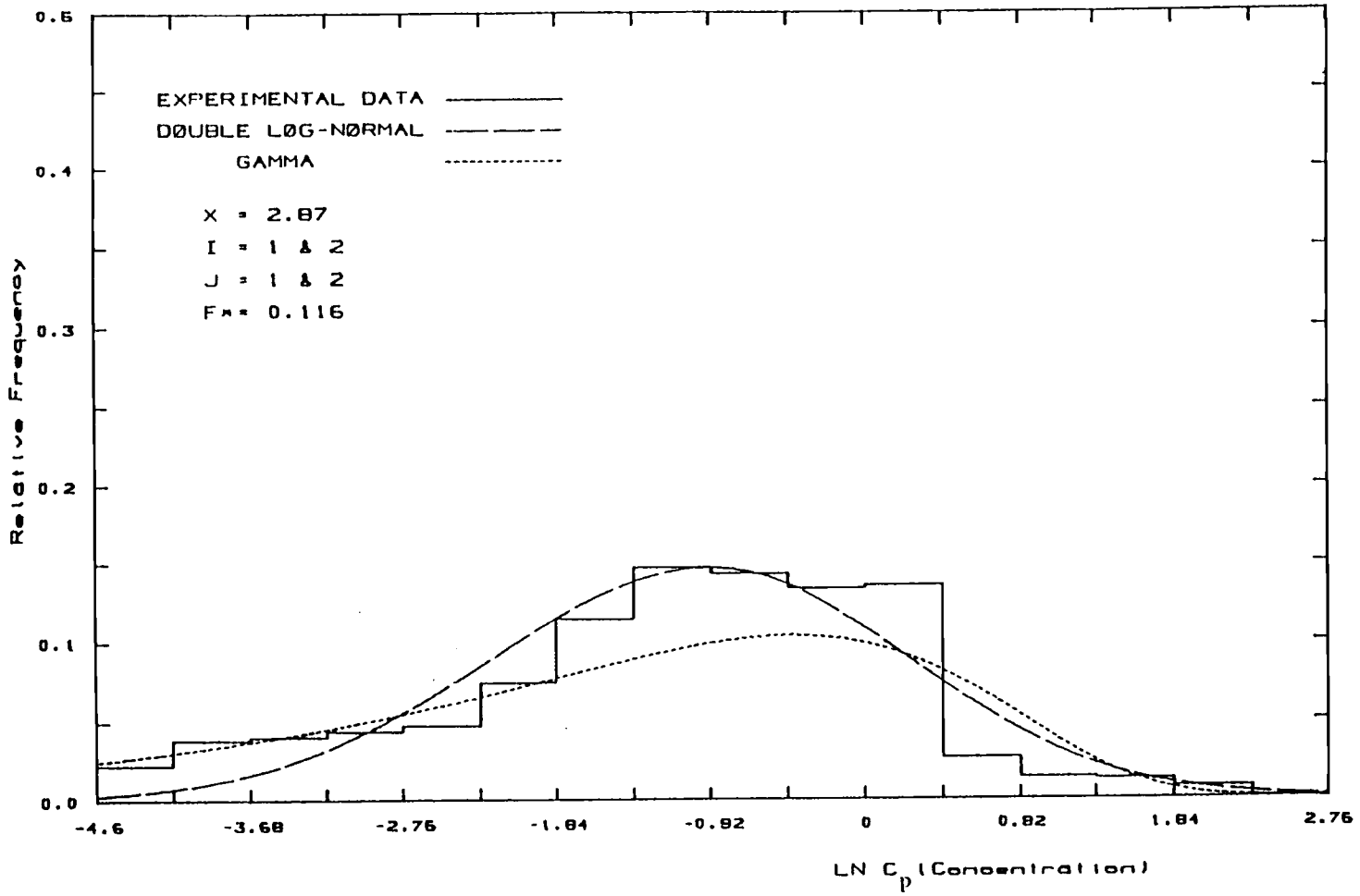

Figure 4 (c) : The probability of occurrences of log-concentration, $P_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at X=2.87 for $F_*=0.116$ case.

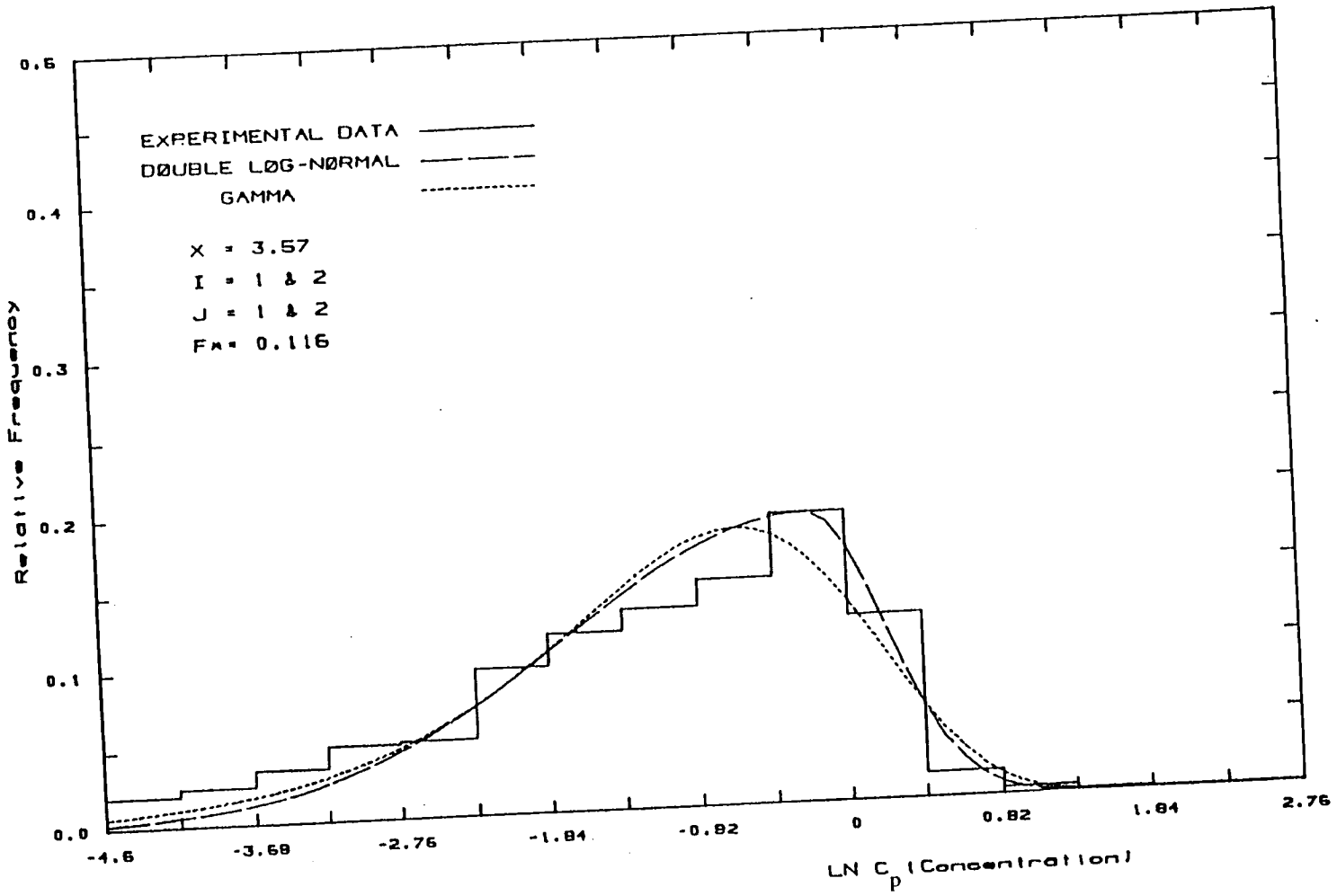

Figure 4 (d) : The probability of occurrences of log-concentration, $P_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=3.57$ for $F_*=0.116$ case.

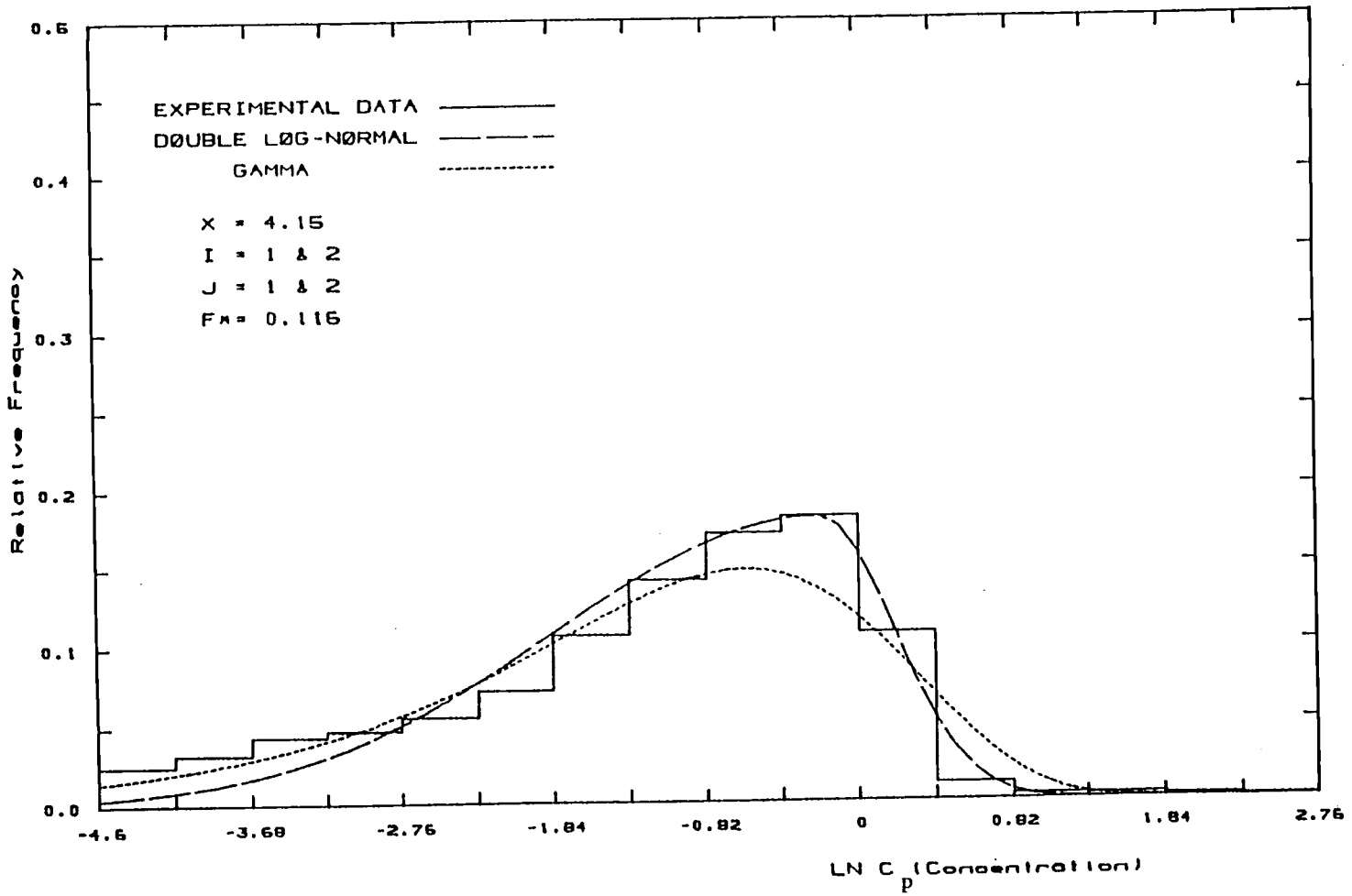

Figure.4 (e) : The probability of occurrences of log-concentration, $P_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at X=4.15 for F_*=0.116 case.

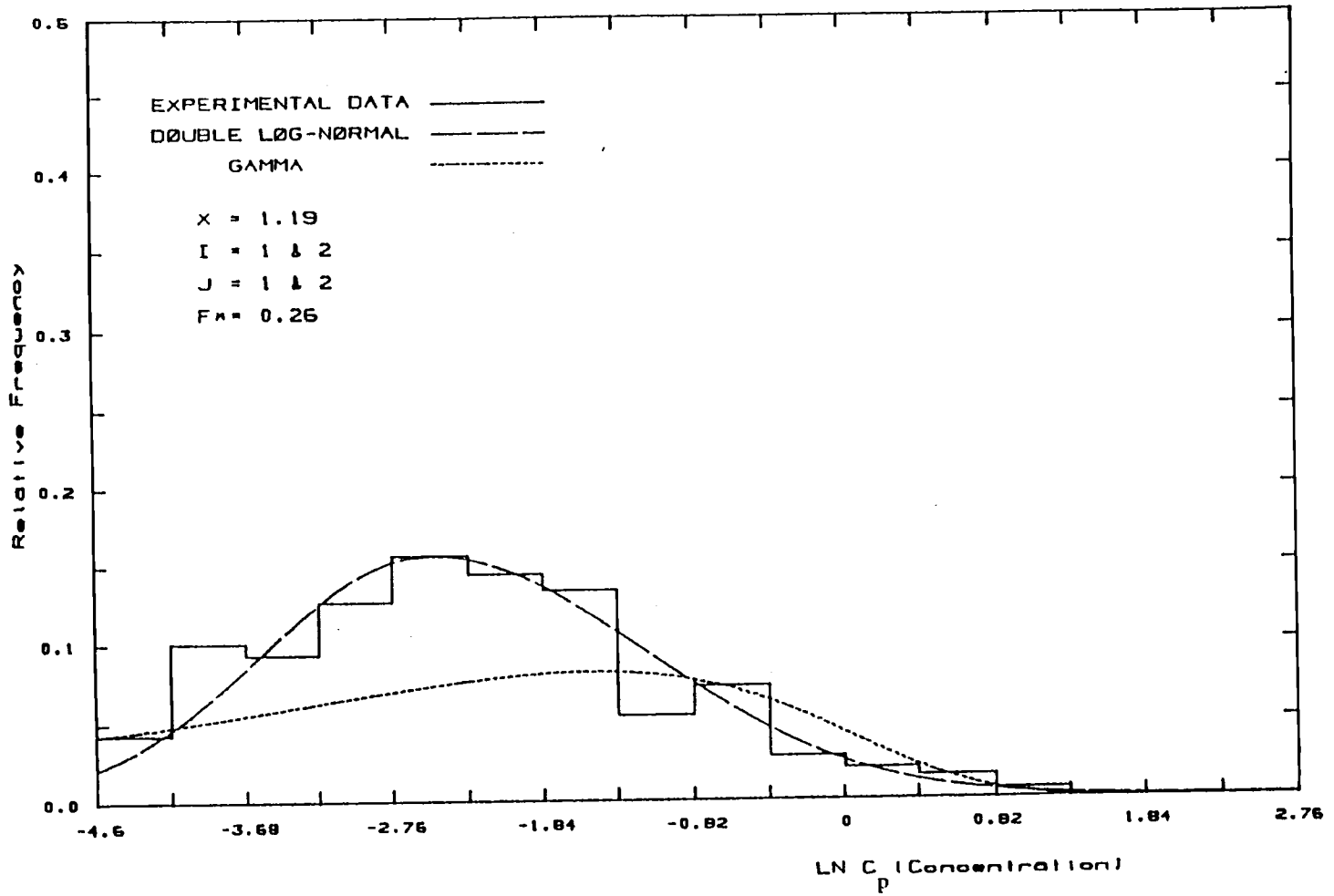

Figure 5 (a) : The probability of occurrences of log-concentration, $P_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=1.19$ for $F_*=0.26$ case.

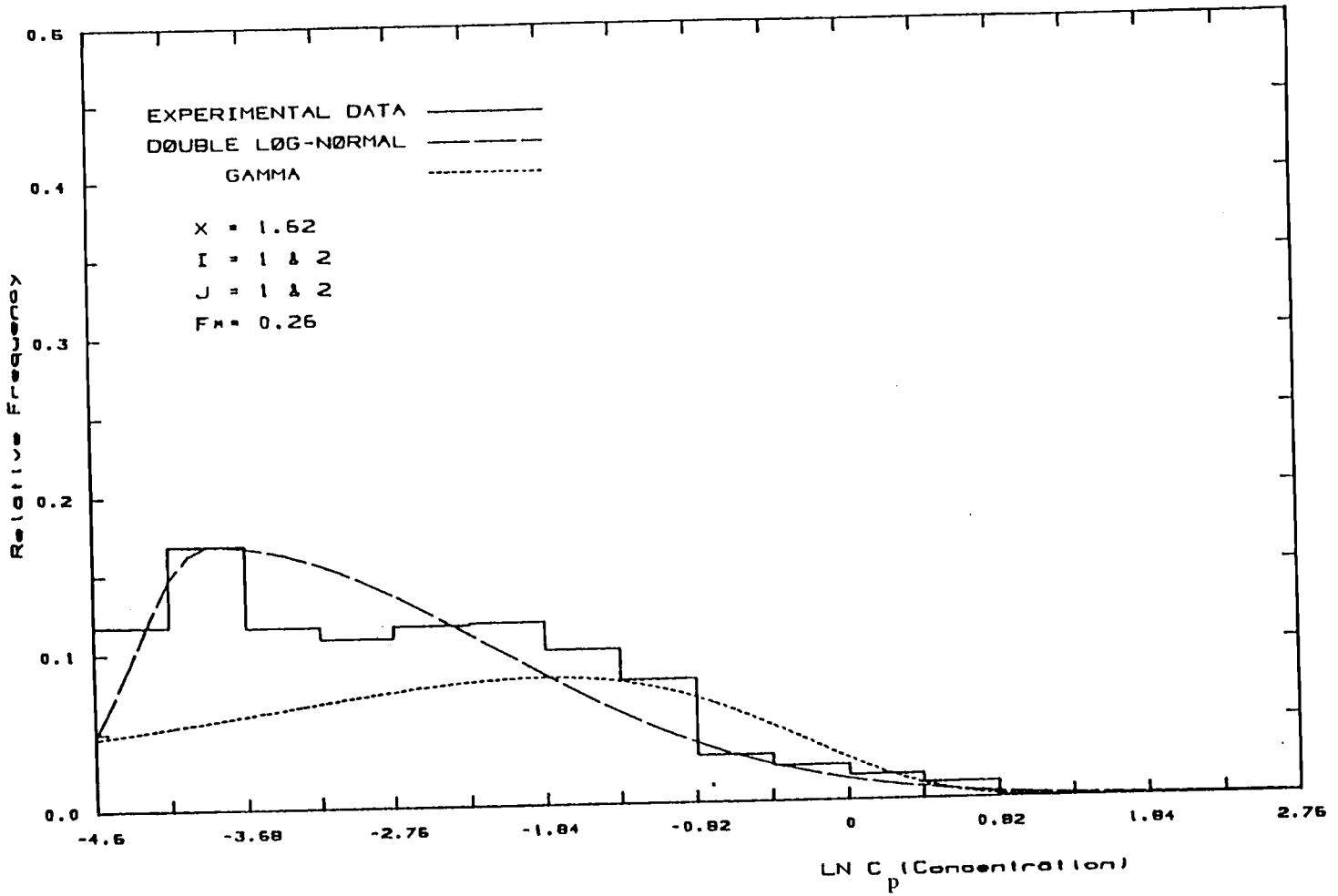

Figure 5 (b) : The probability of occurrences of log-concentration, $P_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=1.62$ for $F_*=0.26$ case.

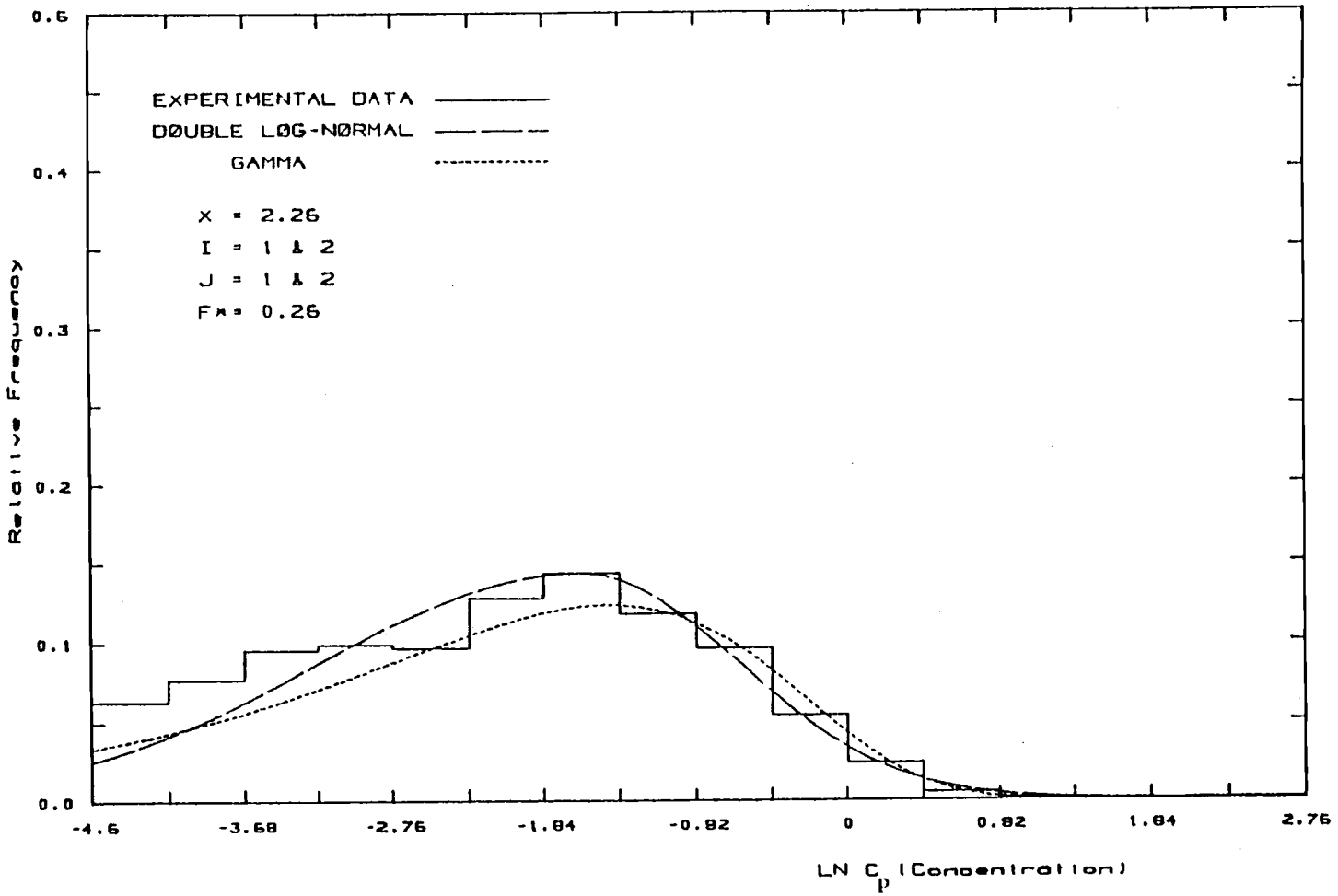

Figure 5 (c) : The probability of occurrences of log-concentration, $P_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=2.26$ for $F_*=0.26$ case.

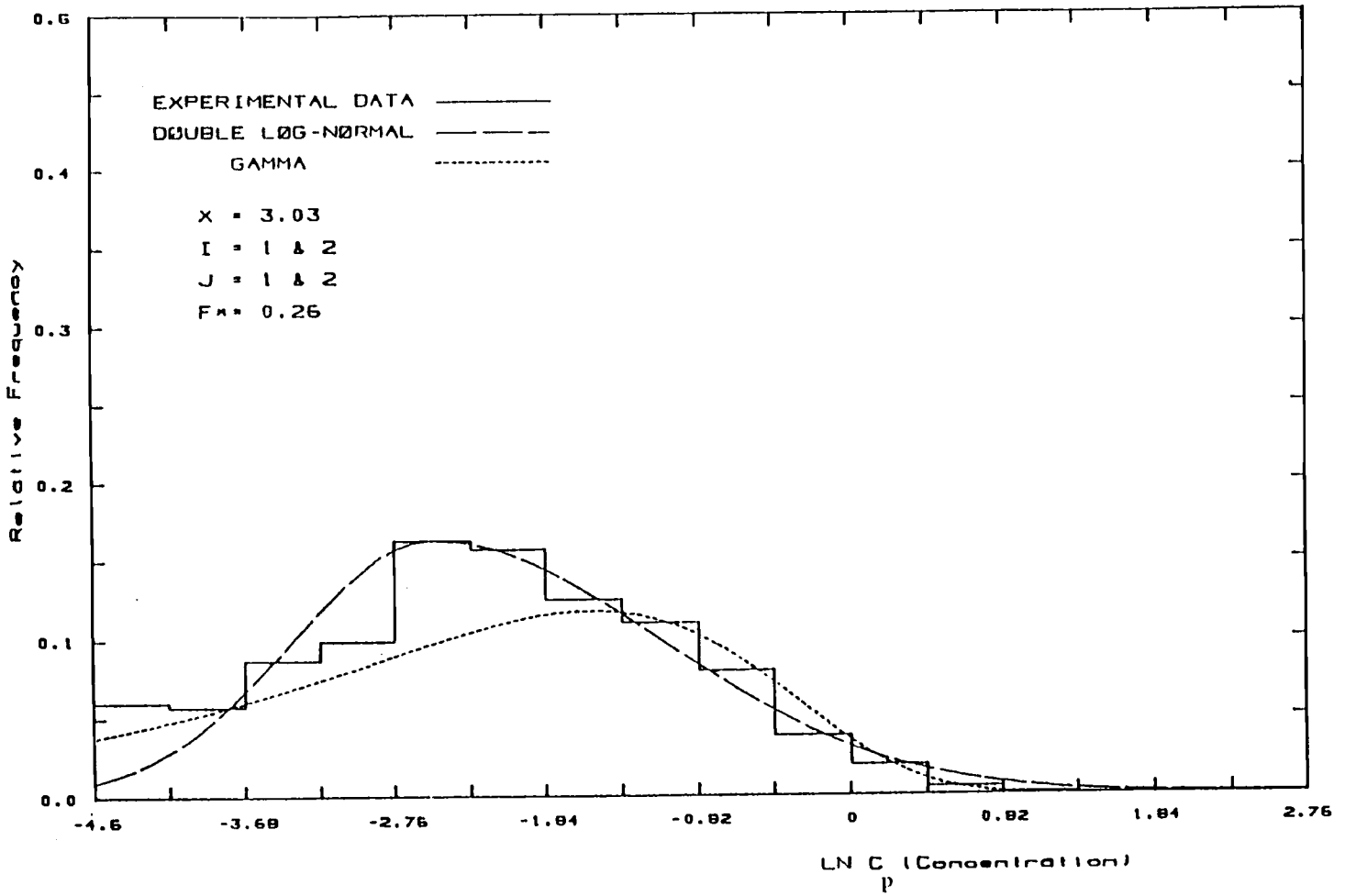

Figure 5 (d) : The probability of occurrences of log-concentration, $P_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=3.03$ for $F_*=0.26$ case.

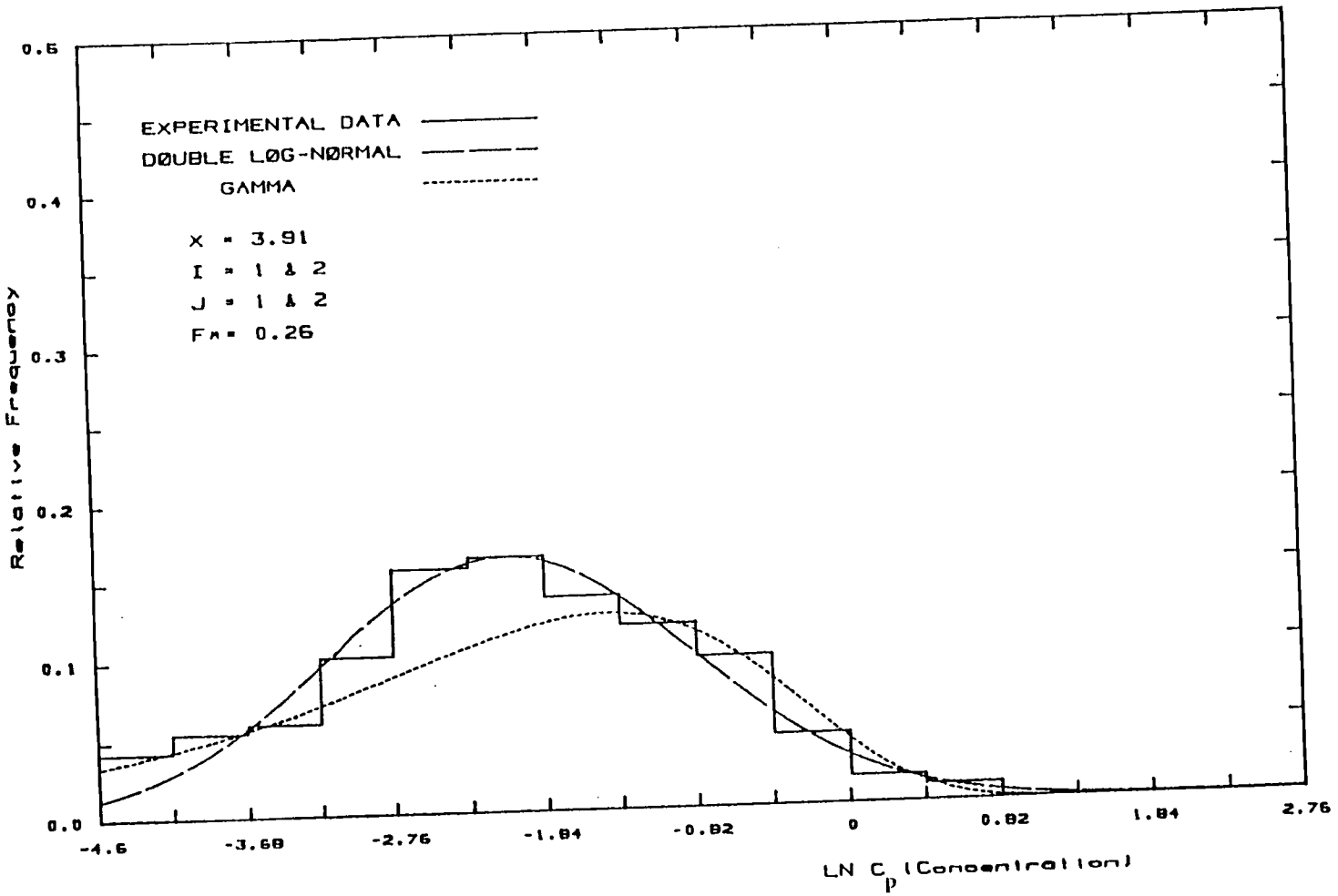

Figure 5 (e) : The probability of occurrences of log-concentration, $P_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at X=3.91 for F_*=0.26 case.

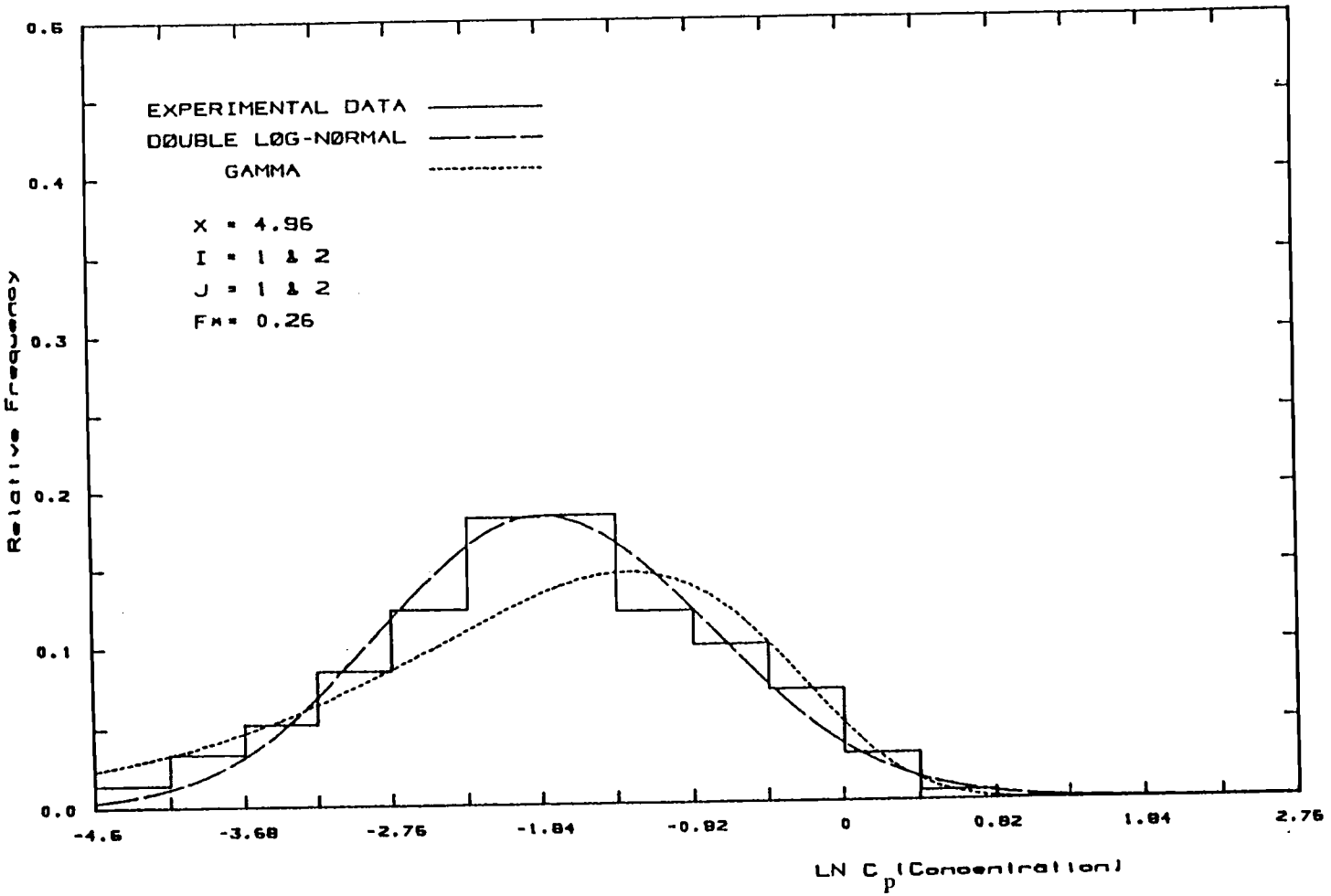

Figure.5 (f) : The probability of occurrences of log-concentration, $P_i(\ln C_p)$, versus the log-concentration, $\ln C_p$, for $|y|<\sigma_y$ and $0<h<0.4$ at $X=4.96$ for $F_*=0.26$ case.

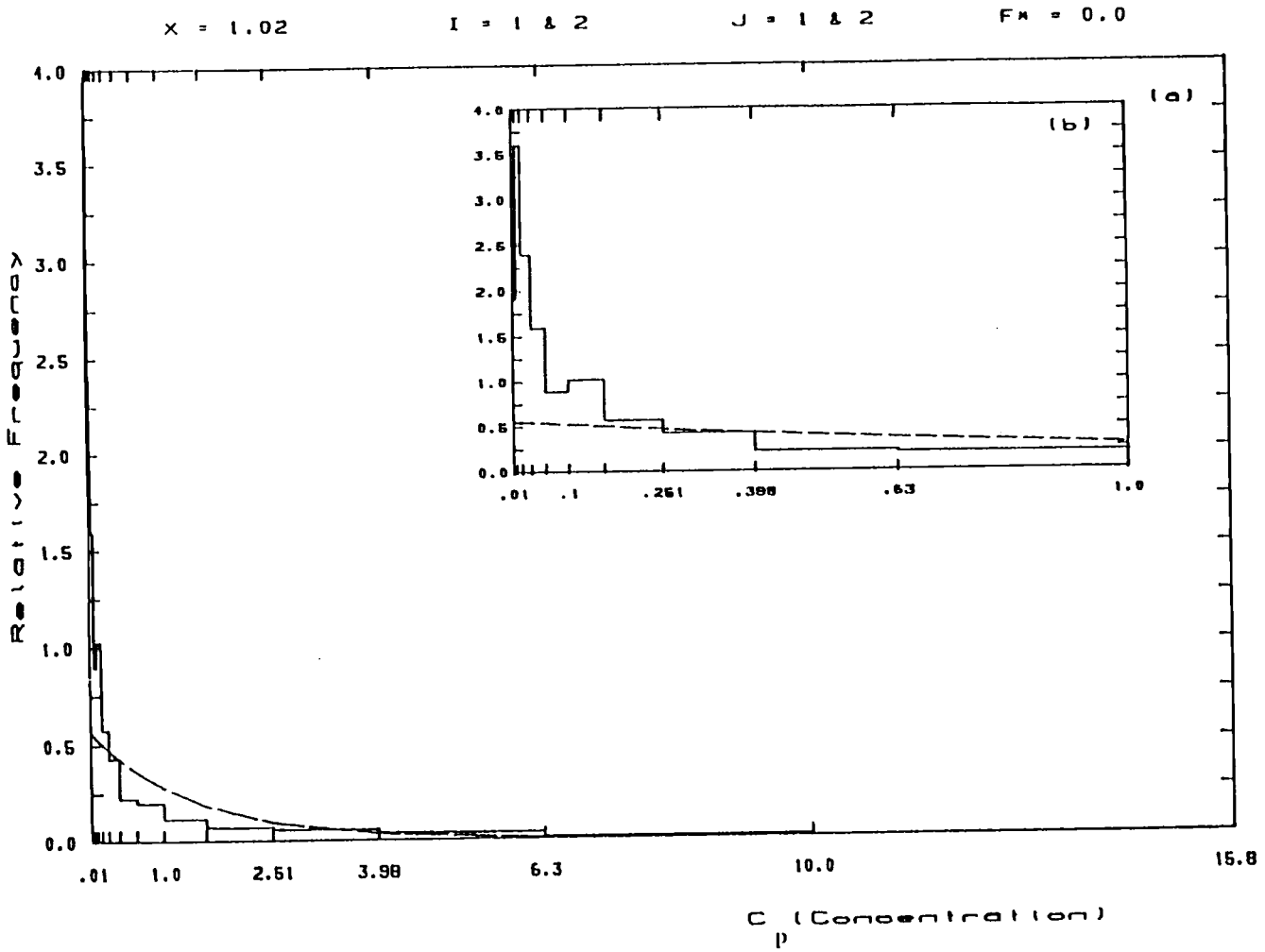

Figure.6 (a) : The probability of occurrences of concentration, $P_i(C_p)$, versus the concentration for $|y|<\sigma_y$ and $0<h<0.4$ for $F_*=0$ case. Inside window is put on the diagram for a better presentation of the exponential probabilities as compared to the experimental data at small concentrations, $C_p<1$.

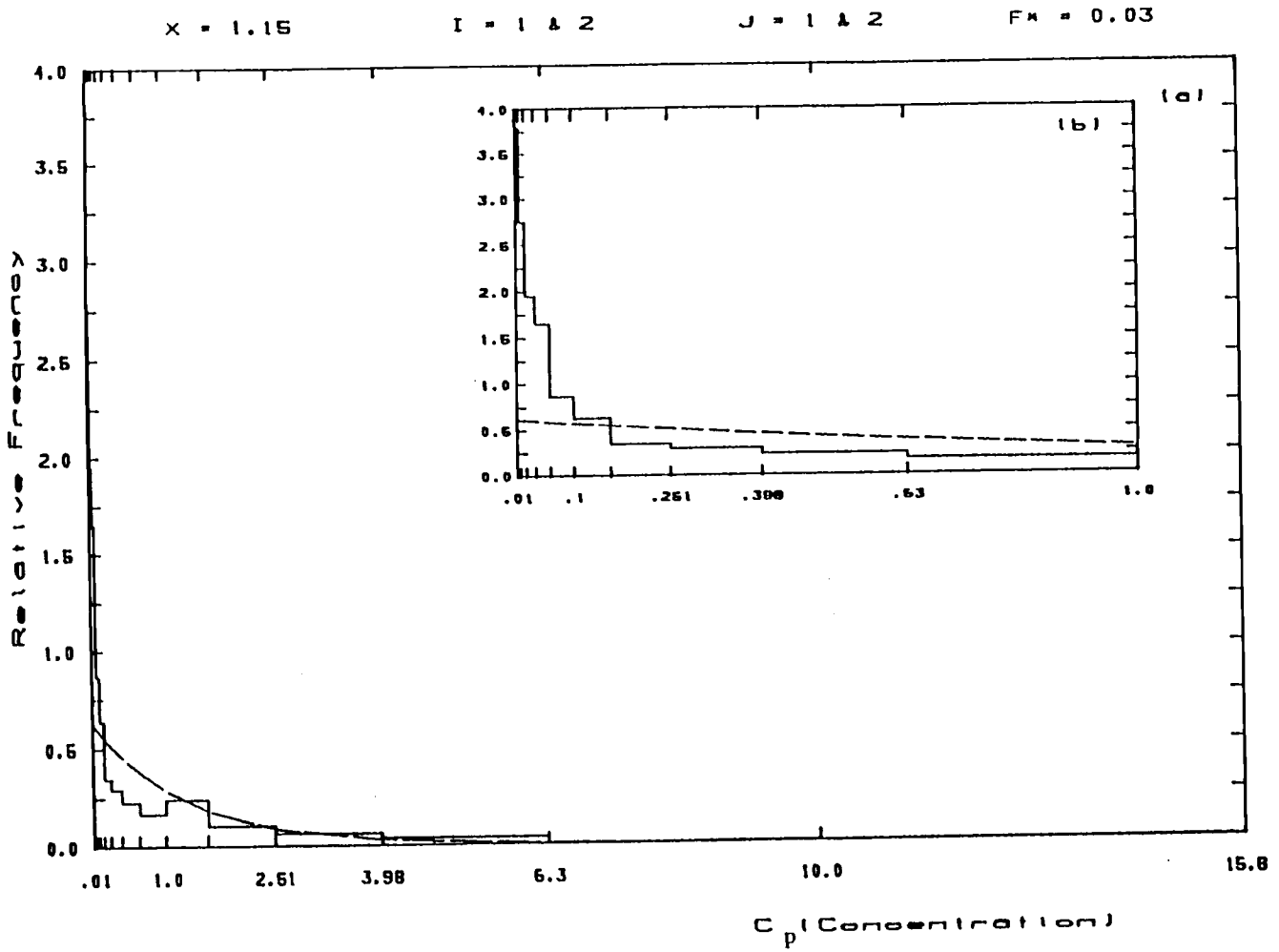

Figure 6 (b) : The probability of occurrences of concentration, $P_i(C_p)$, versus the concentration for $|y|<\sigma_y$ and $0<h<0.4$ for $F_*=0.03$ case. Inside window is put on the diagram for a better presentation of the exponential probabilities as compared to the experimental data at small concentrations, $C_p<1$.

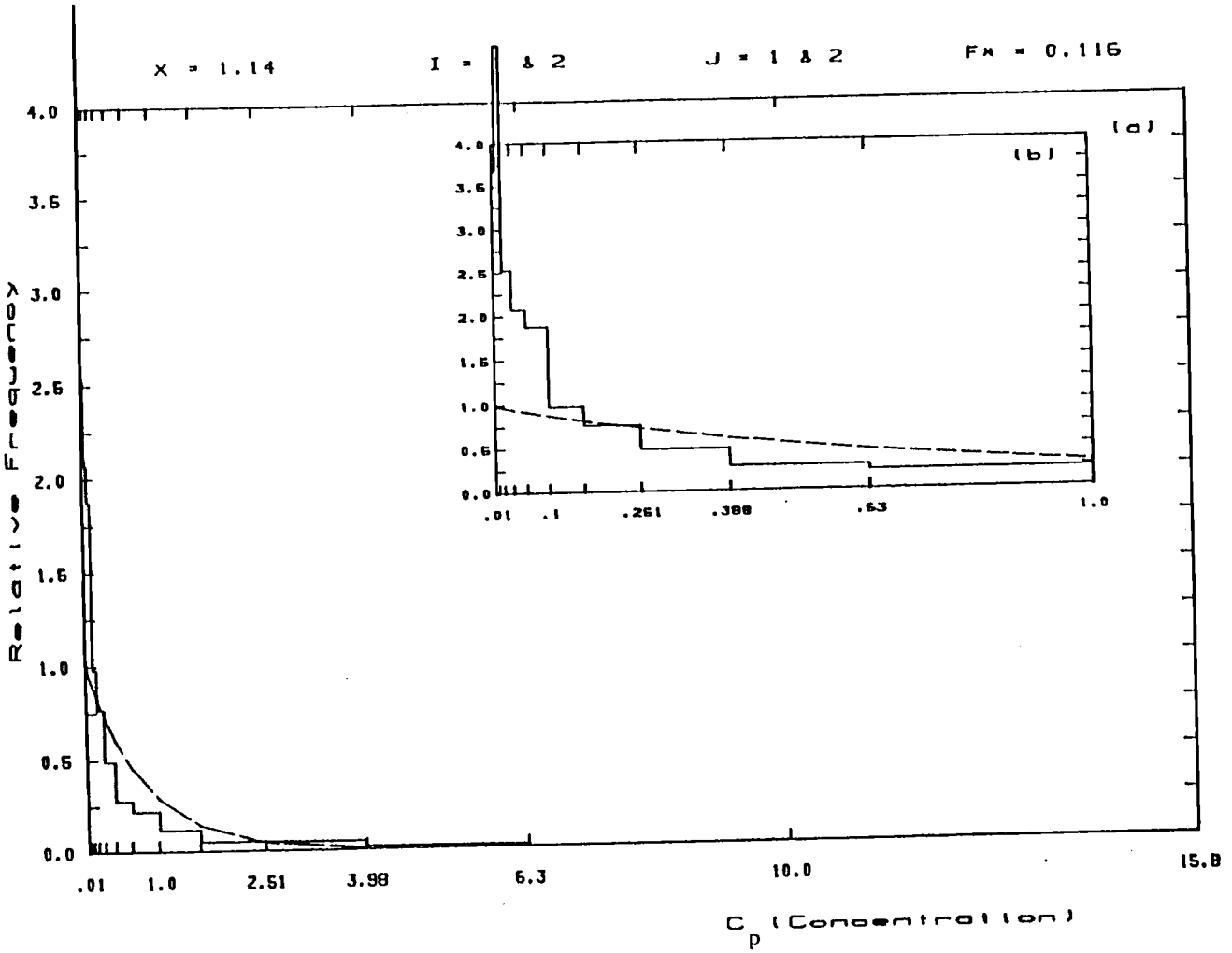

Figure 6 (c) : The probability of occurrences of concentration, $P_i(C_p)$, versus the concentration for $|y|<\sigma_y$ and $0<h<0.4$ for $F_*=0.116$ case. Inside window is put on the diagram for a better presentation of the exponential probabilities as compared to the experimental data at small concentrations, $C_p<1$.

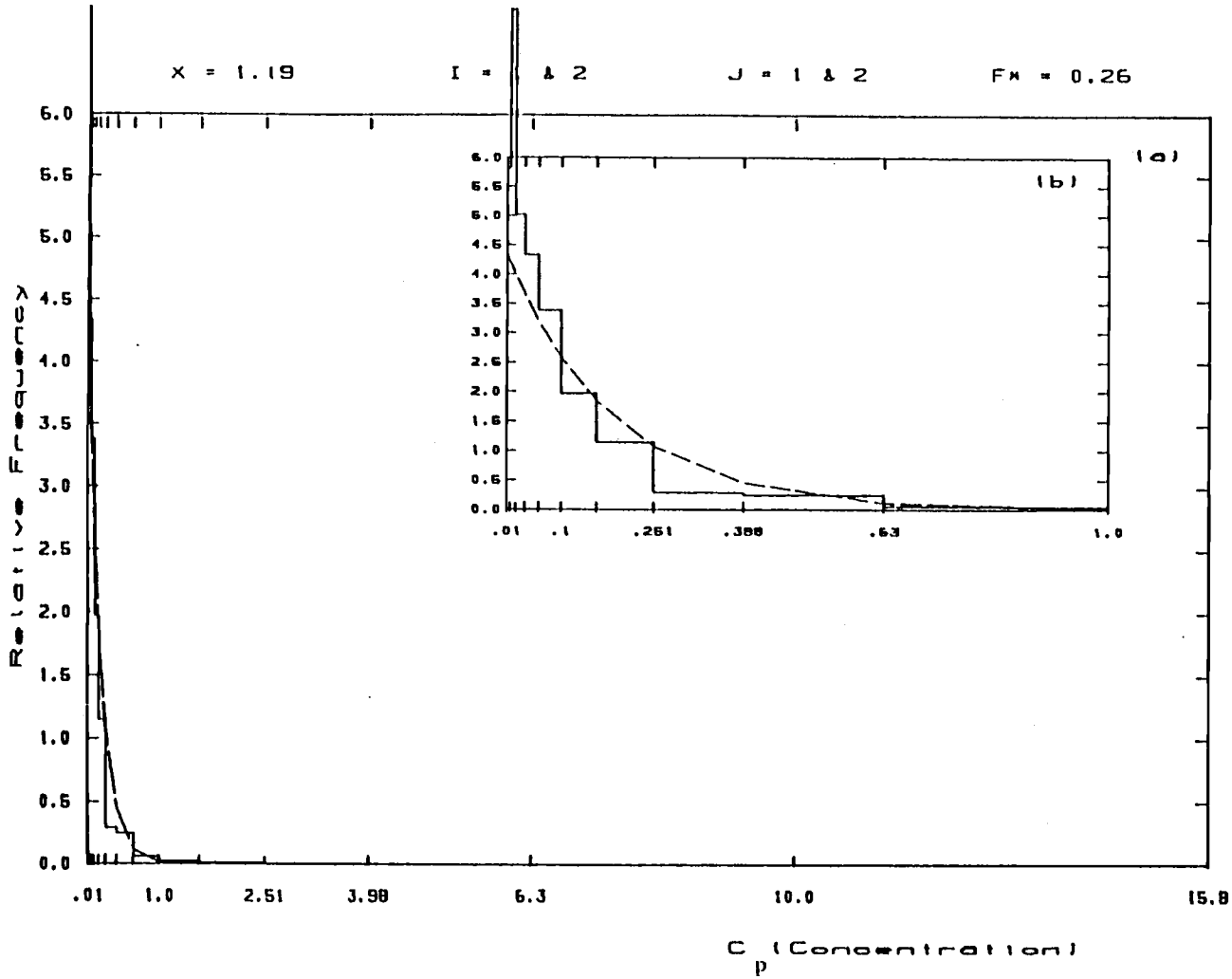

Figure 6 (d) : The probability of occurrences of concentration, $P_i(C_p)$, versus the concentration for $|y|<\sigma_y$ and $0<h<0.4$ for $F_*=0.26$ case. Inside window is put on the diagram for a better presentation of the exponential probabilities as compared to the experimental data at small concentrations, $C_p<1$.

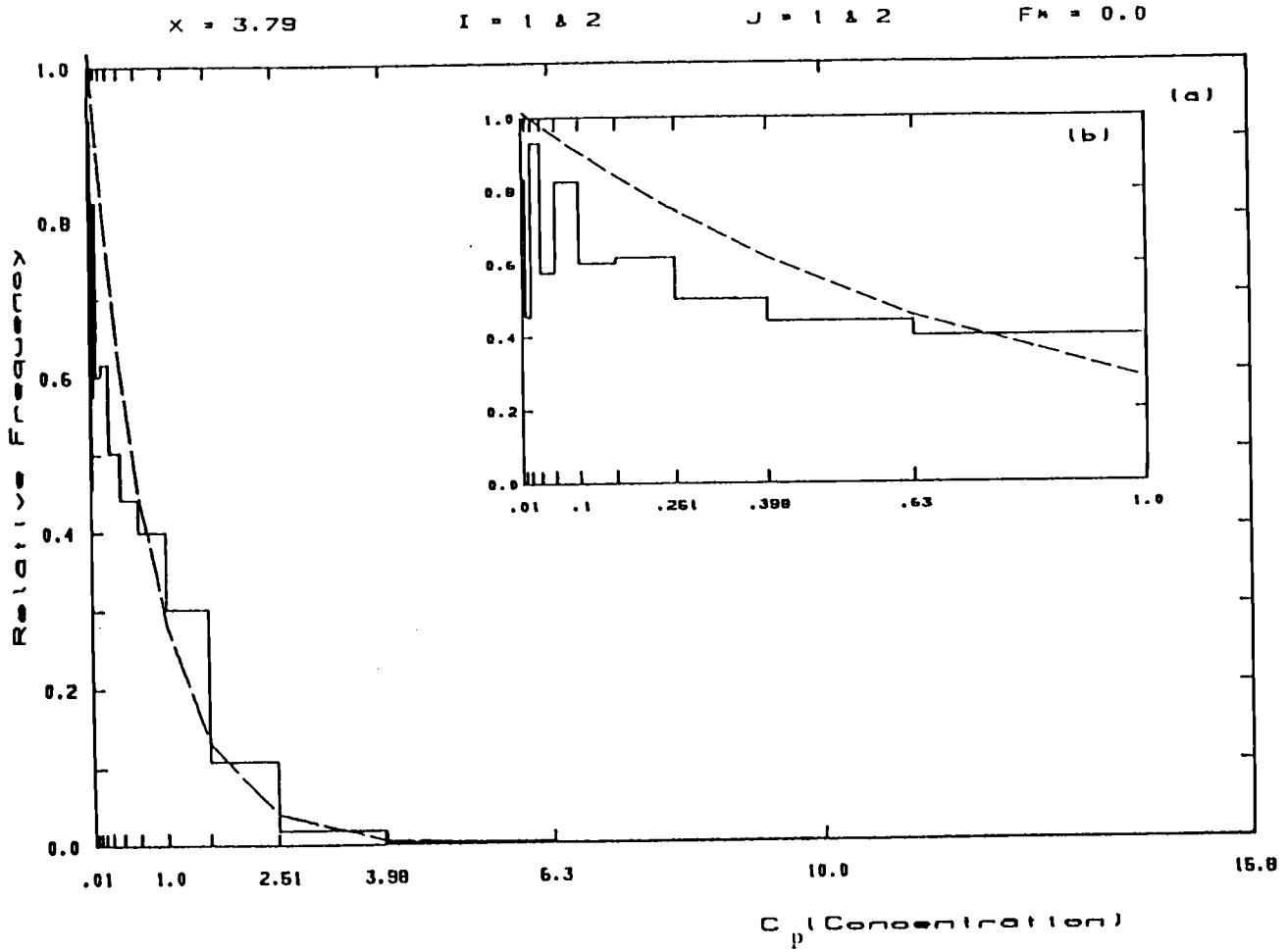

Figure 7 (a) : The probability of occurrences of concentration, $P_i(C_p)$, versus the concentration for $|y|<\sigma_v$ and $0<h<0.4$ for $F_*=0$ case. Inside window is put on the diagram for a better presentation of the exponential probabilities as compared to the the experimental data at small concentrations, $C_p<1$.

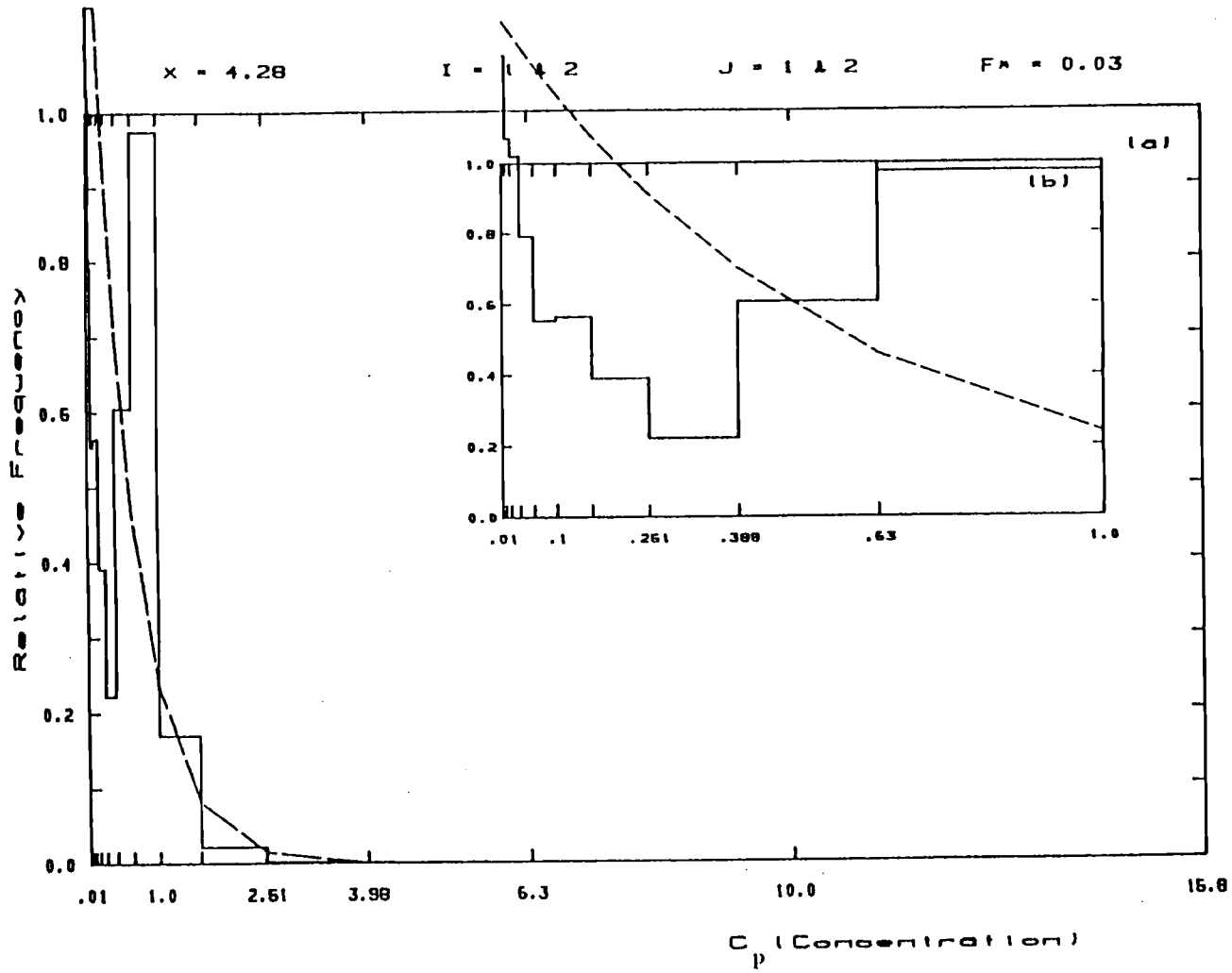

Figure 7 (b) : The probability of occurrences of concentration, $P_1(C_p)$, versus the concentration for $|y|<\sigma_v$ and $0<h<0.4$ for $F_*=0.03$ case. Inside window is put on the diagram for a better presentation of the exponential probabilities as compared to the experimental data at small concentrations, $C_p<1$.

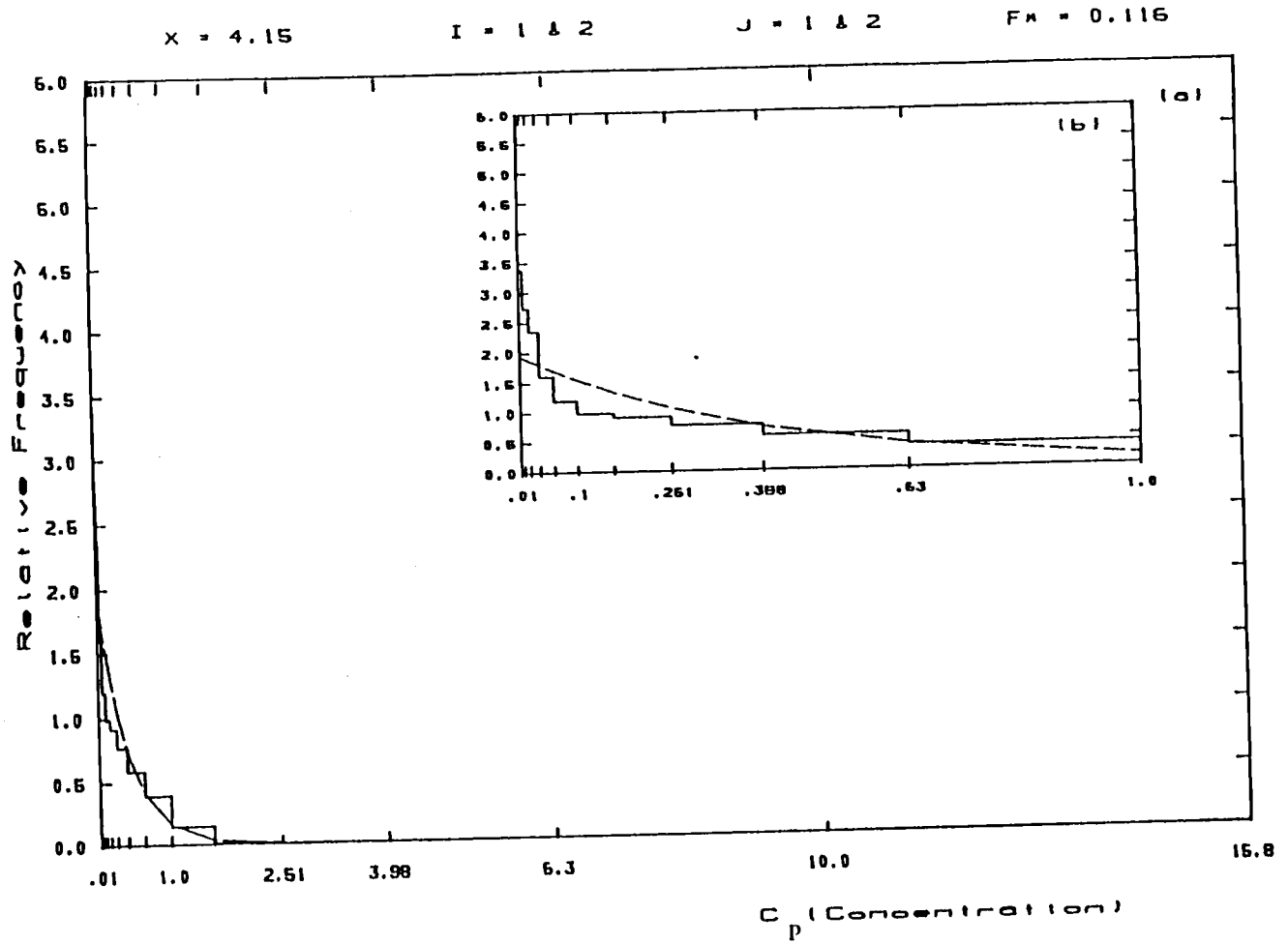

Figure 7 (c) : The probability of occurrences of concentration, $P_i(C_p)$, versus the concentration for $|y|<\sigma_y$ and $0<h<0.4$ for $F=0.116$ case. Inside window is put on the diagram for a better presentation of the exponential probabilities as compared to the experimental data at small concentrations, $C_p<1$.

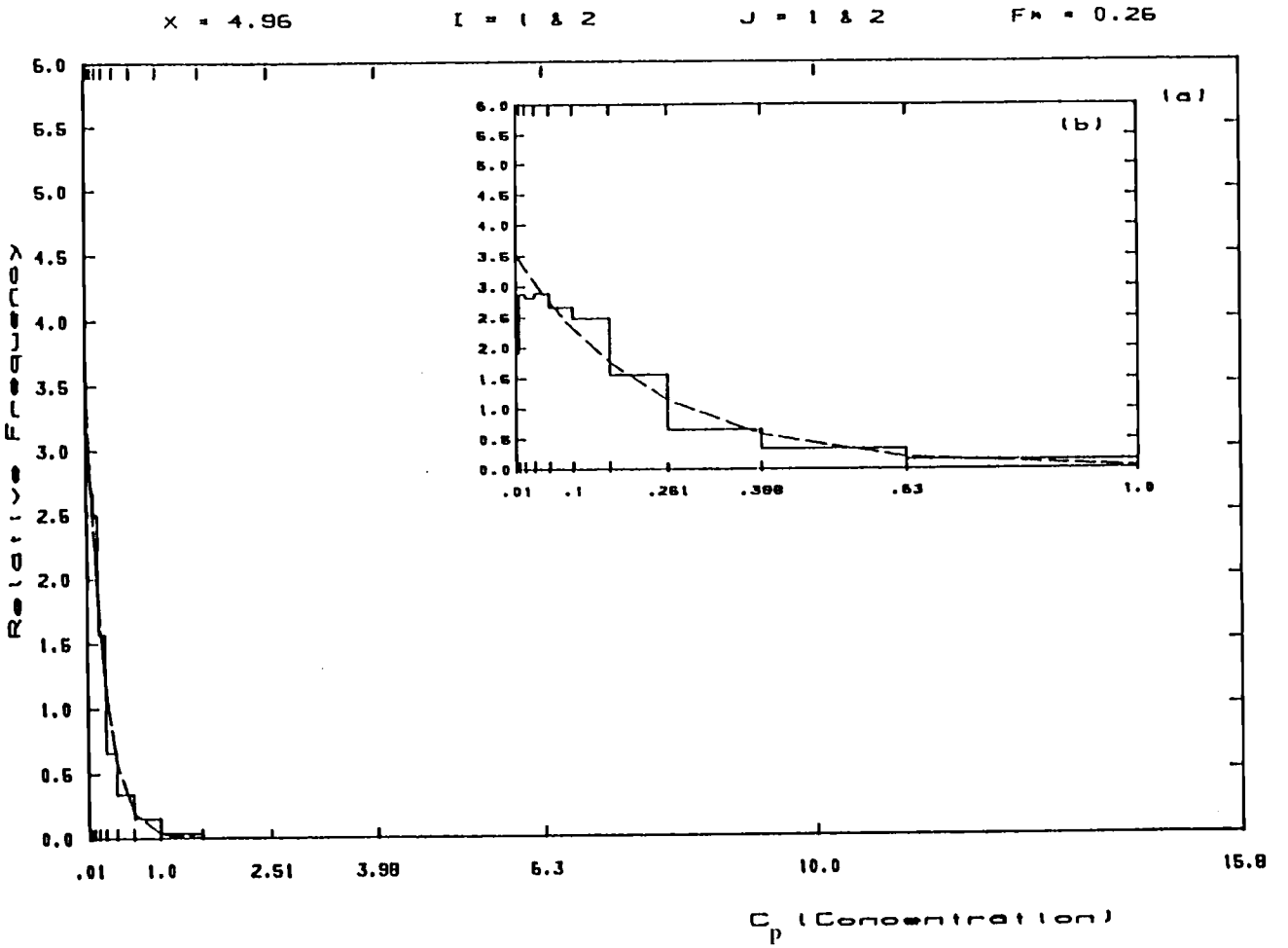

Figure 7 (d) : The probability of occurrences of concentration, $P_1(C_p)$, versus the concentration for $|y|<\sigma_y$ and $0<h<0.4$ for $F_*=0.26$ case. Inside window is put on the diagram for a better presentation of the exponential probabilities as compared to the experimental data at small concentrations, $C_p<1$.

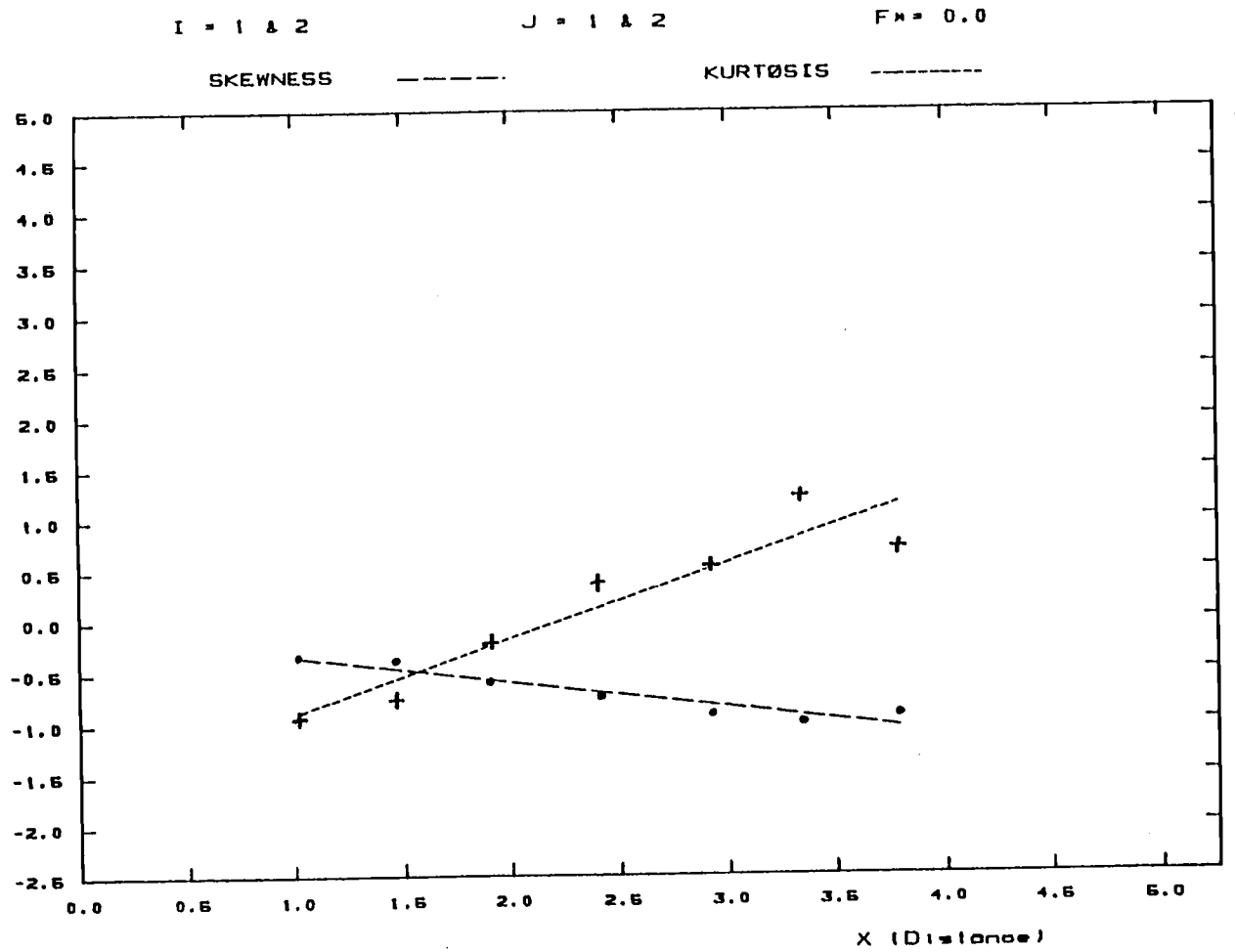

Figure 8 (a) : Skewness and kurtosis values of the experimental data versus downwind distance X for $F_*=0$ case for $|y|<\sigma_y$ and $0<h<0.4$. Broken lines are the linear regression lines.

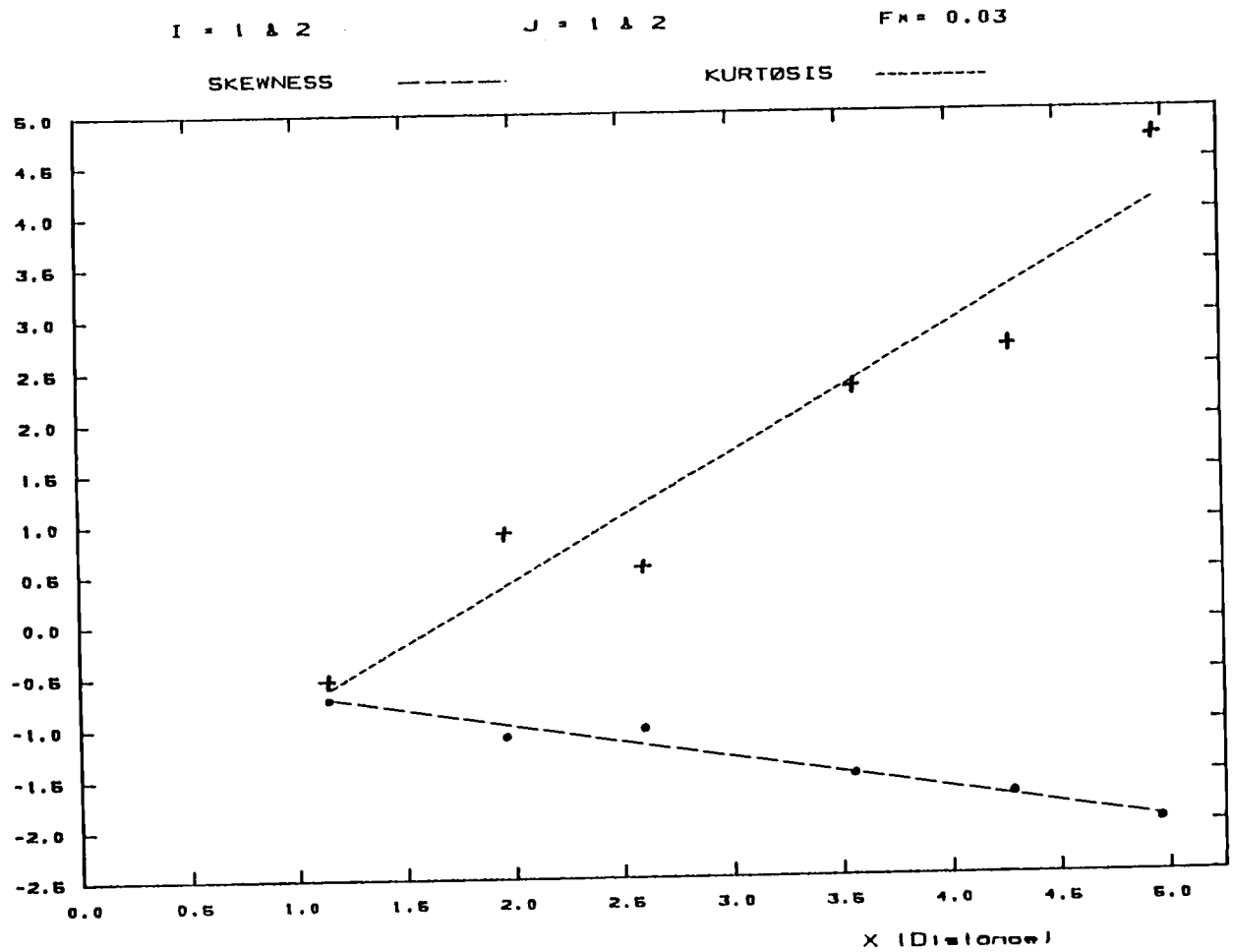

Figure 8 (b) : Skewness and kurtosis of the experimental data versus downwind distance X for F_*=0.03 case for $|y|<\sigma_y$ and $0<h<0.4$. Broken lines are the linear regression lines.

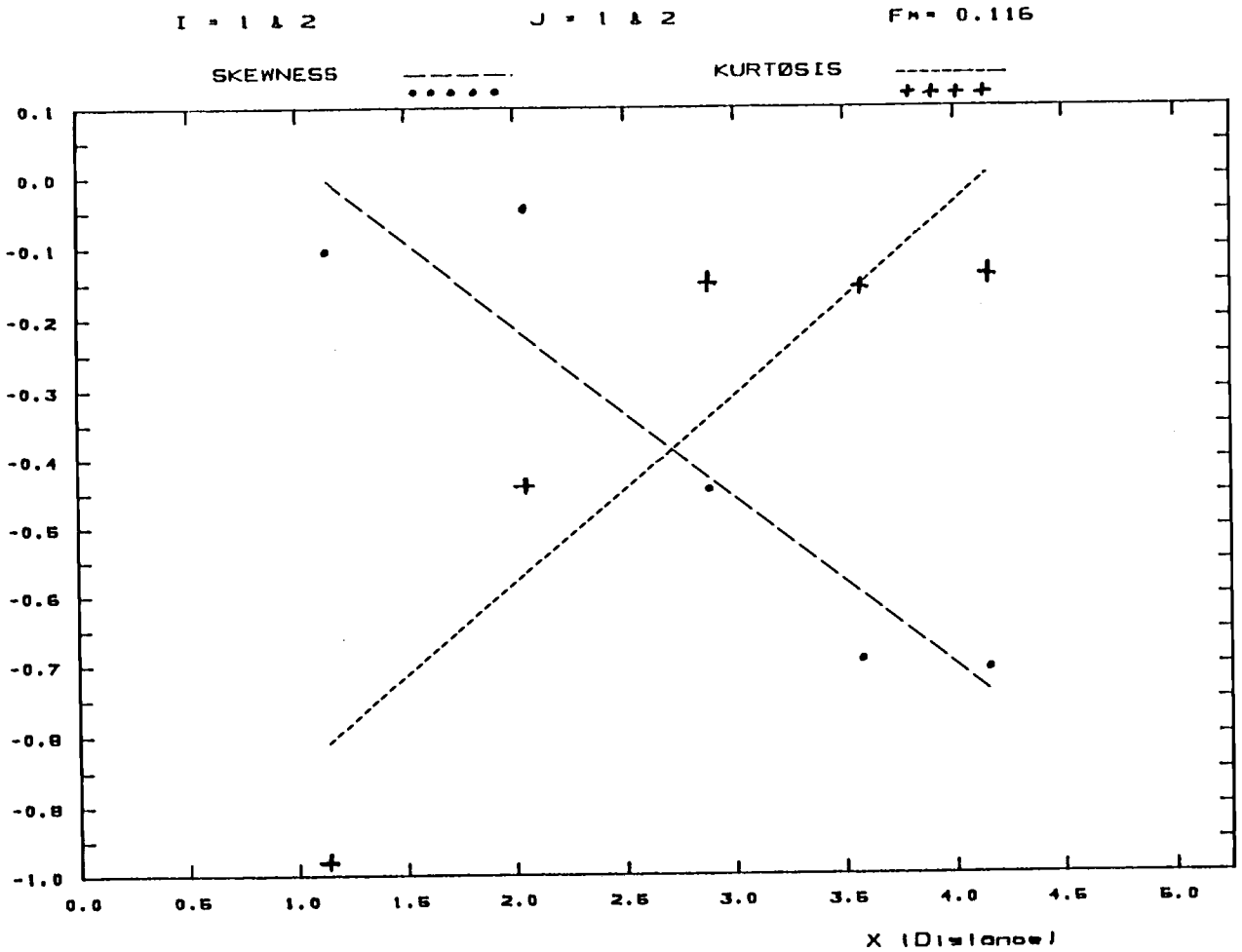

Figure 8 (c) : Skewness and kurtosis values of the experimental data versus the downwind distance X for F_*=0.116 case for $|y|<\sigma_y$ and 0<h<0.4. Broken lines are the linear regression lines.

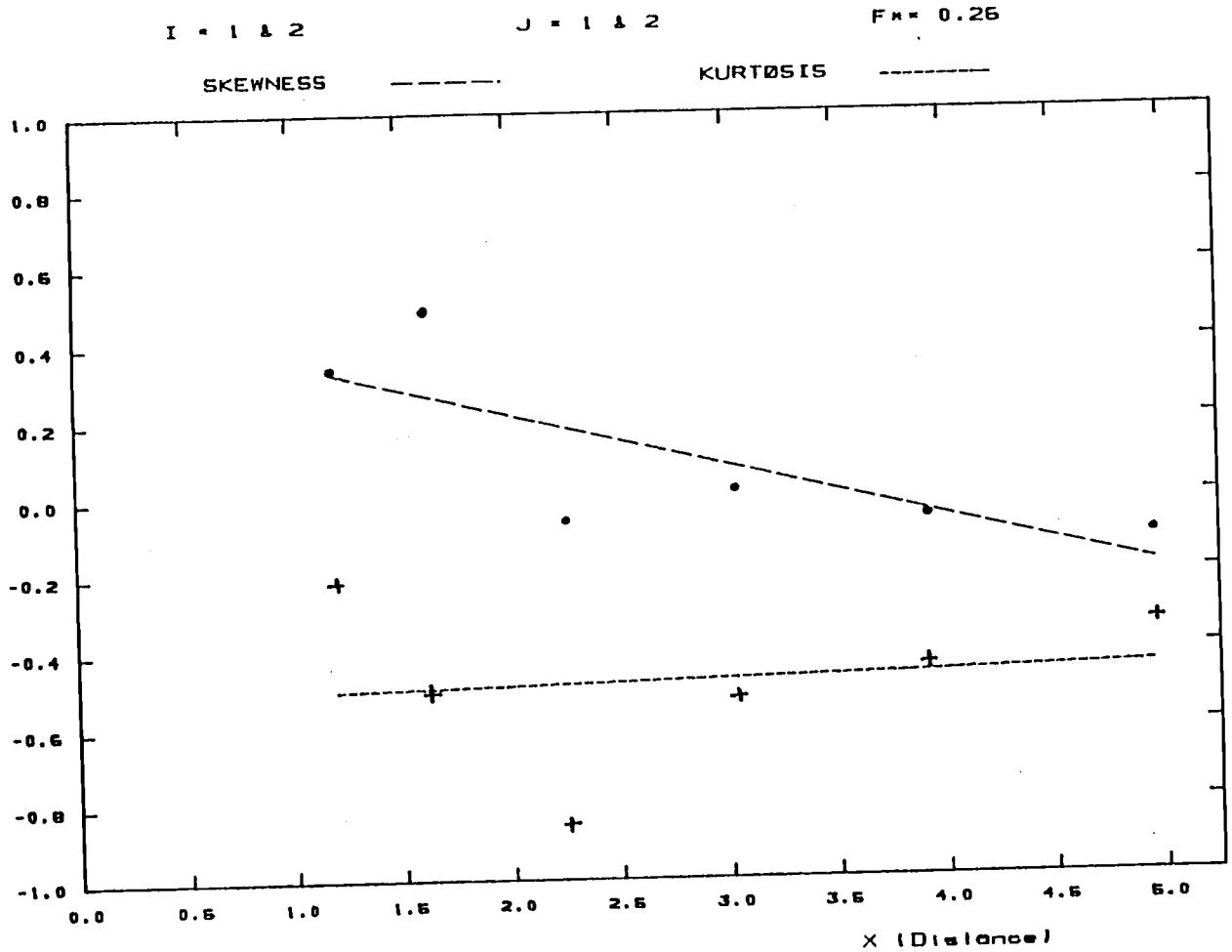

Figure 8 (d) : Skewness and kurtosis values of the experimental data versus the downwind distance X for F_*=0.26 case for $|y|<\sigma_y$ and $0<h<0.4$. Broken lines are the linear regression lines.

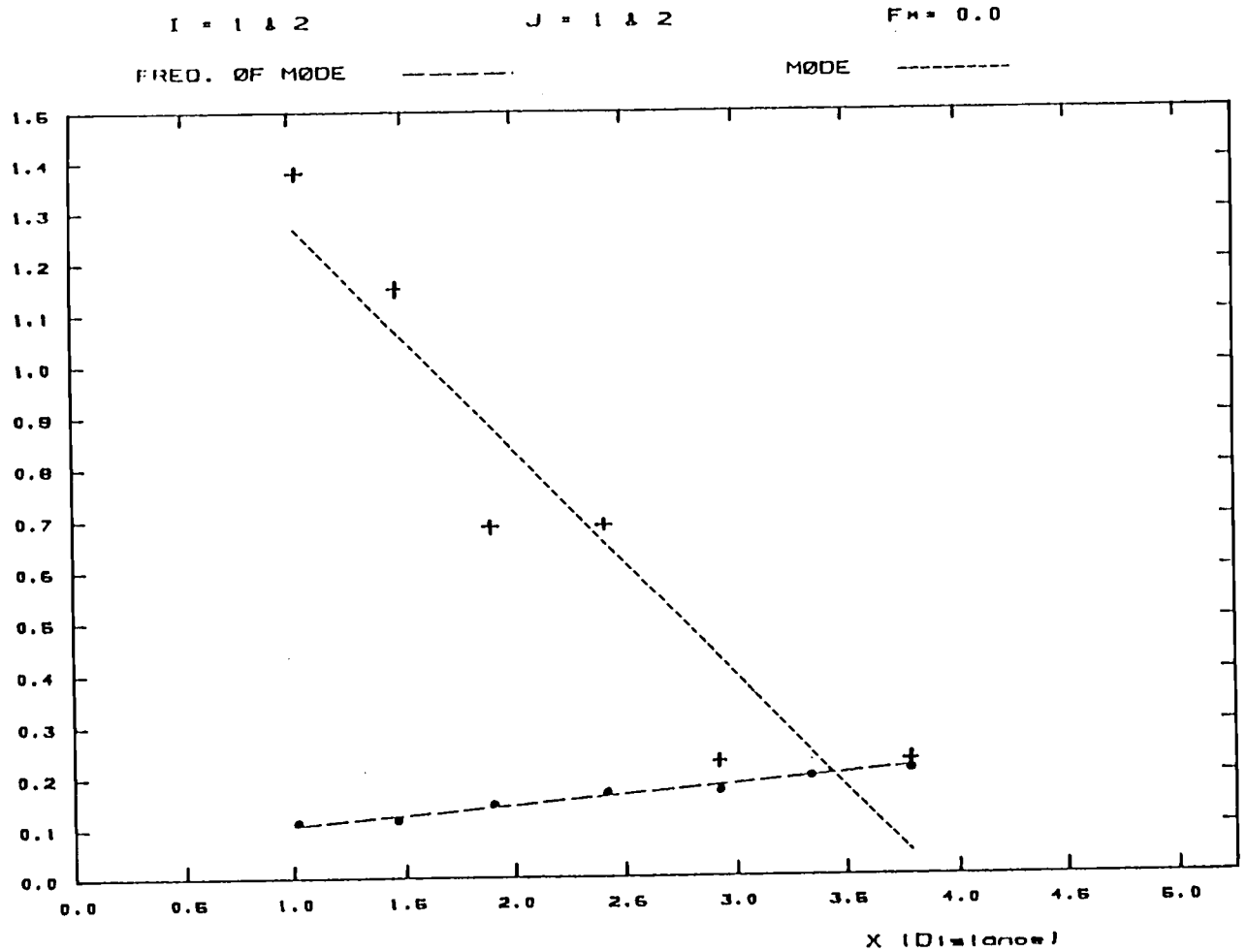

Figure 9 (a) : The mode and its frequency for the experimental data versus the downwind distance X for F_*=0 case for $|y|<\sigma_y$ and 0<h<0.4. Broken lines are the linear regression lines.

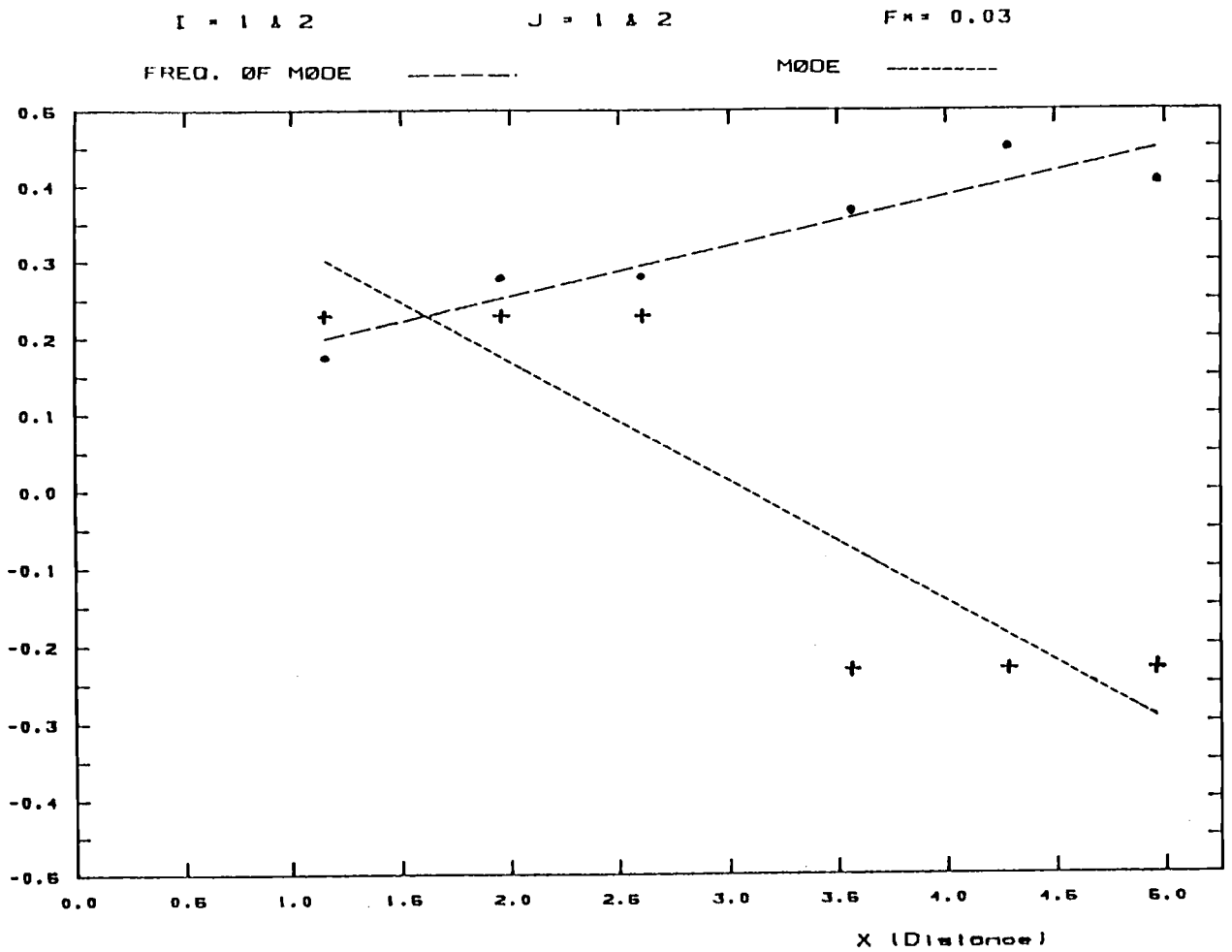

Figure 9 (b) : The mode and its frequency for the experimental data versus the downwind distance X for F_*=0.03 case for $|y|<\sigma_y$ and 0<h<0.4. Broken lines are the linear regression lines.

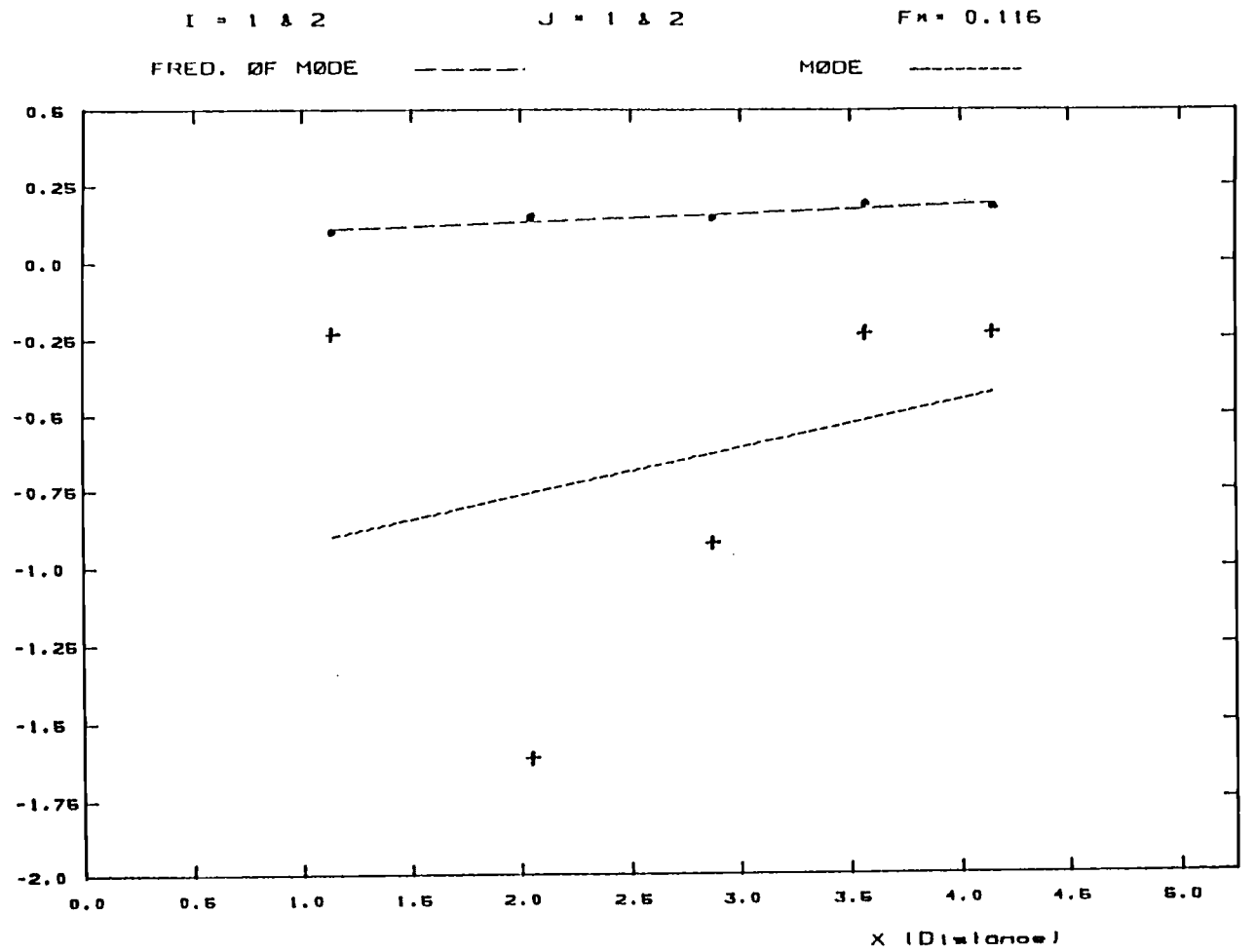

Figure 9 (c) : The mode and its frequency for the experimental data versus the downwind distance X for F_*=0.116 case for $|y|<\sigma_y$ and $0<h<0.4$. Broken lines are the linear regression lines.

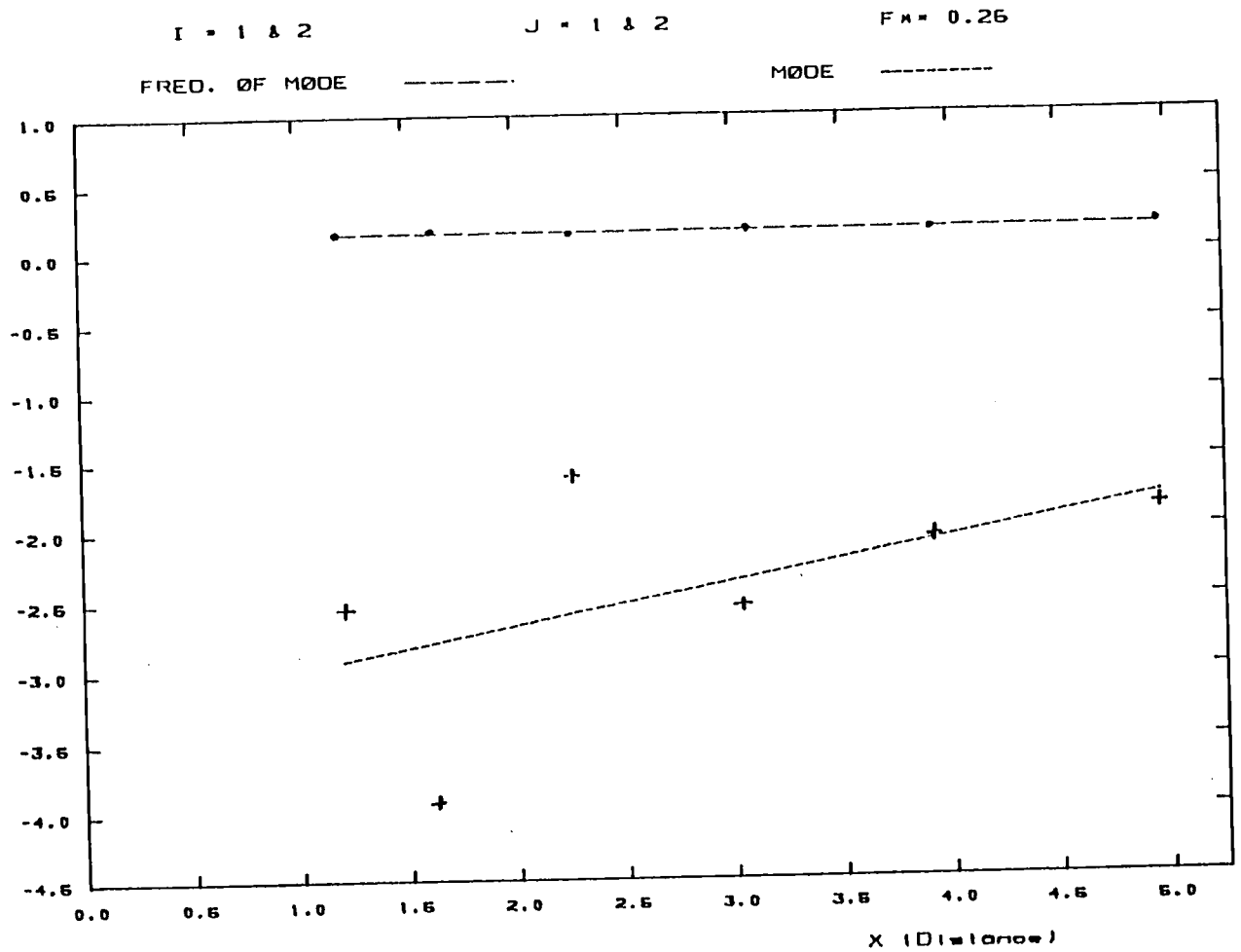

Figure 9 (d) : The mode and its frequency for the experimental data versus the downwind distance X for F_*=0.26 case for $|y|<\sigma_y$ and $0<h<0.4$. Broken lines are the linear regression lines.

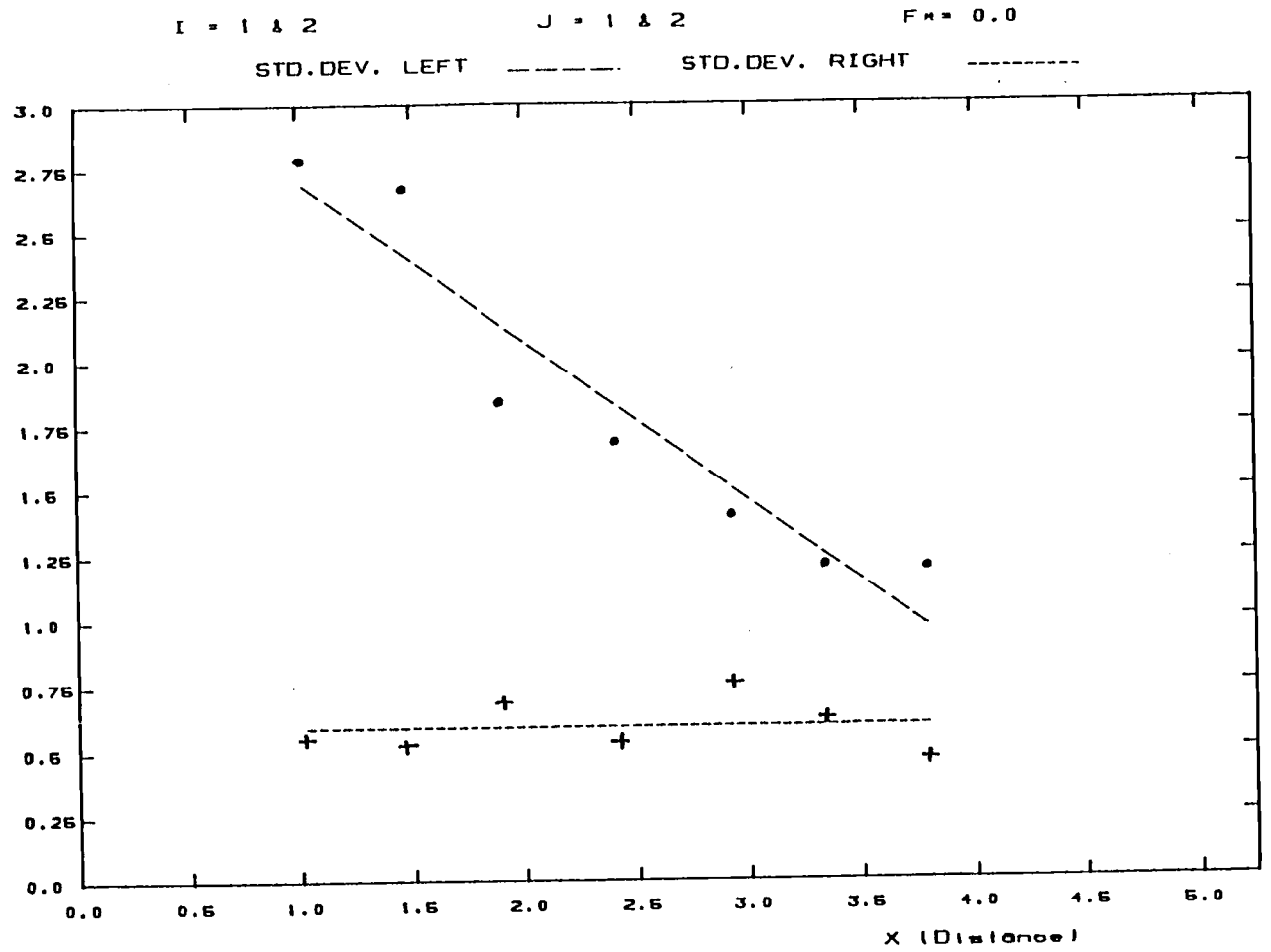

Figure 10 (a) : Both left and right standard deviations of the experimental data versus distance X for $F_*=0$ case for $|y|<\sigma_y$ and $0<h<0.4$. Broken lines are the linear regression lines.

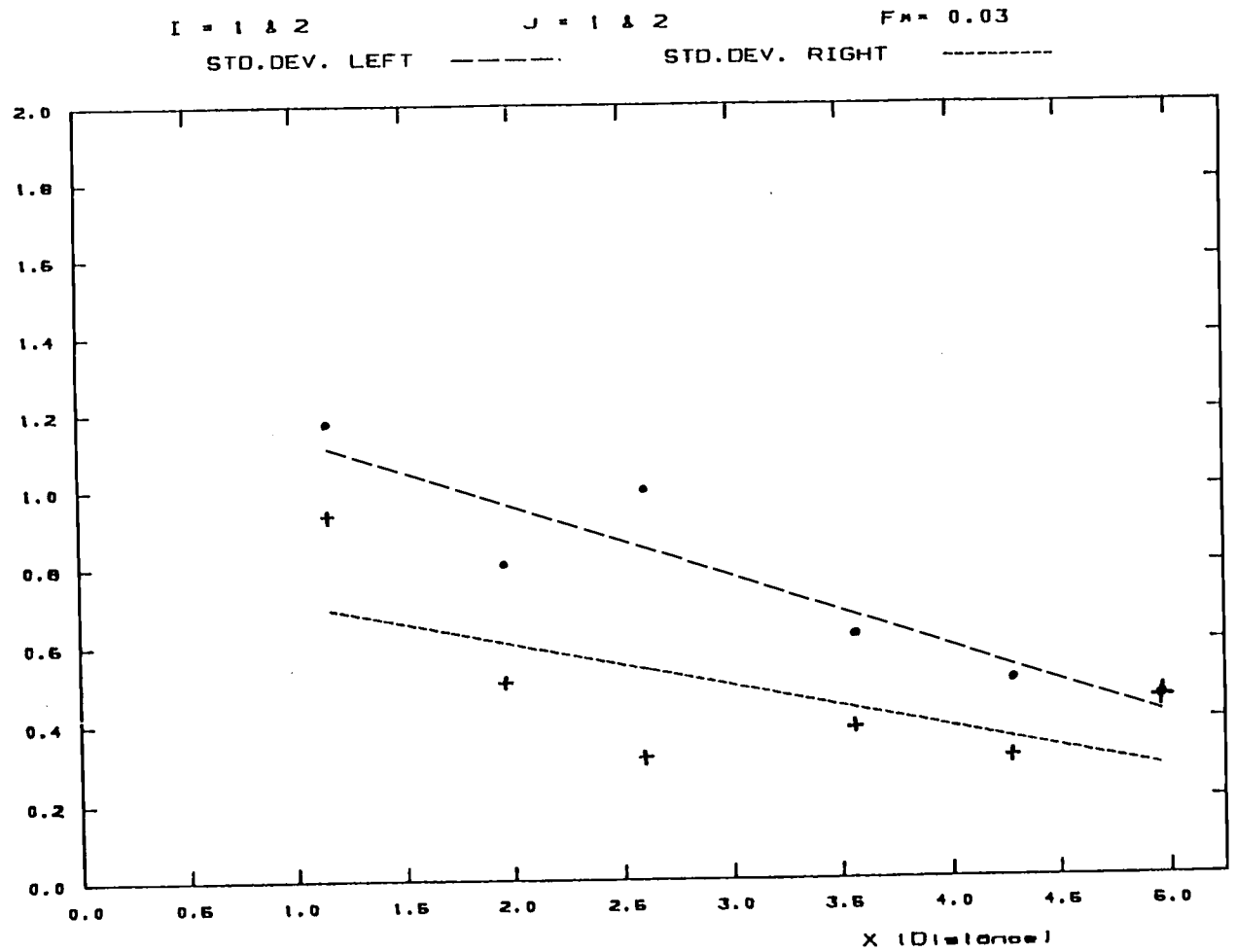

Figure 10 (b) : Both left and right standard deviations of the experimental data versus distance X for F_*=0.03 case for $|y|<\sigma_y$ and 0<h<0.4. Broken lines are the linear regression lines.

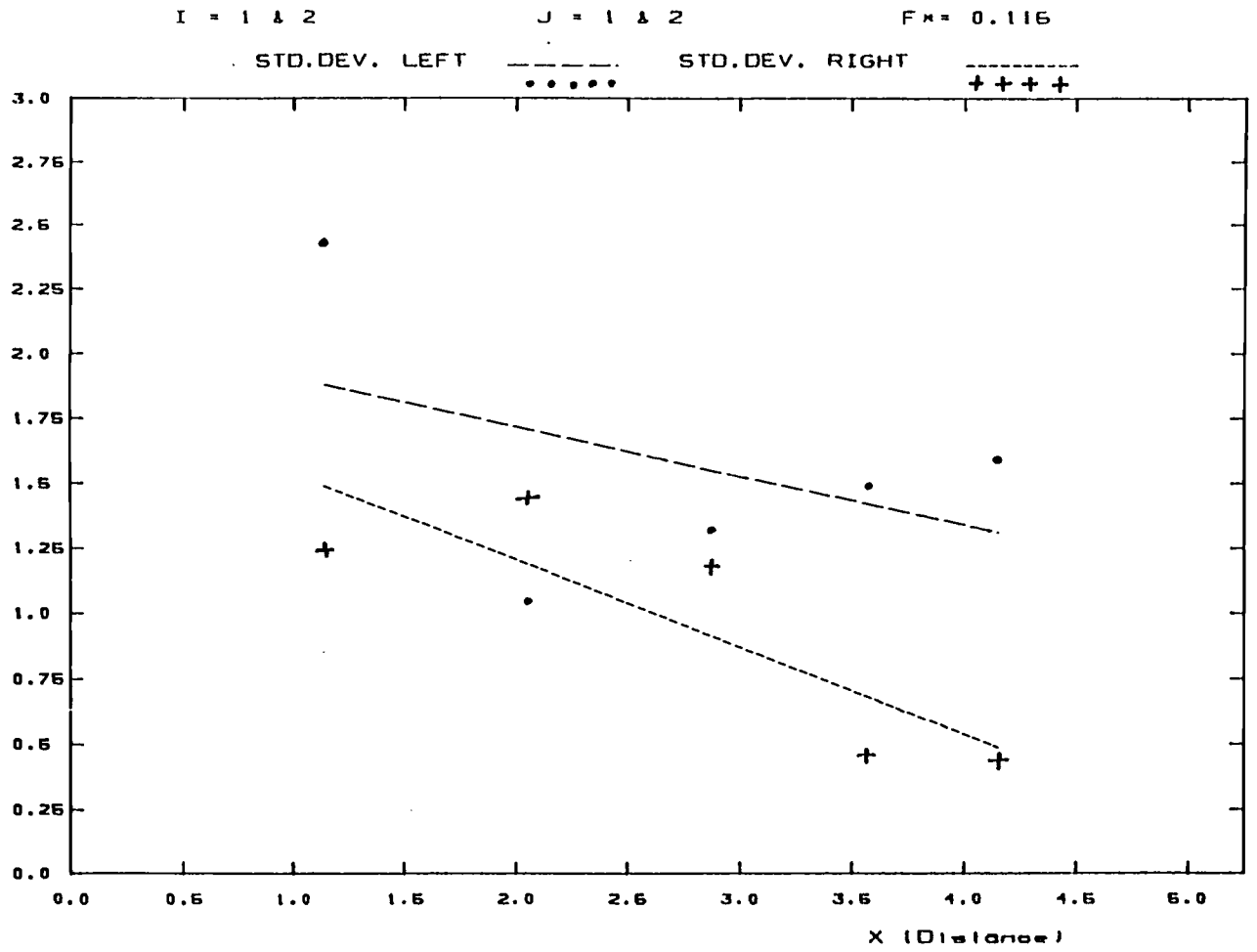

Figure 10 (c) : Both left and right standard deviation of the experimental data versus distance X for F_*=0.116 case for $|y|<\sigma_y$ and $0<h<0.4$. Broken lines are the linear regression lines,

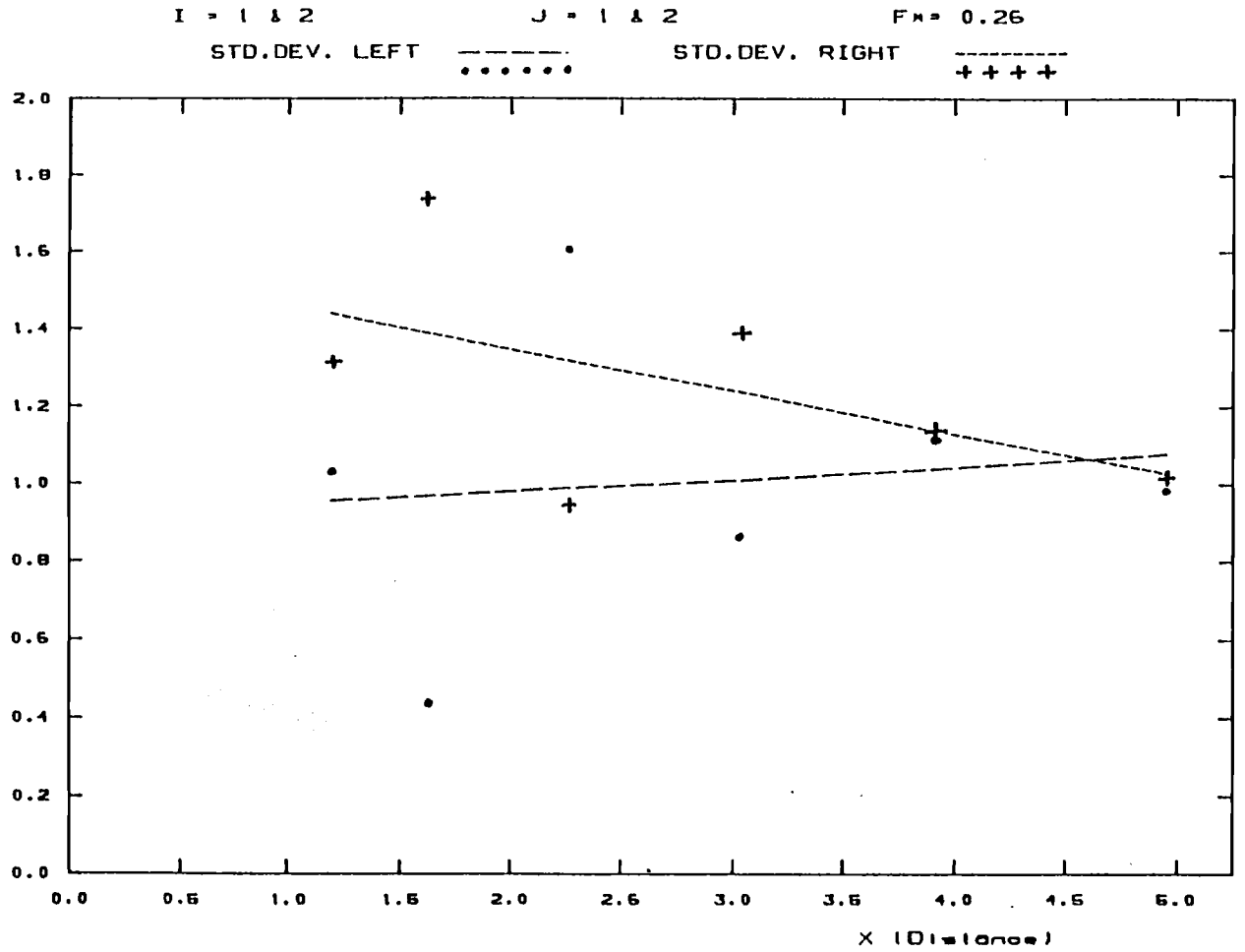

Figure 10 (d) : Both left and right standard deviations of the experimental data versus distance X for F_x=0.26 case for $|y|<\sigma_y$ and 0<h<0.4. Broken lines are the lineat regression lines.

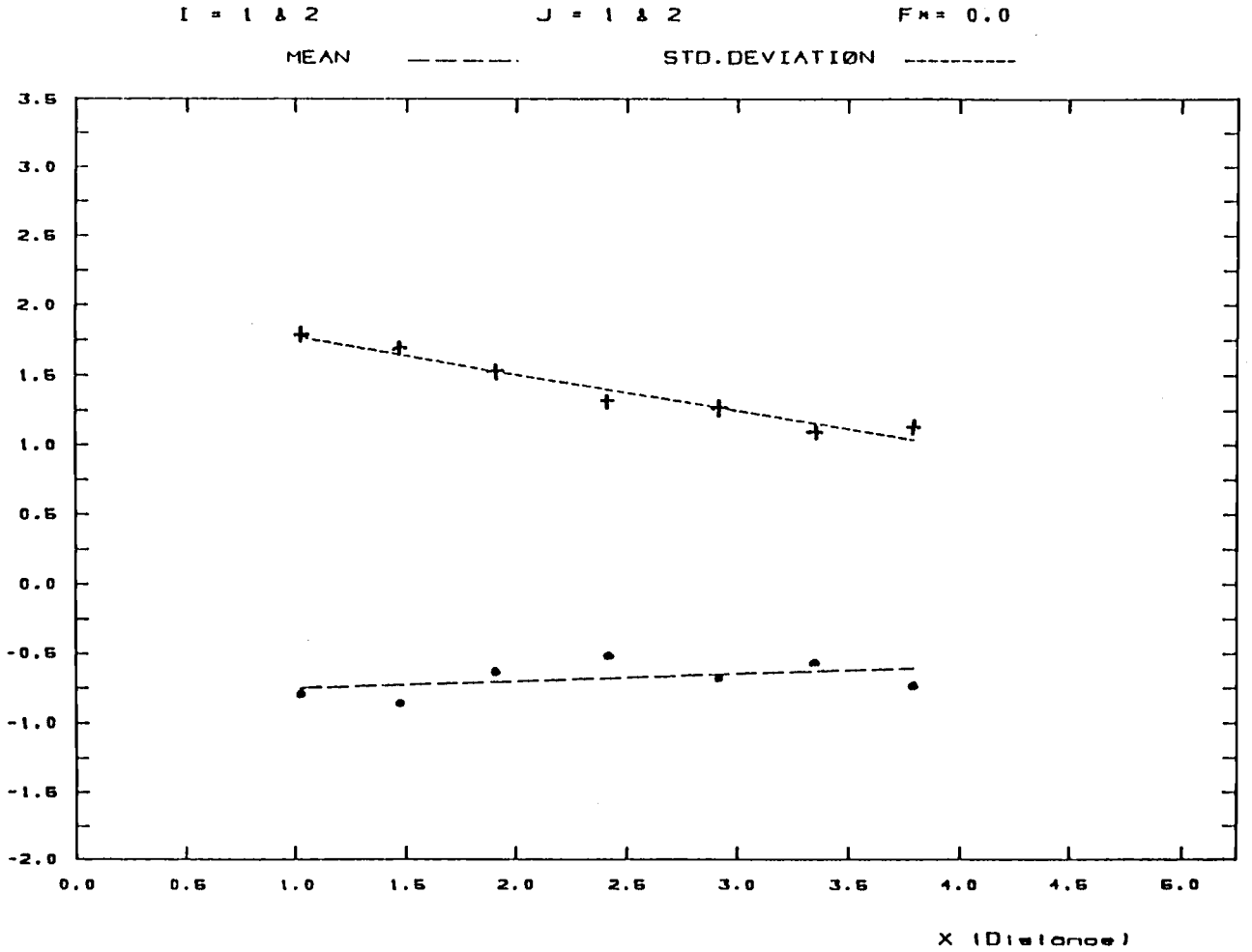

Figure.11 (a) : The mean and the standard deviation of the experimental data versus distance X for F_x=0 case for $|y|<\sigma_y$ and 0<h<0.4. Broken lines are the linear regression lines.

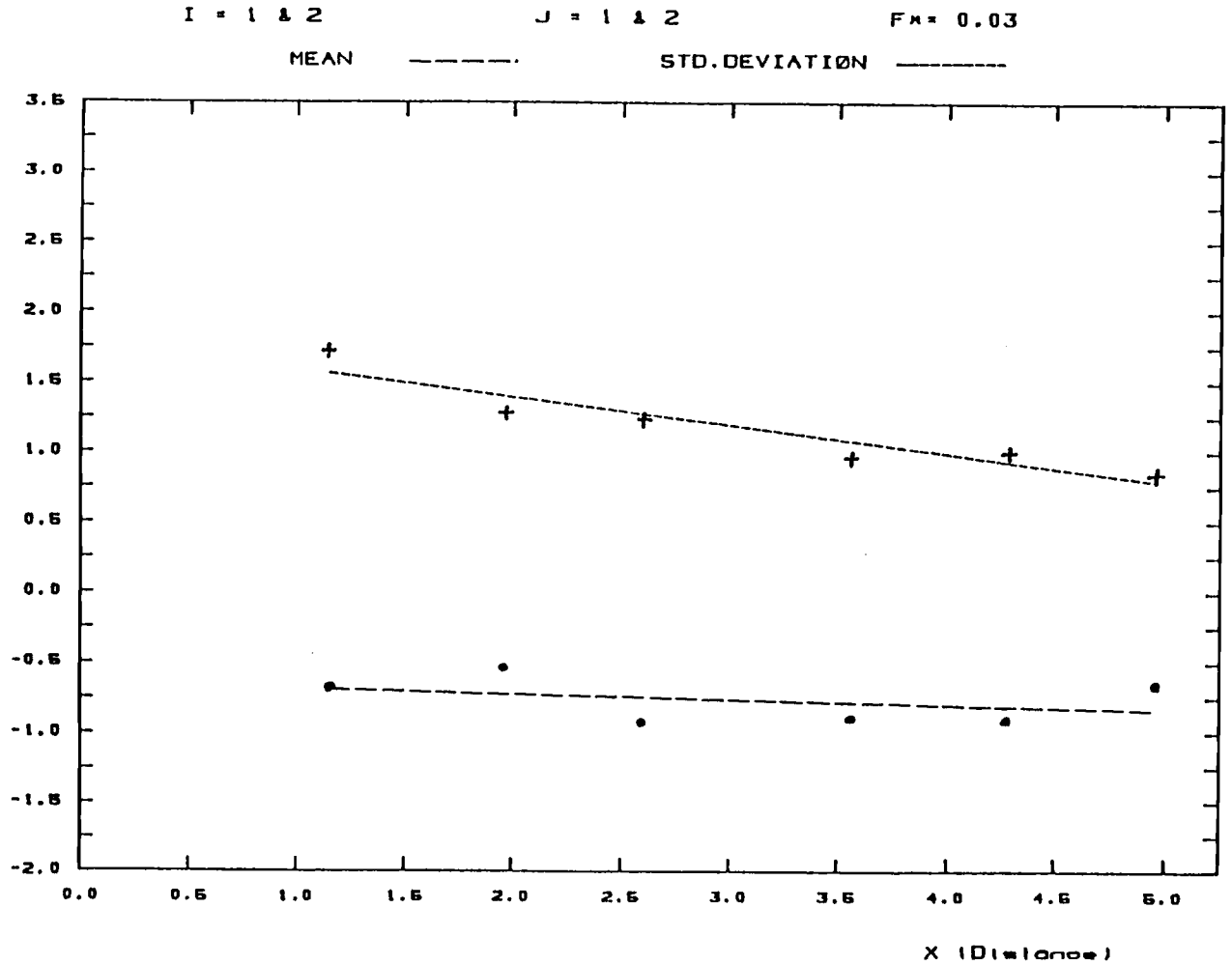

Figure 11 (b) : The mean and the standard deviation of the experimental data versus distance X for F_*=0.03 case for $|y|<\sigma_y$ and 0<h<0.4. Broken lines are the linear regression lines.

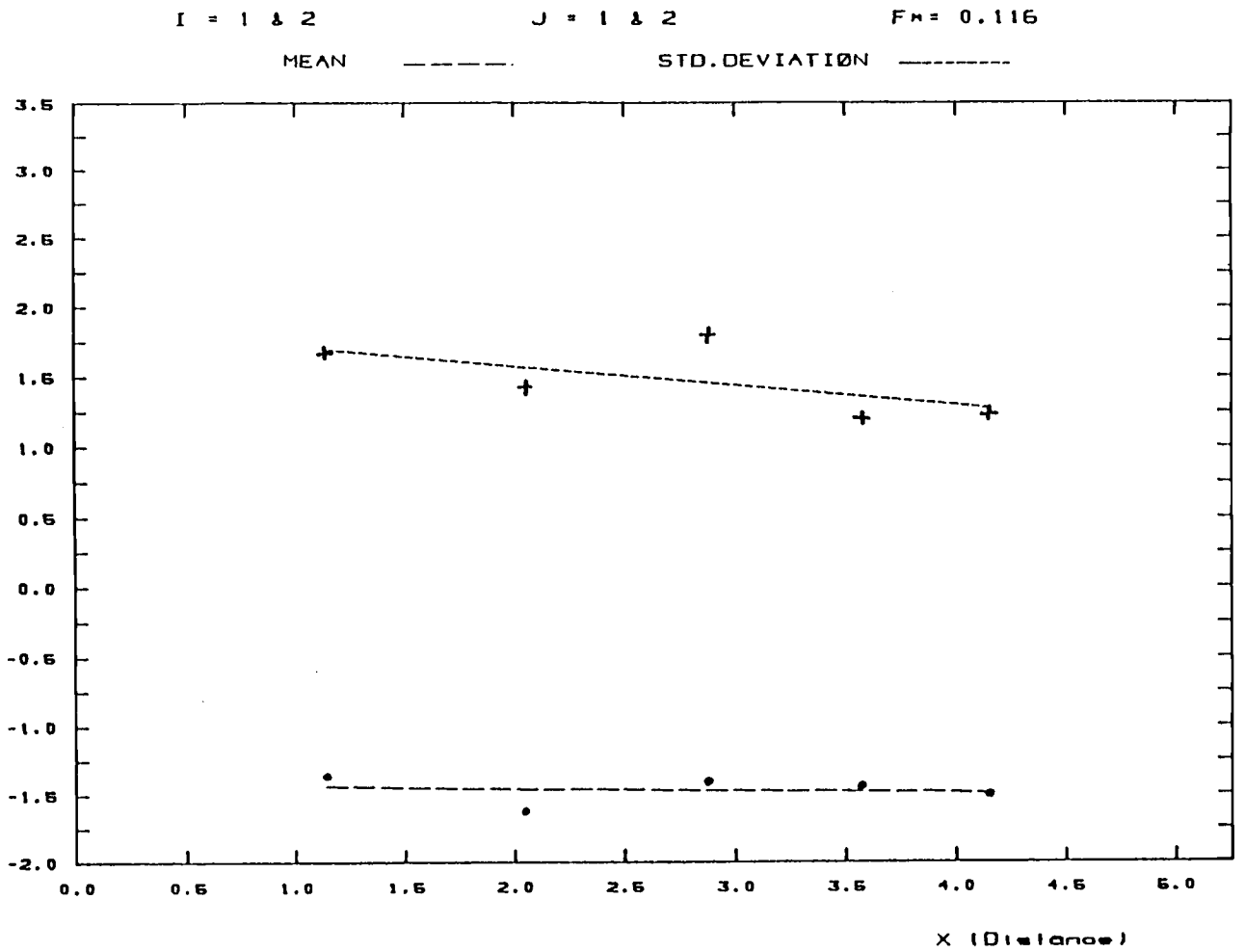

Figure 11 (c) : The mean and the standard deviation of the experimental data versus distance X for F_x=0.116 case for $|y|<\sigma_y$ and 0<h<0.4. Broken lines are the linear regression lines.

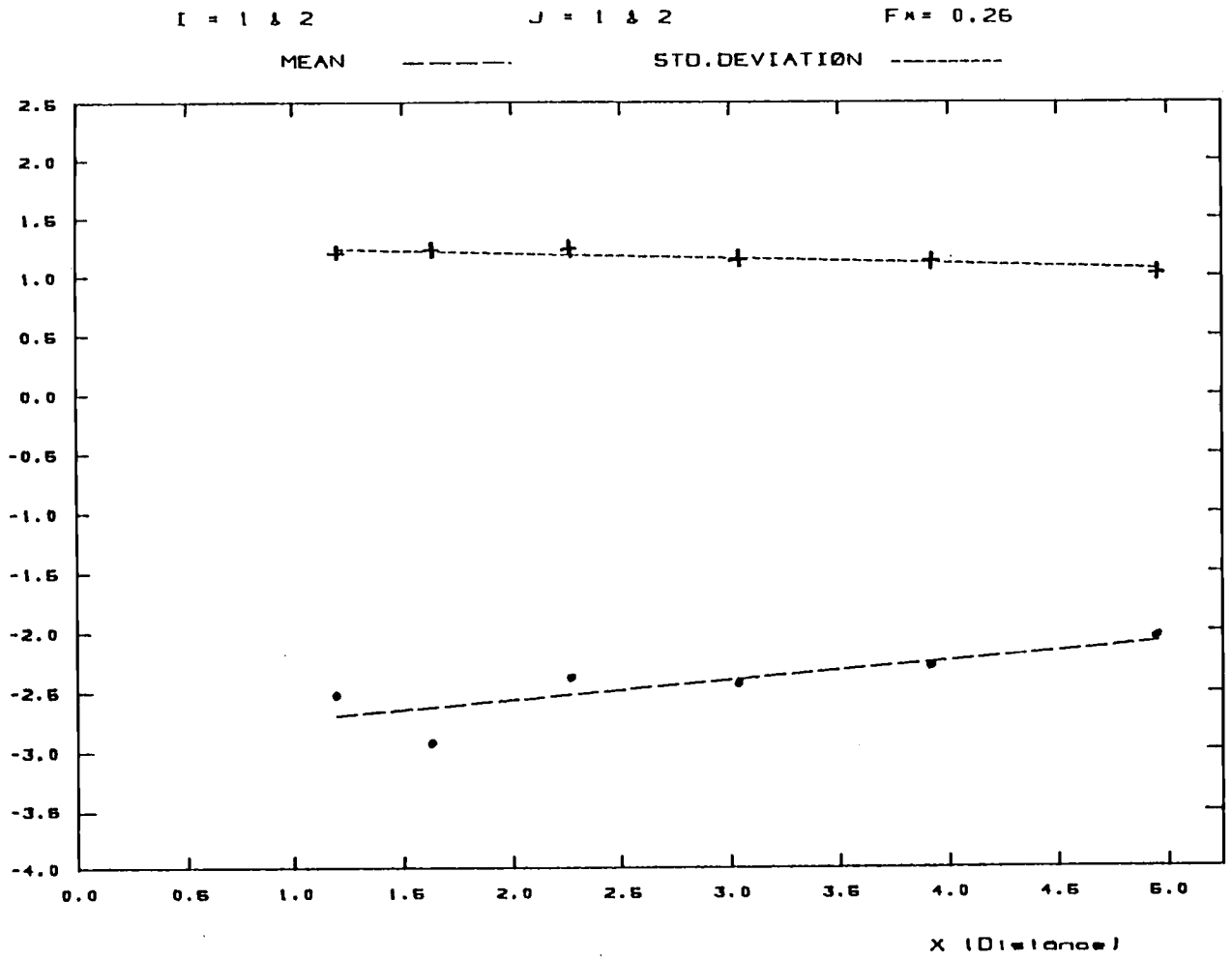

Figure 11 (d) : The mean and the standard deviation of the experimental data versus distance X for F_*=0.26 case for $|y|<\sigma_y$ and 0<h<0.4. Broken lines are the linear regression lines.

	Double-log-normal Distribution				Gamma Distribution				
Distance	Neg.	Pos.	Total	Error Square	Neg.	Pos.	Total	Error Square	N
1.02	-.108	.130	.238	.00500	-.129	.167	.295	.00797	148
1.46	-.097	.118	.215	.00402	-.082	.131	.213	.00389	226
1.90	-.101	.103	.204	.00388	-.032	.089	.121	.00198	263
2.41	-.089	.091	.180	.00556	-.051	.073	.124	.00231	285
2.92	-.097	.097	.195	.00410	-.051	.057	.108	.00114	301
3.34	-.099	.099	.199	.00466	-.035	.038	.073	.00079	241
3.79	-.123	.122	.245	.00870	-.067	.073	.140	.00338	286

Table 1 : Negative and positive differences between the expected and measured frequencies of concentrations, their sum of absolute values and total sum of square of error for the double-log-normal and the gamma distributions at F_*=0 (non-buoyant) case. N is the sample size.

	Double-log-normal Distribution				Gamma Distribution				
Distance	Neg.	Pos.	Total	Error Square	Neg.	Pos.	Total	Error Square	N
1.15	-.238	.239	.477	.01831	-.175	.187	.363	.01110	240
1.96	-.216	.215	.431	.02551	-.158	.163	.321	.01655	340
2.6	-.242	.241	.484	.03518	-.179	.181	.360	.02001	639
3.56	-.218	.217	.435	.02713	-.179	.179	.358	.02151	631
4.28	-.230	.230	.460	.05464	-.259	.259	.518	.04974	448
4.96	-.211	.211	.422	.02559	-.161	.161	.323	.02092	472

Table 2 : Negative and positive differences between the expected and measured frequencies of concentrations, their sum of absolute values and total sum of square of error for the double-log-normal and the gamma distributions at F_*=0.03 (weakly buoyant) case. N is the sample size.

	Double-log-normal Distribution				Gamma Distribution				
Distance	Neg.	Pos.	Total	Error Square	Neg.	Pos.	Total	Error Square	N
1.14	-.093	.134	.227	.00457	-.087	.176	.263	.00632	171
2.05	-.121	.122	.243	.00502	-.060	.220	.280	.01017	333
2.87	-.119	.122	.241	.00633	-.091	.213	.304	.01092	833
3.57	-.125	.128	.252	.00709	-.096	.110	.206	.00457	806
4.15	-.102	.105	.207	.00632	-.076	.116	.192	.00470	980

Table 3 : Negative and positive differences between the expected and measured frequencies of concentrations, their sum of absolute values and total sum of square of error for the double-log-normal and the gamma distributions at F_*=0.116 (buoyant) case. N is the sample size.

	Double-log-normal Distribution				Gamma Distribution				
Distance	Neg.	Pos.	Total	Error Square	Neg.	Pos.	Total	Error Square	N
1.19	-.092	.093	.185	.00667	-.066	.353	.419	.02287	69
1.62	-.133	.139	.273	.01112	-.046	.373	.419	.02547	154
2.26	-.089	.117	.206	.00504	-.015	.142	.157	.00353	209
3.03	-.101	.104	.205	.00510	-.027	.179	.206	.00782	387
3.91	-.081	.088	.169	.00336	-.033	.157	.190	.00695	519
4.96	-.091	.092	.183	.00420	-.073	.140	.214	.00738	577

Table 4 : Negative and positive differences between the expected and measured frequencies of concentrations, their sum of absolute values and total sum of square of error for the double-log-normal and the gamma distributions at F_*=0.26 (highly buoyant) case. N is the sample size.

	Chi-square values	
Distance	D. L. N.	Gamma
1.02	11.4	12.0
1.46	15.1	11.9
1.90	36.6	4.53
2.413	18.4	3.82
2.92	57.8	8.35
3.34	25.7	2.42
3.79	65.0	7.7

Table 5

	Chi-square values	
Distance	D. L. N.	Gamma
1.15	401.0	52.6
1.96	255.0	44.7
2.6	1160.0	129.0
3.56	1310.0	146.0
4.28	405.0	218.0
4.96	756.0	173.0

Table 6

	Chi-square values	
Distance	D. L. N.	Gamma
1.14	8.5	11.4
2.05	86.3	28.1
2.87	237.0	56.8
3.57	178.0	28.3
4.15	190.0	18.1

Table 7

	Chi-square values	
Distance	D. L. N.	Gamma
1.19	3.21	14.7
1.62	18.0	26.6
2.26	16.1	7.4
3.03	69.5	22.7
3.91	44.0	20.8
4.96	52.8	17.6

Table 8

Tables 5-8 : The chi-square values for the double-log-normal and the gamma distributions for different buoyancies. Table 5 is for F_*=0 (non-buoyant), table 6 for F_*=0.03 (weakly buoyant), table 7 for F_*=0.116 (buoyant) and table 8 for F_*=0.26 (highly buoyant) plume cases.

12. REFERENCES

Abramowitz, M. and I.A. Stegun, 1964: Handbook of Mathematical Functions. National Bureau of Standards. Applied Mathematics Series, 55.

Barry, P.J., 1977: Stochastic properties of atmospheric diffusivity. Sulphur and its Inorganic Derivatives in the Canadian Environment. National Research Council of Canada, 313-358.

Benarie, M.M., 1974: The use of the relationship between wind velocity and ambient pollutant concentration distributions for the estimation of average concentrations from gross meteorological data. Environmental Protection Agency Pub.No. 650/4-74-038.

Bencala, K.E. and J.H. Seinfeld, 1976: On frequency distributions of air pollutant concentrations. Atmospheric Environment, 10, 941-950.

Berger, A., J.L. Melice and Cl. Demuth, 1982: Statistical distributions of daily and high atmospheric SO_2 concentrations. Atmospheric Environment, 16, 2863-2877.

Cats, G.J. and A.A.M. Holtslag, 1980: Prediction of air pollution frequency distribution, part-I: The log-normal model. Atmospheric Environment, 14, 255-258.

Deardorff, J.W., 1985: Laboratory experiments on diffusion: the use of convective mixed layer scaling. J. Climate and Appl. Meteor., 24, 1143-1151.

Deardorff, J.W. and G.E. Willis, 1984: Ground level concentration fluctuations from a buoyant and a non-buoyant source within a laboratory convectively mixed layer. Atmospheric Environment, 18, 1297-1309.

Deardorff, J.W. and G.E. Willis, 1985: Further results from a laboratory model of the convective planetary boundary layer. Boundary-Layer Meteor., 32, 205-236.

Deardorff, J.W. and G.E. Willis, 1987: Concentration fluctuations within a laboratory convectively mixed layer. To appear in Atmospheric Environment.

Fackrell, J.E. and A.G. Robins, 1982(a): Concentration fluctuations and fluxes in plumes from point sources in a turbulent boundary layer. J. Fluid Mech., 117, 1-26.

Fackrell, J.E. and A.G. Robins, 1982(b): The effects of source size on concentration fluctuations in plumes. Boundary-Layer Meteor., 22, 335-350.

Gad-el-Hak, M. and J.B. Morton, 1978: Experiments on the diffusion of smoke in isotropic turbulent flow. AIAA Journal, 17, No.6, 558-562.

Gifford, F.A., 1974: The form of the frequency distributions of air pollutant concentrations. Environmental Protection Agency Pub.No. 650/4-74-038.

Hanna, S.R., 1984: The exponential probability distribution function and concentration fluctuations in smoke plumes. Boundary-Layer Meteor., 29, 361-375.

Hunt, W.F., 1972: The precision associated with the sampling frequency of log-normally distributed air pollutant measurements. J. of Air Poll. Control Assoc., 22, 687-691.

Jeffrey, G.H. and F.B. Smith, 1974: The prediction of high concentrations of sulfur dioxide in London and Manchester air. Environmental Protection Agency Pub. No. 650/4-74-038.

Knox, J.B. and R. Lange, 1974: Surface air pollutant concentration frequency distributions: Implications for urban modelling. J. of Air Poll. Control Assoc., 24, 48-53.

Knox, J.B. and R.I. Pollack, 1974: An investigation of the frequency distribution of air pollutant concentrations. Environmental Protection Agency Pub. No. 650/4-74-038.

Larsen, R.I., 1969: A new mathematical model of air pollutant concentration, averaging time and frequency. J. of Air Poll. Control Assoc., 19, 24-30.

Lindgren, B.W., 1962: Statistical Theory. Department of Math. Institute of Tech. University of Minnesota. 427 pp.

Lynn, D.A., 1974: Fitting curves to urban suspended particulate data. Environmental Protection Agency Pub. No. 650/4-74-038.

Nieuwstadt, F.T.M., 1980: Application of mixed layer similarity to the observed dispersion from a ground-level source. J. Appl. Meteor., 19, 157-162.

Ott, W.R. and D.T. Mage, 1976: A general purpose univariate probability model for environmental data analysis. Comput. and Ops. Res., 3, 209-216.

Sawford, B.L., 1983: The effect of Gaussian particle-pair distribution functions in the statistical theory of concentration fluctuations in homogeneous turbulence. Quart. J. of R. Meteor. Soc., 109, 339-354.

Shoji, H. and T. Tsukatani, 1973: Statistical model of air pollutant concentration and its application to the air quality standards. Atmospheric Environment, 7, 485-501.

Simpson, R.W., N.J. Daly and A.J. Jakeman, 1983: The prediction of maximum air pollution concentrations for TSP and CO using Larsen's model and the ATDL model. Atmospheric Environment, 14, 255-258.

Venkatram, A., 1982: A framework for evaluating air quality models. Boundary-Layer Meteor., 24, 371-385.

Willis, G.E. and J.W. Deardorff, 1974: A laboratory model of the unstable planetary boundary layer. J. Atmos. Sci., 31, 1297-1307.

Willis, G.E. and J.W. Deardorff, 1976: A laboratory model of diffusion into the convective planetary boundary layer. Quart. J. of R. Meteor. Soc., 102, 427-445.

Willis, G.E. and J.W. Deardorff, 1981: A laboratory study of dispersion from a source in the middle of the convectively mixed layer. Atmospheric Environment, 15, 109-117.

Willis, G.E. and J.W. Deardorff, 1983: On plume rise within a convective boundary layer. Atmospheric Environment, 17, 2435-2447.

Wilson, D.J., 1982: Predicting risk of exposure to peak concentrations in fluctuating plumes. Alberta Environment Technical Report, Dec. 1982, 90 pp.

Wilson, D.J. and B.W. Simms, 1985: Exposure time effects on concentration fluctuations in plumes. Dept. of Mechanical Engineering, University of Alberta, Edmonton, Canada. Report No. 47, 163 pp.

Zimmer, C.E. and R.I. Larsen, 1965: Calculating air quality and its control. J. of Air Poll. Control Assoc., 15, 565-572.