APPLICATION OF THE INVERSE GAUSSIAN DISTRIBUTION TO REGIONAL FLOW ANALYSIS FOR THE ISLAND OF NEWFOUNDLAND

SUELYNN ELIZABETH DIGNARD

Your file Votre référence
ISBN: 0-612-89620-X
Our file Notre référence
ISBN: 0-612-89620-X

APPLICATION OF THE INVERSE GAUSSIAN DISTRIBUTION TO REGIONAL FLOW ANALYSIS FOR THE ISLAND OF NEWFOUNDLAND

Prepared by

©Suelynn Elizabeth Dignard, B. Eng.

A Thesis submitted to the School of Graduate Studies
in partial fulfillment of the requirements for the degree
of Master's of Engineering

Faculty of Engineering and Applied Science
Memorial University of Newfoundland

April, 2003

St. John's **Newfoundland** **Canada**

Abstract

This thesis analyzes how the Inverse Gaussian distribution compares to other statistical distributions for fitting flow data from Newfoundland rivers. Although the 2-parameter Inverse Gaussian distribution is as flexible as most 3-parameter distributions, it is not currently used for statistical analysis of hydrologic data. Theories are available for regression analysis as well as hypothesis testing that is based on the Inverse Gaussian assumption.

A hydrological study of Newfoundland rivers was performed to assess application of the Inverse Gaussian distribution to flow and regional analysis techniques. For high flows, the Inverse Gaussian distribution was compared to the Generalized Extreme Value (GEV), the 3-Parameter Lognormal (3PLN), the Extreme Value (EV) and the Lognormal (LN) distributions. For low flows, comparison was made to the 3-Parameter Weibull (W3), 2-Parameter Weibull (W2), Extreme Value (EV) and the Lognormal (LN) distributions.

The analysis confirmed the Inverse Gaussian distribution is a suitable candidate for flood analysis of high flows. However, the distribution does not perform well for low flow analysis. Using the Akaike Information Criterion (AIC) as an indicator of suitability, the

Inverse Gaussian distribution is significantly better than both the 2-parameter and 3-parameter distributions considered in this analysis for high flows. The results of the study include regional flood frequency curves based on the Inverse Gaussian distribution for two distinct regions within the island of Newfoundland. These curves are suitable for use in addition to and for comparison with other regional flood analysis techniques for Newfoundland streamflow data.

There are two main recommendations derived from this study. The first is to apply the Inverse Gaussian distribution to streamflow data in areas other than the island of Newfoundland. Such application of the distribution to both high and low flows of other areas would determine if the suitability for high flows, or the inappropriateness for low flows, is limited to Newfoundland streamflow data. The second main recommendation is to develop an approximation for the inverse of the Inverse Gaussian distribution, similar to the approximation for the more popular Gaussian, or Normal distribution. Such an approximation would enhance the use and capabilities of the Inverse Gaussian distribution in flood frequency analysis, as well as many other statistical applications where the inverse of this robust distribution is required.

Acknowledgements

My sincere thanks to my supervisor, Dr. Leonard Lye, for all his guidance, help, support and patience that led to the successful preparation of this thesis.

I am thankful to my fiancé, Bob Kirkland, for his love and support during the preparation of this thesis.

Thanks are due to the Faculty of Engineering and Applied Science, Memorial University of Newfoundland for financial support that was provided.

Table of Contents

Abstract ii
Acknowledgments iv
Table of Contents v
List of Appendices vii
List of Tables viii
List of Figures ix
List of Acronyms x

1.0 Introduction 1

2.0 The Inverse Gaussian Distribution 5
2.1 Background and Theory 5
2.2 Regionalisation Using the Inverse Gaussian Distribution 9
2.2.1 Test for Homogeneity 10
2.2.2 Regional Flow Frequency Analysis using the Index Flow Method 14
2.2.3 Regional Estimate of Q_{avg} 15

3.0 Data Preparation and Single Site Analysis Technique 17
3.1 Study Area 17
3.2 Data Preparation 22
3.3 Single Site Analysis 22
3.3.1 High Flow Analysis 23
3.3.2 Low Flow Analysis 28
3.3.3 Use of the Akaike Information Criterion Method for Comparing Performance of Statistical Distributions 31

4.0 Use of the Inverse Gaussian Distribution for Flow Estimation - Results and Discussion 33
4.1 Single Site Analysis Results 33
4.2 Regionalisation 40
4.2.1 Test for Homogeneity 41
4.2.2 Index Flow 44
4.2.3 Average Flow Estimate 51
4.3 Out of Sample Testing 56

5.0 Recommended Methodology and Application ... 64
5.1 Regionalisation Method Using the Inverse Gaussian Distribution ... 64
5.2 Application of Regionalisation Results ... 66

6.0 Conclusions and Recommendations ... 68
6.1 Conclusions ... 68
6.2 Recommendations ... 69

7.0 References ... 71

List of Appendices

Appendix A: Single Site Analysis AIC Results – High Flow Analysis 75

Appendix B: Single Site Analysis AIC Results – Low Flow Analysis 77

Appendix C: Single Site Analysis Results – High Flow Analysis Data Plots 89

Appendix D: Single Site Analysis Results – Low Flow Analysis Data Plots 138

List of Tables

Table 3-1 Gauging Stations Utilized in Analysis 19

Table 4-1 Summary Statistics of AIC Ranking by Distribution 34

Table 4-2 Frequency of AIC Rankings by Distribution 35

Table 4-3 Results of Regionalisation Technique 43

Table 4-4 Index Flow Flows 46

Table 4-5 Regional Flow Frequency Factors 48

Table 4-6 Inverse Gaussian Regional Scale Parameters 50

Table 4-7 Test Station Flow Estimates using the Proposed Regionalisation Technique 57

Table 4-8 Test Station Flow Estimates using Inverse Gaussian Single Site Analysis Technique 58

Table 4-9 Test Station Flow Estimates using the Top AIC Ranking Distribution ... 59

Table 4-10 Median Absolute Percent Difference between Regional Inverse Gaussian Technique and Top Single Site Analysis Distribution 60

Table 4-11 Median Absolute Percent Difference between Regional Inverse Gaussian Technique and Top Single Site Analysis, by Region 60

Table 4-12 Median Absolute Percentage Difference Between Frequency Analysis Estimates and Regression Equation Estimates for the Independent Data Sets (Government of Newfoundland and Labrador, 1999) 61

Table 4-13 Comparison of Flood Factors Results 62

List of Figures

Figure 2-1 Inverse Gaussian Distribution Probability Density Functions (Johnson and Kotz, 1970) 8

Figure 3-1 Location of Environment Canada Gauging Stations 21

Figure 4-1 Bar Chart of AIC Ranking for High Flow Analysis 36

Figure 4-2 Bar Chart of AIC Ranking for Low Flow Analysis 36

Figure 4-3 Index Flow Plot – North (Y) Region 47

Figure 4-4 Index Flow Plot – South (Z) Region 47

Figure 4-5 Regional Growth Curve - North (Y) and South (Z) Regions 50

Figure 4-6 Relationship of Mean Annual Daily Maximum Flows to Drainage Area – North (Y) Region 53

Figure 4-7 Relationship of Mean Annual Daily Maximum Flows to Drainage Area – South (Z) Region 53

Figure 4-8 Residual and Normal Plots for North (Y) Region Regression Residuals. 54

Figure 4-9 Residual and Normal Plots for South (Z) Region Regression Residuals. 55

List of Acronyms

AIC	Akaike Information Criterion
CFA	Consolidated Frequency Analysis
CI	Confidence Interval
DA	Drainage Area
GEV	Generalized Extreme Value
EV	Extreme Value Distribution
IG	Inverse Gaussian Distribution
LN	Log Normal Distribution
MAR	Mean Annual Runoff
MLE	Maximum Likelihood Estimate
Q_{avg}	Mean Annual Daily Extreme Flow
RFFA	Regional Flood Frequency Analysis
T	Return Period
EC	Environment Canada
W2	Weibull 2-Parameter Distribution
W3	Weibull 3-Parameter Distribution
Z	Standard Normal Variate
3PLN	3-Parameter Log Normal Distribution

1.0 Introduction

Environment Canada (EC) currently (January 2003) has 65 flow measuring stations on Newfoundland streams. However, this encompasses a very small portion of the stream network covering the island. The province of Newfoundland and Labrador lies between the 46^{th} and 61^{st} parallels with the bulk of the island portion of the province below the 50^{th} parallel. The land area of the Island of Newfoundland is 111 390 km^2 and is located in the Gulf of St. Lawrence. The topography of the island consists of a very rugged, rocky terrain. The mean annual runoff (MAR) ranges from approximately 700 – 900 mm in the north-central area of the island to 1300 – 2100 mm in the southwestern region of the island. This runoff flows through a vast network of streams, rivers, ponds and lakes that cover the Island of Newfoundland.

When an engineering project is undertaken one of the first required tasks is a flow analysis of the local watershed to determine the design flow. This is true for many engineering applications, including the design of roads, bridges and dams. If the engineering project is for water supply or power generation, a reliable estimate of the quantity of water available for consumption or power and energy production is essential. This is determined through detailed flow frequency analysis. If the river system has an Environment Canada (EC) flow gauging station, the flow frequency task is relatively easy, as there is direct data available for the analysis. However, for most new engineering projects on the Island of Newfoundland the streams in the study area are often not in the Environment Canada network of gauged streams. When data is not available for the stream in question nearby streams with similar characteristics are used to develop an appropriate data set for the ungauged stream. Regional flow frequency analysis for the Island of Newfoundland can provide a set of equations or flow factors from which the desired return period flows for an ungauged station are easily calculated and compared.

The Department of Environment and Lands has conducted a Regional Flood Frequency Analysis for the Island of Newfoundland (Government of Newfoundland and Labrador, 1999). This current regional flood analysis technique is based on annual maximum instantaneous flow data and uses a combination of two distributions; the General Extreme Value and the 3-Parameter Log-Normal. These distributions were used to develop a set of four regression equations for regional analysis of four distinct regions in Newfoundland. This study proposes the Inverse Gaussian distribution as an alternate single statistical distribution, using annual maximum daily flows.

For this study a flow analysis was conducted for the Island of Newfoundland. The analysis investigates application of the Inverse Gaussian distribution to both high and low Newfoundland streamflow data.

For high flows, comparison is made with the following statistical distributions:

3-Parameter Distributions

- Generalized Extreme Value (GEV)
- 3-Parameter Lognormal (3PLN)

2-Parameter Distributions

- Extreme Value (EV)
- Lognormal (LN) distributions.

For low flows, comparison is with:

3-Parameter Distributions

- 3-Parameter Weibull (W3)

2-Parameter Distributions

- 2-Parameter Weibull (W2)
- Extreme Value (EV)
- Lognormal (LN)

The contents of this thesis document are summarized by chapter as follows. Chapter 2 presents the theory of the Inverse Gaussian distribution along with the method of regionalisation using the Inverse Gaussian distribution and the expectations of the study. Chapter 3 summarizes the data preparation and preliminary analysis, while Chapter 4 presents the results and a discussion of the use of the Inverse Gaussian distribution for flow estimation. Chapter 5 discusses the methodology and application of the Inverse Gaussian distribution to flow frequency analysis in Newfoundland and Chapter 6 provides the conclusions and recommendations of the study. A list of all references used for this research is provided in Chapter 7.

2.0 The Inverse Gaussian Distribution

2.1 Background and Theory

The Inverse Gaussian distribution is also referred to as a first passage time distribution of Brownian motion with positive drift. The distribution is thought to have originated with Tweedie who laid the foundation research work for this distribution (Johnson & Kotz, 1970 and Chhikara & Folks 1989). The name 'Inverse Gaussian' was given to this distribution as a result of Tweedie noticing the inverse relation between the cumulant generating functions of these distributions and those of the Normal, or Gaussian, distribution. This distribution is also commonly referred to as the Wald distribution after Wald who also derived the same class of distribution (Johnson & Kotz, 1970 and Chhikara & Folks 1989).

Unlike the Normal density function, which requires a symmetrical distribution, the Inverse Gaussian density function represents a wide class of distributions, ranging from symmetrical to highly skewed distributions. In nature, observed data is rarely symmetrical, but most often skewed to some degree. Transformations are often made to asymmetrical data in an attempt to approximate the data set using the symmetrical

Normal density function. The Lognormal and the 3-Parameter Lognormal are examples of distributions based on transformations of the data. Where possible, it is more desirable to perform analysis on the acquired data set using appropriate distributions that are capable of approximating the original, untransformed data set. Chhikara and Folks (1989) believe that the Inverse Gaussian distribution, when appropriate, can be applied to meet the requirements for skewed data analysis.

The Inverse Gaussian probability function may be written as:

$$(1) \qquad f(x;\mu,\lambda) = \sqrt{\frac{\lambda}{2\pi}}x^{-3/2}\exp(-\frac{\lambda(x-\mu)^2}{2\mu^2 x}, x > 0, \mu > 0, \lambda > 0$$

where: μ = mean

$$\frac{1}{\lambda} = \frac{1}{n}\sum_{1}^{n}\left(\frac{1}{X_i} - \frac{1}{\overline{X}}\right) \quad \text{or} \quad \lambda = \frac{\mu^3}{Var(X)}$$

n = sample size

X = random variable

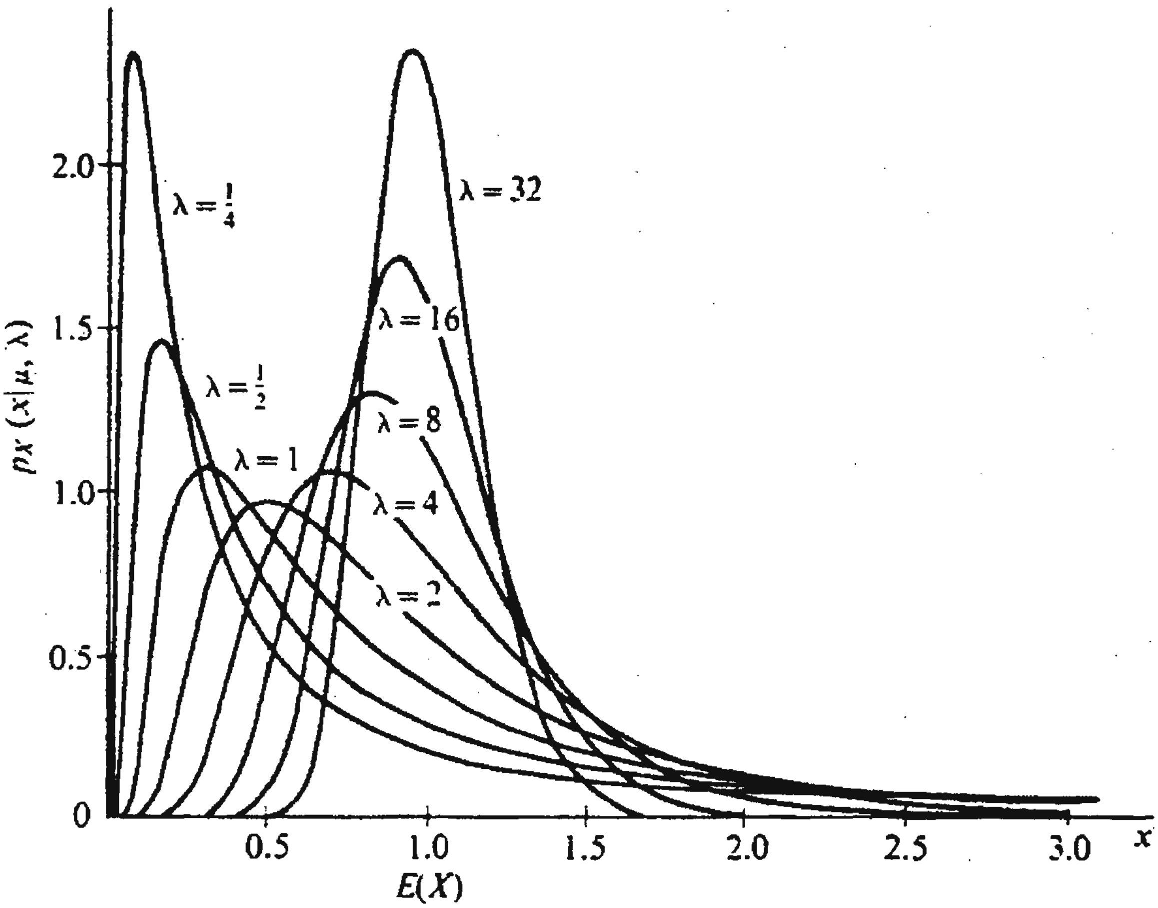
X_i = random sample

$\overline{X}$ = sample mean

The probability density functions of the Inverse Gaussian family are unimodal and skewed. The parameter μ is a measure of location, equaling the population mean E(X), and λ is a reciprocal measure of dispersion, equaling μ^3/Var(X) (Koziol, 1989). As discussed above, the Inverse Gaussian distribution density function represents a wide class of distributions, ranging from symmetrical to highly skewed as the shape factor (ϕ) ranges from 0 to infinity. Figure 2-1 (Johnson and Kotz, 1970) illustrates the wide range of shapes that the Inverse Gaussian distribution can represent. Using the two parameters of the distribution, the shape factor is calculated by $\phi = \mu / \lambda$. With $\mu = 1.0$ and λ ranging from 0.25 to 32, the resulting family of curves is shown.

Like the Normal distribution, the Inverse Gaussian distribution does not have a simple solution for the cumulative distribution function as integration of the probability density function is not possible. However, Koziol's (1989) Handbook of Percentage Points of the Inverse Gaussian Distribution contains tables pertaining to the cumulative distribution of the probability density function. These tables can be used to determine the random variate based on the distribution parameters. The tables are based on what Koziol (1989) describes as "remarkable that the Inverse Gaussian can be expressed in terms of the standard normal distribution function Φ by the following equation":

$$(2) \qquad F_X(x) = \Phi[(\lambda / x)^{1/2}(-1 + x / \mu)] + \exp(2\lambda / \mu)\Phi[-(\lambda / x)^{1/2}(1 + x / \mu)]$$



Inverse Gaussian Density Functions
$[E(x) = \mu = 1]$

Figure 2-1 **Inverse Gaussian Distribution Density Functions (Johnson and Kotz, 1970)**

The percentage points presented in Koziol (1989) are determined using an accurate algorithm for numerical calculation of the standard normal cumulative distribution combined with a variant of Newton's method. Software packages, such as BestFit (Palisade, 1993-96), are available to automate the tedious task of searching through tables of values for each percentile of interest. The software estimates the parameters of the distribution from the sample data, and provides estimates of the variate. The desired percentile can by entered into the software program, and an estimate of the resulting variate is determined. For the purposes of this study, the software is a much more efficient method of determining flow estimates using the Inverse Gaussian distribution.

2.2 Regionalisation Using the Inverse Gaussian Distribution

Although the 2-parameter distribution is as flexible as most 3-parameter distributions, it is not currently used for statistical analysis of hydrologic data. However, the Inverse Gaussian distribution is ready to apply as theories are available for regression analysis as well as hypothesis testing that is based on the Inverse Gaussian assumption. As well, the two parameters required are very easily calculated using the Maximum Likelihood Estimators (MLE) method. This section describes how the Inverse Gaussian distribution is used for regional flow frequency analysis.

Regional flow frequency analysis comprises a series of steps. The first step is to ensure that the data for the study region is considered homogeneous. This step is described in Section 2.2.1. Next, a method to standardize the data sets is required in order to facilitate comparisons, analyses and generalizations on the data. This study utilizes the Index Flow Method, and is described in Section 2.2.2. Finally, a technique to relate an ungauged basin within the region to the regional analysis is required. This is described in Section 2.2.3.

2.2.1 Test for Homogeneity

Regionalisation depends on the data set belonging to a homogeneous region. The Inverse Gaussian distribution's analysis of variance methodology can be used to test a specified region for homogeneity. The hypothesis that 'all populations have the same λ with different, unspecified means', as described by Chhikara and Folk (1989), is used to examine the homogeneity of regions for analysis using the Inverse Gaussian distribution. The hypothesis is evaluated using the likelihood ratio, which can be approximated using a modified test statistic M/C, described by the following equations.

(3) $$M = f \ln\left(\frac{V}{f}\right) - \sum f_i \ln\left(\frac{V_i}{f_i}\right)$$

(4) $$C = 1 + \frac{1}{3(k-1)}\left[\sum \frac{1}{f_i} - \frac{1}{\sum f_i}\right]$$

where: $f_i = n_i - 1$ and $f = \sum (n_i - 1)$

$V_i = \sum_j (1/X_{ij} - 1/\overline{X}_i)$ and $V = \sum V_i$

M/C is distributed approximately as chi-square with (k-1) degrees of freedom

Using equations (3) and (4) to calculate M/C, the desired outcome is that M/C < $\chi^2_{k-1,\alpha}$, where α is the desired significance level. This study will consider the 0.1, 0.05 and 0.01 significance levels.

The test for homogeneity as described above for the Inverse Gaussian distribution will be applied to Newfoundland stream flow data for rivers with 13 or more years of record. Rivers with greater than 13 years of record will be used to develop the regionalisation technique and those with 10-12 years of data will verify the results. The test for homogeneity of the stream flow data for the island of Newfoundland was evaluated according to the following three categories.

1. Entire region as a whole. This is the desired outcome – that all Newfoundland streamflow data sets fall within a single homogeneous region. However, if the

test for homogeneity indicates that this is not the case, the streamflow data for Newfoundland will be subdivided into the required number of distinct regions.

2. North and south (Environment Canada Y and Z) regions. The test for homogeneity will be applied to two distinct regions. This north/south split will follow the regions specified in the original (1984) Regional Flood Frequency Analysis of Newfoundland Rivers (Government of Newfoundland and Labrador, 1984).
3. Four distinct regions. The streams within the province will be split into four regions as outlined by the current (1999) Regional Flood Frequency Analysis of Newfoundland Rivers (Government of Newfoundland and Labrador, 1999). This scenario is the least desired outcome, but will be applied in the event that conditions of homogeneity for one of the above scenarios is not met.

This study uses previously defined regions from the Government of Newfoundland's RFFA analyses and tests for homogeneity using the Inverse Gaussian distribution's analysis of variance. The RFFA studies use the method of subjective judgement based on causative factors of flood flows (1984 study) and specific mean annual peak flow (1990 and 1999 studies) to divide the province into regions that are located in geographically proximity. There are a number of other methods commonly used for delineating homogeneous regions. The most common methods, as described in Hosking and Wallis (1997) are described by the following categories:

- **Geographical**

  Geographically defined regions are those enclosed by political, administrative or physiographic boundaries, and are typically arbitrary and subjective in nature.

- **Subjective Partitioning**

  Subjective region delineation is typically based on analysis of site characteristics, such as time of flood, nature of distribution, mean annual precipitation, mean annual or flood per unit area.

- **Objective Partitioning**

  Objective region delineation is based on measured site characteristics. The region is determined based on the site's measured characteristic falling above or below a pre-determined threshold value for that characteristic.

- **Multivariate Statistical Analysis**

  Other multivariate statistical analysis of catchment characteristics or flood statistics has also been used to delineate regions of similar sites.

While there are other methods in practice today, these are among the most commonly applied methods. A key to any regional grouping is an effective method to test for regional homogeneity. As described above, the hypothesis that 'all populations have the same λ with different, unspecified means' can be applied using the Inverse Gaussian distribution theory to test for homogeneity of the assumed regions.

2.2.2 Regional Flow Frequency Factors using the Index Flow Method

The Index Flow Method uses a standardized data set to compare and provide regional estimates of data sets. A standardized data set (i.e. the index flow data) is obtained by dividing each flow ordinate by the mean for the particular data set.

The desired return period flows for each river within the homogeneous region are estimated using an appropriate statistical distribution, which is the Inverse Gaussian distribution for this study. As discussed above, the return period flow estimates for each river are standardized using the respective river's average flow. The data sets are then plotted using the Index Flow Method, or Q_T/Q_{avg} (Index Flow) versus T (Return Period).

From the data set, the regional flow frequency curve is estimated and plotted on this Index Flow plot. For this study there was a wide range of record lengths. For each return period the weighted average of the data set was determined using the record length as the weighting factor. The result is a regional curve with a mean of 1 and a regional estimate of λ.

The Method of L-Moments, as described by Pokhrel (2002) and Hosking (1997) is becoming an increasing popular method for estimating growth curves. However, as described in detail in Hosking (1997) the regional average mean, L-CV and L-skewness are required. Again, the mean is 1.0, however two additional parameter estimates are required in order to develop the regional growth curve compared to just one for the Inverse Gaussian / Index Flow Method.

2.2.3 Regional Estimate of Q_{avg}

The regional Index Flow Plot described in 2.2.2 above is used to obtain the flow frequency factor of a desired return period. The regional curve, from which this flow frequency factor is determined, has a mean of 1 and a regional value of λ. This value is used in conjunction with the average flow (Q_{avg}) of the ungauged basin to estimate the desired return period flow for that basin.

For an ungauged basin, the average flow is another unknown value. However, the basins within the region can be used to estimate the average flow of an ungauged site. There are many factors contributing to the flow within a basin. Richter (1996), Pokhrel (2002) and RFFA (Government of Newfoundland and Labrador, 1999) use additional factors in the regression analysis to develop multi-parameter regression equations. The additional factors that are used depend on the particular region. However, the drainage area of the basin has the largest influence on runoff. In an attempt to simplify application of the methodology developed in this study, and to reduce the number of 'unknowns' requiring estimation, the drainage area is the only factor considered in this study for estimating the average flow. Regression of the mean flow versus drainage area using all rivers in the region provides a relationship by which the average flow at an ungauged site can be estimated using only the drainage area.

3.0 Data Preparation and Single Site Analysis

This study considers application of the Inverse Gaussian distribution to both daily maximum and daily minimum annual flows. Section 3 describes the study area, the data preparation and the single site analysis performed on both the high and low flow data sets.

3.1 Study Area

The analysis is for the Island of Newfoundland and all basins gauged by Environment Canada with 10 or more years of record were considered for the study. Regulated basins, basins with unknown drainage areas and basins with uncertainty in the data set were excluded.

A total of 58 gauging stations were selected for use in this study. A complete list of these rivers, along with the corresponding years of record and drainage area is provided in Table 3-1, and the locations are shown in Figure 3-1. The following rivers were excluded from the study.

- Gander River at the outlet of Gander Lake (02YQ002)

  Period of record (1923 – 1939) was excluded. The record was a manual gauge, perhaps maintained to different standards, making the data not directly comparable to the other gauges. However, the Gander River at Big Chute (02YQ001) provides good coverage of the same basin (Richter, 1994).

- Grey River (02ZD002)

  Fisheries culverts at the upstream divide of Grey River basin releases an unknown volume of water from Meelpaeg Reservoir into the basin (Richter, 1994).

- Indian Brook at Indian Falls (02YM001)

  Culvert diverts an unknown amount of water from the basin into Birchy Lake (Richter, 1994).

- Indian Brook Diversion (02YM002)

  Drainage area is unknown.

- Lewaseechjeech Brook (02YK002)

  Period of record prior to 1973 excluded from the analysis. The gauge was removed in 1967 and reinstalled in 1973 because the outlet of Little Grand Lake, located just upstream of the gauge, was blasted.

- Waterford River at Kilbride (02ZM008)

  Basin is subject to increasing urbanization.

- Waterford River at Mount Pearl (02ZM010)

  Basin is subject to increasing urbanization.

- Leary Brook at St. John's (02ZM017)

  Basin is subject to increasing urbanization.

- Virginia River at Pleasantville (02ZM018)

  Basin is subject to increasing urbanization.

- Virginia River at Cartwright Place (02ZM019)

  Basin is subject to increasing urbanization.

- Leary Brook at Prince Philip Drive (02ZM020)

  Basin is subject to increasing urbanization.

Table 3-1 Gauging Stations Utilized in Analysis

Truncated Station ID	EC Station ID	Station Name	Drainage Area (km^2)	EC Count
YA1	02YA001	Ste. Genevieve River near Forresters Point	306.0	26
YC1	02YC001	Torrent River at Bristol's Pool	624.0	36
YD1	02YD001	Beaver Brook near Roddickton	237.0	20
YD2	02YD002	Northeast Brook near Roddickton	200.0	15
YE1	02YE001	Greavett Brook above Portland Creek Pond	200.0	11
YF1	02YF001	Cat Arm River above Great Cat Arm	611.0	15
YH1	02YH001	Bottom Creek near Rocky Harbour	611.0	10
YJ1	02YJ001	Harrys River below Highway Bridge	640.0	27
YK2	02YK002	Lewaseechjeech Brook at Little Grand Lake	470.0	39
YK3	02YK003	Sheffield River at Sheffield Lake	470.0	12
YK4	02YK004	Hinds Brook near Grand Lake	529.0	24
YK5	02YK005	Sheffield Brook near Trans Canada Highway	391.0	23
YK7	02YK007	Glide Brook below Glide Lake	391.0	11
YL1	02YL001	Upper Humber River near Reidville	2110.0	67

YL4	02YL004	South Brook at Pasadena	2110.0	12
YL5	02YL005	Rattler Brook near McIvers	2110.0	10
YM3	02YM003	South West Brook near Baie Verte	93.2	15
YN2	02YN002	Lloyds River below King George IV Lake	469.0	14
YO6	02YO006	Peters River neat Botwood	177.0	14
YO7	02YO007	Leech Brook near Grand Falls	177.0	11
YO8	02YO008	Great Rattling Brook above Tote River Confluence	177.0	11
YO10	02YO010	Junction Brook near Badger	469.0	10
YP1	02YP001	Shoal Arm Brook near Badger Bay	63.8	13
YQ1	02YQ001	Gander River at Big Chute	4444.0	46
YQ4	02YQ004	Northwest Gander River near Gander Lake	4444.0	12
YR1	02YR001	Middle Brook near Gambo	275.0	36
YR2	02YR002	Ragged Harbour River near Musgrave Harbour	399.0	18
YR3	02YR003	Indian Bay Brook near Northwest Arm	554.0	14
YS1	02YS001	Terra Nova River at Eight Mile Bridges	1365.0	34
YS3	02YS003	Southwest Brook at Terra Nova National Park	36.7	28
YS5	02YS005	Terra Nova River at Glovertown	36.7	10
ZA1	02ZA001	Little Barachois Brook near St. George's	343.0	17
ZA2	02ZA002	Highlands River at Trans-Canada Highway	72.0	13
ZA3	02ZA003	Little Codroy River near Doyles	139.0	13
ZB1	02ZB001	Isle aux Morts River below Highway Bridge	205.0	33
ZC2	02ZC002	Grandy Brook below Top Pond Brook	230.0	13
ZE1	02ZE001	Salmon River at Long Pond	2640.0	22
ZF1	02ZF001	Bay du Nord River at Big Falls	1170.0	45
ZG1	02ZG001	Garnish River near Garnish	205.0	37
ZG2	02ZG002	Tides Brook below Freshwater Pond	166.0	18
ZG3	02ZG003	Salmonier River near Lamaline	115.0	15
ZG4	02ZG004	Rattle Brook near Boat Harbour	42.7	14
ZH1	02ZH001	Pipers Hole River at Mothers Brook	764.0	43
ZH2	02ZH002	Come by Chance River near Goobies	43.3	27
ZJ1	02ZJ001	Southern Bay River near Southern Bay	67.4	19
ZJ2	02ZJ002	Salmon Cove River near Champneys	67.4	12
ZK1	02ZK001	Rocky River Near Colinet	300.0	47
ZK2	02ZK002	Northeast River near Placentia	89.6	16
ZK3	02ZK003	Little Barachois River near Placentia	89.6	12
ZK4	02ZK004	Little Salmonier River near North Harbour	89.6	12
ZL3	02ZL003	Spout Cove Brook near Spout Cove	10.8	16
ZL4	02ZL004	Shearstown Brook at Shearstown	10.8	12
ZL5	02ZL005	Big Brook at Lead Cove	10.8	10
ZM6	02ZM006	Northeast Pond River at Northeast Pond	3.6	42
ZM9	02ZM009	Seal Cove Brook near Cappahayden	53.6	16
ZM16	02ZM016	South River near Holyrood	10.8	12
ZN1	02ZN001	Northwest Brook at Northwest Pond	53.3	29
ZN2	02ZN002	St. Shotts River neat Trepassey	53.3	10

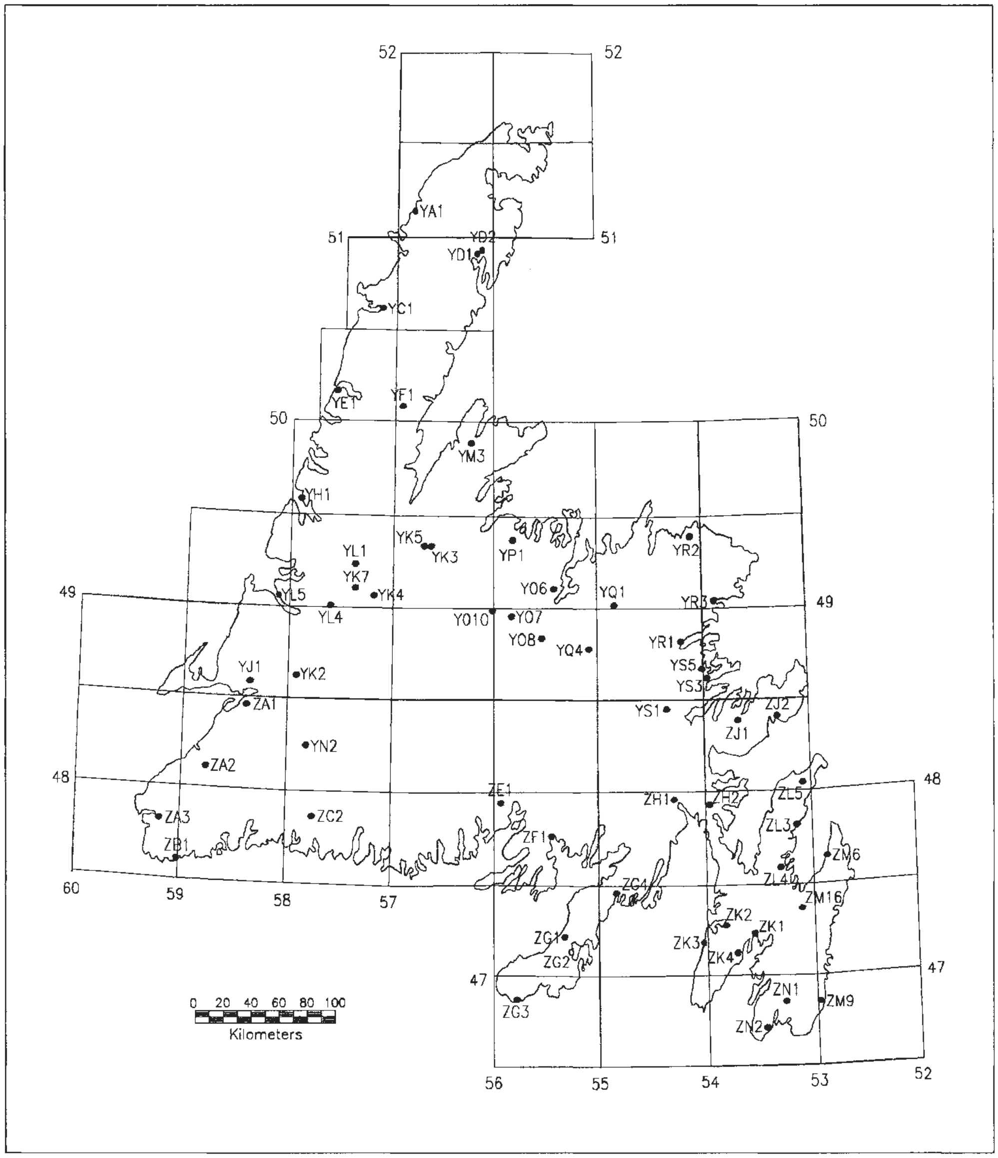


Figure 3-1 Location of Environment Canada Gauging Stations

3.2 Data Preparation

With the list of rivers established as described in the previous section, the data for each of these gauging stations was gathered and manipulated into a format usable for the analysis.

A database of Environment Canada's gauged rivers is available on CD-ROM. The CD-ROM, issued annually, is called HYDAT and data can be readily exported in a variety of usable formats (Environment Canada, 1996). Once the necessary data is exported from this database, only minor manipulation is required.

The data sets are then ready for use in the analysis. The remainder of this thesis describes the analysis.

3.3 Single Site Analysis

Single site analysis is used in this study to assess how the Inverse Gaussian distribution performs; that is, how well it fits the data and how it compares to other distributions normally used in the analysis of annual flow data. The following two sub-sections discuss the single site analysis for annual maximum and minimum flow data, respectively.

3.3.1 High Flow Analysis

For high flows, the single-site analysis compares the Inverse Gaussian distribution to four other statistical distributions that are more commonly used for flood frequency analysis. The objective of the comparison is to determine if the Inverse Gaussian distribution is an effective model for flood frequency analysis.

Flood frequency analysis is a major component of water resource or hydrotechnical engineering. Environment Canada (1988 and 1993) has produced the Consolidated Frequency Analysis (CFA) program which uses the GEV, 3PLN and Log-Pearson Type 3 distributions to analyze floods of various return periods (up to the 500 year flood). Many Canadian hydrologists use this program in conjunction with HYDAT. Newfoundland rivers can be modeled quite adequately using either the GEV or the 3PLN distributions and Newfoundland water resource engineers generally use these two distributions to perform flood analysis and complete flood studies. However, this study presents the 2-parameter Inverse Gaussian distribution as an alternate statistical model for Newfoundland floods.

To complete the single-site analysis, the parameters for each of the five distributions in the study are required. Available software packages are utilized for this task; the CFA software for estimating parameters of the 3-parameter distributions, and BestFit (Palisade, 1993-96) for the 2-parameter distributions evaluated in this study. BestFit is a statistical software package that fits data to a wide range of statistical distributions.

CFA: 3 Parameter Distributions

The General Extreme Value and the 3-Parameter Lognormal distributions are among the most popular 3-parameter distributions used in flood frequency analysis for Canadian rivers. Both are included in the CFA software program used Canada-wide for flood frequency analysis. Therefore the CFA software program is used in this analysis to estimate the parameters of both the GEV and 3PLN distributions.

The Maximum Likelihood Estimate (MLE) of the statistical distribution parameters can be obtained for both the 3PLN and GEV distributions using the CFA software package. Once the parameters are known, an estimate of the required flows, based on the plotting position, or cumulative probability is easily determined. The Return Period, T is calculated by $T = (1 / (1-P))$, where P is the cumulative probability.

The full spectrum of desired return period flow estimates could not be determined directly from the software program. However, it is simple to express both of these distributions in terms of the variate, which is the discharge (Q) in flow frequency analysis. The transform for the GEV and 3PLN distributions are illustrated in equations (5) and (6), respectively (Environment Canada, 1993). Using an Excel spreadsheet and the MLE parameters determined using the CFA program, the estimated flows for these distributions are determined for the desired return periods.

(5) $$x_i = \xi - \frac{\alpha}{K}\{[-\ln(P_i)]^K - 1\}$$

where: x_i = discharge estimated by the GEV distribution

ξ, α, K = GEV distribution parameters

P_i = Cunnane plotting position P_i=(i-0.2)/(n+0.4)

(6) $$x_i = \alpha + \exp(SZ_i + M)$$

where: x_i = discharge estimated by the 3PLN distribution

a, S, M = 3PLN distribution parameters

Z_i = standard normal variate which can be approximated by:

$$Z_i = 5.0633(P_i^{0.135} - (1 - P_i)^{0.135})$$

P_i = Cunnane plotting position (P=(i-0.2)/(n+0.4))

BestFit: 2 Parameter Distributions

The Extreme Value (Gumbel) and the Lognormal distributions are among the most popular 2-parameter distributions for flood frequency analysis. Both of these, as well as the Inverse Gaussian distribution are included in the BestFit software program. The BestFit software is capable of calculating the required MLE's of the parameters for all three of these 2-parameter distributions based on the observed flood data. This program can then be utilized to obtain the discharge variates by inputting the desired plotting positions. The BestFit package was therefore used for all three 2-parameter distributions to determine the flow estimates of each desired return period.

Verification of the results from BestFit is obtained by comparing the flood estimates for the EV and LN distribution from the BestFit software to the estimates computed using the appropriate transformation, as described below in equations (7) and (8), respectively.

(7) $$x_i = \mu - \alpha * \ln(-\ln(P_i))$$

where: x_i = discharge corresponding to the plotting position as estimated by the EV distribution

μ, α = distribution parameters

P_i = Cunnane plotting position

(8) $$x_i = \exp(\sigma Z + \mu)$$

where: x_i = discharge estimated by the LN distribution

σ, μ = LN distribution parameters

Z_i = standard normal variate which can be approximated by:

$$Z_i = 5.0633(P_i^{0.135} - (1 - P_i)^{0.135})$$

P_i = Cunnane plotting position (P=(i-0.2)/(n+0.4))

The results obtained from the software program and the results using the above equations were an identical match for the EV distribution. The results obtained from the software program and the results using the above equation for the LN distribution were also comparable, but not identical. This difference is not significant, and is explained by the estimation of the standard normal variate, Z, used in the above equation.

The results of the Inverse Gaussian distribution estimates as determined from the software package are also verifiable by using Koziol's (1989) book of percentage points for the Inverse Gaussian distribution. Using the Inverse Gaussian parameters, the discharge estimates for various probabilities are determined. The results obtained using the tables are slightly different from those using the software. However, the results are accurate to the first decimal place. This slight difference is most likely due to the difference in the significant figures used in each method. Similar to the other two distributions, the results indicate that the BestFit software adequately calculates the Inverse Gaussian distribution variates.

Comparison of Distribution Performance

The above provides a method of estimating return period flows for each of the five statistical distributions used in the high flow analysis. Section 3.3.3 describes a methodology to assess and compare the performance of the various distributions. However, first Section 3.3.2 describes the single site analysis technique for low flows.

3.3.2 Low Flow Analysis

For low flows, the single-site analysis also compares the Inverse Gaussian distribution to four other statistical distributions that are more commonly used for low flow frequency analysis. The objective of the comparison is to determine if the Inverse Gaussian distribution is an effective model for frequency analysis of low flows.

Like flood frequency analysis, low flow analysis is a major component of water resource or hydrotechnical engineering. This study compares the Inverse Gaussian distribution with some of the more popular distributions that are used for frequency analysis of low flows. This study investigates the 2-parameter Inverse Gaussian distribution as an alternate statistical model for Newfoundland low flow analysis.

To conduct the single-site analysis for low flows, the parameters of the five distributions in the study are required. The software package BestFit was utilized for this task, as discussed below.

BestFit: 2 Parameter Distributions

The Extreme Value (Gumbel) and the Lognormal distributions are among the most popular 2-parameter distributions for low flow analysis as well. The method to derive distribution parameters and estimate the desired return period low flows for these two distributions and the Inverse Gaussian was described above in Section 3.3.1.

The other two distributions included in the low flow analysis are the 2-parameter and 3-parameter Weibull distributions. The software packages available for use were not capable of readily estimating the 3-parameter Weibull distribution. However, as described by Zanakis (1978), there is a simple method of determining the MLE estimates of the parameters. This study utilizes the simple estimating method described by Zanakis (1978) to determine the location (third) parameter of the 3-parameter Weibull distribution. With this value known, the distribution is essentially a 2-parameter distribution and obtaining the MLE estimate of the remaining two parameters is possible using Best Fit.

The simple method for estimating the Weibull location parameter is described below in equation (9).

(9) $$\bar{a} = \frac{(y_1 y_n - y_2^2)}{(y_1 + y_n - 2y_2)}$$

where: $\bar{a}$ = simple MLE estimate for 3-parameter Weibull location parameter

y_1= minimum value in data set

y_2 = second smallest value in data set

y_n = largest value in data set

The flow estimates of the Weibull distribution are obtained by manipulating the cumulative distribution function of the Weibull distribution as described below in equation (10).

(10) $$x_i = a + b[-\ln(1-6)]^{1/c}$$

where: x_i = discharge estimated by the Weibull distribution

a = location parameter (0 for the 2-parameter distribution)

b = scale parameter

c = shape parameter

Comparison of Distribution Performance

The above provides a method of estimating return period flows for each of the five statistical distributions used in the low flow analysis. Section 3.3.3 describes the method to assess and compare the performance of the various distributions.

3.3.3 Use of the Akaike Information Criterion Method for Comparing Performance of Statistical Distributions

The Akaike Information Criterion (AIC) is recommended by Chow and Watt (1992) as an aid in selecting a statistical distribution for flood frequency analysis and by Lawal and Watt (1996) for selecting a statistical distribution to model low flows.

The use of a statistical distribution with additional parameters may visually indicate that the distribution provides a slightly better fit to the data set. However, there are increased uncertainties with the parameter estimates and the derived flow estimates that may make the distribution less desirable. The AIC selection criterion is a means to evaluate whether the improved fit of the 3-parameter distribution compensates for the penalty associated with the uncertainty of third parameter. The AIC is calculated using equation (11), as follows:

(11) $$AIC = (-2)LL(x;\hat{\theta}) + 2m$$

where: AIC = Akaike Information Criterion

$LL(x;\hat{\theta})$ = log-likelihood function of the statistical model

m = number of independently adjusted parameters

To use the AIC as a selection method when several models are considered, the distribution that results in the lowest AIC is selected. As discussed in Lawal and Watt (1996) and Chow and Watt (1992), the AIC is decreased by reducing '$LL(x;\hat{\theta})$' with a better fit and increased by '2m' when adding more model parameters. For this study, the AIC is calculated for each distribution, for each data set. Using the AIC as a guide, each distribution is ranked on how well the distribution estimates the observed data set. The results are summarized and discussed in Section 4.

4.0 Use of the Inverse Gaussian Distribution for Flow Estimation - Results and Discussion

The technique for single site analysis and the method of regionalisation using the Inverse Gaussian distribution are described in the previous sections. This section discusses the results of applying the described methodology.

4.1 Single Site Analysis Results

The techniques and methods described in Section 3 were used to determine how the Inverse Gaussian distribution compares to other distributions for flow analysis of Newfoundland rivers. As discussed in Section 3.3.3, the Akaike Information Criterion was used to determine how the distributions compare for both the low flow and the high flow analysis. Appendices A and B contain the single site analysis AIC values for high and low flows, respectively. A total of 58 stations with 10 or more years of record are included in the single site analysis for high flows. For low flows, 29 stations with 15 or more years of record were used. Stations with 10 through 14 years of record were not included for low flows as the study sufficiently shows the methodology is not appropriate for low flow analysis of Newfoundland rivers.

Tables 4-1 and 4-2 below summarize the AIC ranking of each distribution using summary statistics (sum, mean, median and mode) and frequency of ranking by distribution, respectively. Figures 4-1 and 4-2 graphically illustrate the results shown in Table 4-2 as bar charts.

Table 4-1 Summary Statistics of the AIC Ranking of Each Distribution

Distribution	Sum of Ranks	Mean Rank	Median Rank	Mode Rank
High Flows (≥ 10 years of record)				
IG	101	1.7	2	1
LN	132	2.3	2	2
EV	179	3.1	3	3
GEV	217	3.7	4	5
3PLN	241	4.2	4	5
Low Flows (≥ 15 years of record)				
IG	117	4.0	4	5
LN	90	3.1	3	4
EV	77	2.7	2	2
W2	50	3.5	1	1
W3	101	1.7	3	3

Note: a rank of 1 indicates the lowest AIC value, while a rank of five indicates the highest AIC. A lower AIC value is desirable, therefore the lower numbers in the above table indicate the more suitable distribution.

Table 4-2 Frequency of AIC Ranking for the Five Distributions

Distribution	Frequency of Each Rank by Distribution				
	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5
High Flow (≥ 10 years of record)					
IG	45 %	40 %	12 %	3 %	0 %
LN	14 %	53 %	26 %	5 %	2 %
EV	19 %	2 %	45 %	21 %	14 %
GEV	17 %	2 %	10 %	31 %	40 %
3PLN	5 %	3 %	7 %	40 %	45 %
Low Flows (≥ 15 years of record)					
IG	10 %	3 %	7 %	31 %	48 %
LN	3 %	31 %	17 %	48 %	0 %
EV	14 %	38 %	28 %	10 %	10 %
W2	72 %	3 %	14 %	0 %	10 %
W3	0 %	24 %	34 %	10 %	31 %

The results shown in Tables 4-1 and 4-2 indicate that based on the AIC selection method, the 2-parameter distributions more adequately fit the observed data sets. For high flows, the AIC value indicates that the Inverse Gaussian distribution is the most appropriate candidate for analyzing the data, ranking first and second 45% and 40% of the time, respectively. The results indicate that the Inverse Gaussian distribution significantly out-performs the other statistical distributions for analysis of high flows. However, the opposite is true for the low flow analysis. The results indicate that the Inverse Gaussian distribution is not a suitable candidate for low flow analysis of Newfoundland rivers, and is in fact the least suitable statistical distribution ranking in last place 48% of the time and second to last 31% of the time.

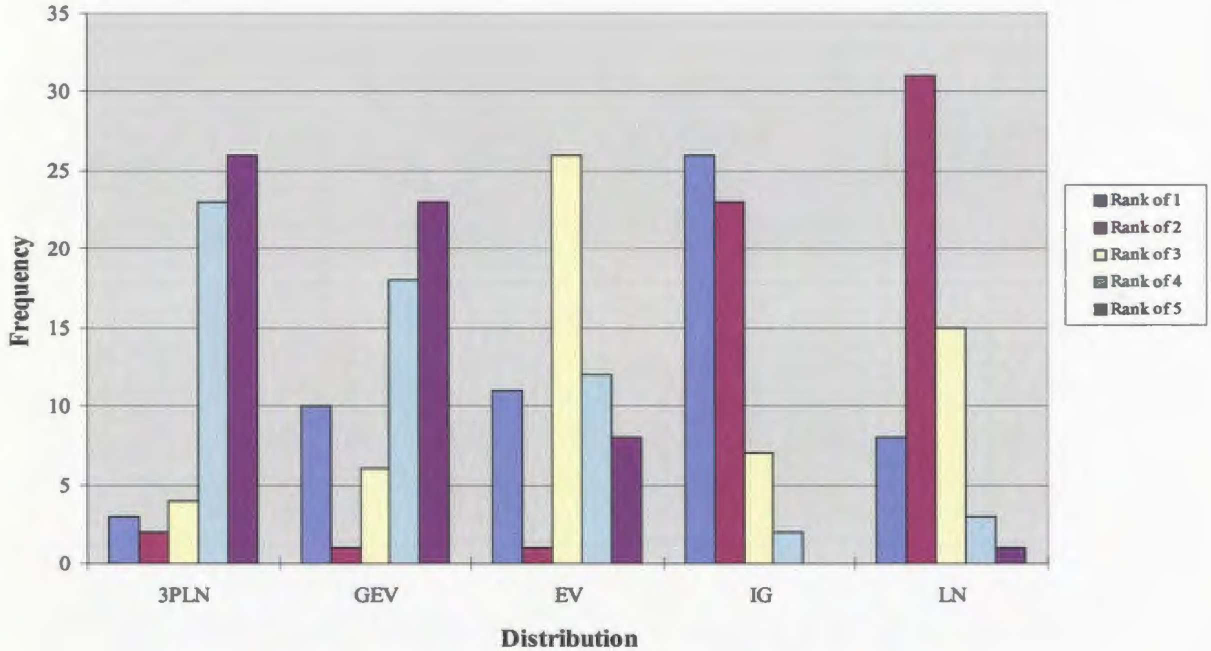


Figure 4-1 Bar Chart of AIC Ranking for High Flow Analysis

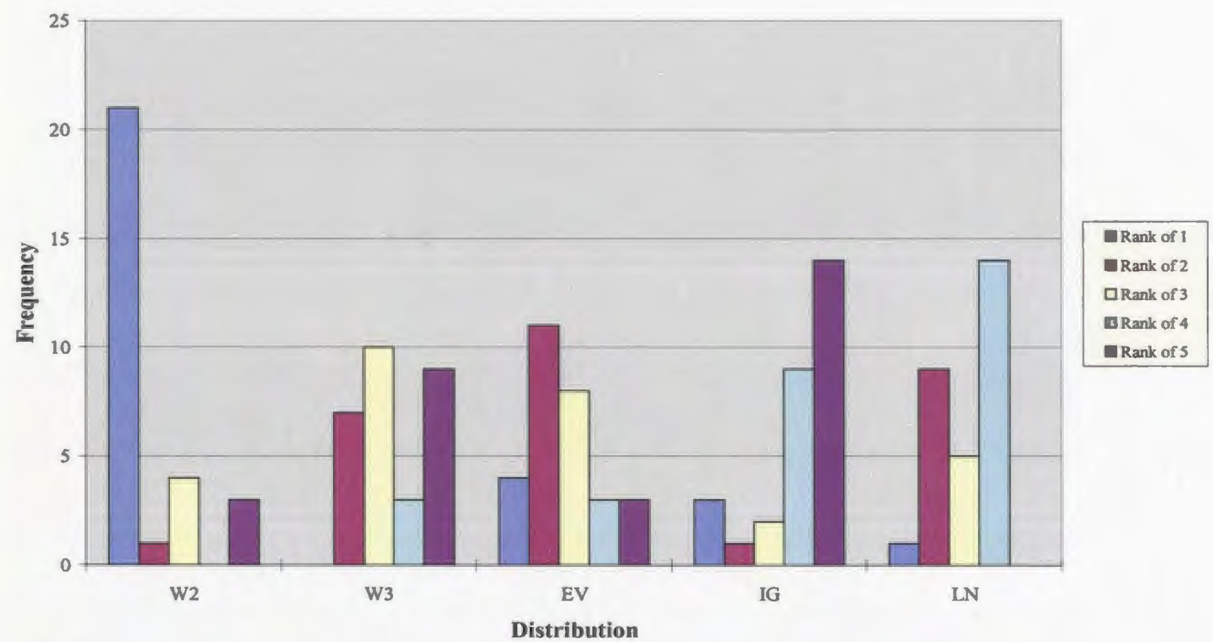


Figure 4-2 Bar Chart of AIC Ranking for Low Flow Analysis

The bar charts shown in Figures 4-1 and 4-2 provide a graphical interpretation of the results for the high and low flow analysis, respectively. Figure 4-1 shows how the Inverse Gaussian distribution most often ranks first, with a few occasions where other distributions provide a better estimate of the observed data for the single site analysis. The 3-parameter distributions, the 3PLN and GEV, are trailing behind the 2-parameter distributions. Conversely, Figure 4-2 shows how poorly the Inverse Gaussian distribution performs in the low flow analysis compared with the other 2-parameter distributions and the 3-parameter Weibull distribution.

The ranking of each distribution based on the AIC values indicates that the 3-parameter distributions typically do not overcome the penalty associated with having the third parameter. Overall, the 2-parameter distributions have a lower AIC value than the 3-parameter distributions, with the 2-parameter Inverse Gaussian distribution performing extremely well for high flows, but extremely poorly for low flows.

Although the Inverse Gaussian Distribution was not considered in their analysis, these results follow closely with results by Lawal and Watt (1996) and Chow and Watt (1991) in their use of the AIC to compare suitability of distributions for the low and high flows, respectively. Their investigation of the AIC as a measure of the goodness of fit and cost of parameter uncertainty indicates that the 'improved' goodness of fit does not overcome the uncertainty associated with the third parameter. In both Lawal and Watt (1996) and

Chow and Watt (1991) the 2-Parameter Lognormal Distribution is clearly the distribution of choice based on the AIC.

In both the high flow and low flow analysis of long-term hydrometric stations across Canada by Chow and Watt (1991), and Lawal and Watt (1996) respectively, there is only one Newfoundland station considered. For the low flow analysis the station is 02YL001 (Upper Humber River near Reidville) with 55 years of record. For the high flow analysis, the station is 02ZK001 (Rocky River near Colinet) with 37 years of record. Obviously, these studies did not include the Inverse Gaussian distribution in the analysis. However, we can still compare the analyses. For the low flow analysis, this study had 65 years of record for the 02YL001 station. In comparison with Lawal and Watt, the results are similar with the first two ranked distributions being the same. In Lawal and Watt, the 3-parameter Weibull and 2-parameter Weibull ranked third and fourth, respectively. In the present analysis, the 3- and 2- parameter Weibull distributions were actually reversed and the Inverse Gaussian ranked between the two.

For the high flow analysis, Chow and Watt found that for 02ZK001, the 3PLN distribution ranked first, with the LN, EV1 and Normal as second, third and fourth place respectively. With an additional 8 years of data, the 3-parameter GEV distribution was ranked in first place, with IG, 3PLN, LN and EV falling in behind. These results support the finding that the Inverse Gaussian distribution is a strong candidate for flood frequency analysis and overall is a strong competitor against 3-parameter distributions.

Visual investigation is often a key factor in determining how well a statistical distribution fits a particular data set. For each station in the study, the five distributions are plotted along with the observed data set for comparison. These plots are shown in Appendix C and D for the high flow and low flow analyses, respectively. Examination of these plots suggests that all five of the distributions are typically very reasonable in fitting the observed data set. The third parameter in the 3-parameter distributions provides more flexibility in fitting the curved nature of some of the observed data sets. However, as identified by the 3-parameter distributions having a higher AIC value, the added flexibility is typically not sufficient to overcome the uncertainty associated with the third parameter.

For the low flows, the extreme low value of the observed data is underestimated approximately 35% of the time, with no real distinction between how often the 2-parameter underestimates the low flow compared to the 3-parameter distributions. For the high flows, the extreme maximum value of the observed data is underestimated approximately 30% of the time by the 2-parameter distributions and about 20% of the time by the 3 parameter distributions. The remainder of the time the estimates are either very close to the observed data or slightly above the observed data. It is necessary to note that a factor of safety is generally added to any engineering design, especially when using estimates of an unknown factor such as flood frequency based on historical information. Considering the factor of safety, the fact that the 2-parameter distributions seem to underestimate the flow 10% more often than the 3-parameter distributions does not mean that these 2-parameter distributions should be discarded as an inadequate method of flood frequency analysis. Knowing that 2-parameter distributions are more likely to underestimate the high return period flows approximately 10% of the time, the factor of safety used in engineering analysis can include this for estimates and designs.

4.2 Regionalisation

The single site analysis indicates that the Inverse Gaussian distribution is a strong candidate for high flow analysis, but is inappropriate for low flow analysis of Newfoundland streamflow data. Based on this conclusion, only the annual maximum daily flows are assessed for regionalisation using the Inverse Gaussian distribution. The remainder of this thesis concentrates on the analysis of the high flows, or annual maximum daily flows. The low flow analysis conclusion is that the Inverse Gaussian distribution is not a suitable candidate for such analysis of Newfoundland streamflow data.

Regionalisation requires three steps.

1. First, the region of analysis must pass the test of homogeneity - that is the data within the region must be homogeneous.

2. Second, a relationship of the desired return period flows within the region is required.

3. Finally, the relationship between the average flow and the representative basin characteristics must be known.

The following sub-sections describe the analysis results for each of these three stages. Gauged stations with 13 or more years of record are used for the regionalisation described here. The remaining stations with 10-12 years of record are used for testing the results.

4.2.1 Test for Homogeneity

The methodology for using the Inverse Gaussian distribution to test for homogeneity was discussed in Section 2.2. This section discusses how the test for homogeneity described in Section 2.2.1 is applied to the Newfoundland streamflow data.

As discussed in Section 2.2.1, there are many ways to group data to perform regionalisation. One of the most common methods is by subjective judgement based on causative factors within a geographical location. This is the method used for the Regional Flood Frequency Analyses of Newfoundland prepared by the Department of Water Resources. Another popular method of regionalisation is the use of scatter plots to identify regional trends in data. However, one of the most useful applications of a regional flood frequency analysis is to estimate desired return period streamflow data for an ungauged station. This study uses previously defined regions from the Government of Newfoundland's RFFA analysis (1984 and 1999) and tests for homogeneity using the Inverse Gaussian distribution's analysis of variance methodology.

The key to regionalisation is finding the appropriate level of segregation of the study area. For this reason, segregation of gauged stations into separate regions is accomplished by beginning with larger scale regions than the current RFFA of Newfoundland. In this analysis, the streamflow network for the island of Newfoundland is considered as a single region, as two regions divided into approximately north and south regions as defined in the original RFFA of 1988, and finally as four separate regions as identified in the current RFFA, 2000.

As discussed previously in Section 2.2.2, regionalisation using the Inverse Gaussian distribution can be evaluated using the hypothesis that 'all populations have the same λ

with different, unspecified means'. Table 4-3 contains the results of applying this hypothesis to the entire island as a single region, the island as two distinct regions and the island as four distinct regions, using gauged stations with 13 or more years of data.

Table 4-3 Results of Regionalisation Technique

Region	M/C Value	M/C Significance Level		
		0.05	0.025	0.01
Island as one region				
Island of Newfoundland	70.5	54.6	58.1	62.4
Island as two distinct regions				
North (Y- Region)	29.7	30.1	32.9	36.2
South (Z-Region)	18.4	30.1	32.9	36.2
Island as Four Separate Regions				
NE	19.1	19.7	21.9	24.7
NW	8.9	15.5	17.5	20.1
SE	9.3	16.9	19.0	21.7
SW	8.6	15.5	17.5	20.1

Where possible the preference is to use the least number of regions required to describe the data set. As shown in Table 4-3, using the Inverse Gaussian distribution test for homogeneity, the entire river network on the island of Newfoundland is not a single homogeneous region. However, the separation of Newfoundland rivers into just two distinct regions, a North Region and a South Region, does pass the significance test at all three levels of significance, including the desired 5% level of significance. The four regions as identified in the most recent RFFA for the Island of Newfoundland are indeed homogeneous using this test, as identified in the RFFA. However, since the segregation of two regions passes the test for homogeneity, the remainder of this study focuses on the island of Newfoundland as two distinct regions for streamflow analysis.

This follows closely with the results of Pokhrel (2002). For instantaneous peak flow estimation, Pokhrel found that the division of Newfoundland stream gauging stations into two regions based on Environment Canada Y and Z subregions is appropriate. Using the method of L-moments, Pokhrel shows that the two regions are both statistically and operationally homogeneous. Similarly, Richter (1994) does not support the four regions of the 1989 RFFA study. However, the two regions of North and South based on the Y-Z division are supported based on the distribution of specific flood and average daily maximum flood across the island of Newfoundland.

4.2.2 Index Flow

An Index Flow approach to frequency analysis is used to derive regional growth curves for predicting extreme floods as a function of average annual maximum daily flows. The previous section proves that the Island of Newfoundland's streamflow gauging stations belong to two separate regions dividing the island into approximately north and south homogeneous regions. Within each region a relationship between the streamflow data is determined using the Index Flow approach. Data from gauged catchments within the region are analyzed to determine the magnitude of a relatively frequent flood, such as the average annual maximum daily flood, and then to estimate the ratios of less frequent floods to the more common event. This approach is based on the assumption that a small flood such as an average annual maximum flood can be estimated with reasonable confidence. The estimate of the average annual flood is discussed in the following section, Section 4.2.3. This section concentrates on the development of the Index Flows for a known region.

The estimate of larger floods is based on the compilation of ratios of less frequent floods to the more common event for all gauged basins within a region. For each region, the ratios of flows with estimated return periods to the average maximum daily flows is then calculated for return periods ranging from 2 to 10 000 years. These are known as index flows and are presented in tabular form in Table 4-4 and graphically in Figures 4-3 and 4-4 for the north and south regions, respectively.

Table 4-4 Index Flow Flows

Station	Return Period									
	2	5	10	20	50	100	200	500	1000	10 000
YA1	0.956	1.230	1.402	1.561	1.758	1.901	2.041	2.222	2.356	2.793
YC1	0.951	1.243	1.429	1.601	1.817	1.974	2.128	2.328	2.477	2.963
YD1	0.958	1.227	1.396	1.551	1.744	1.884	2.020	2.196	2.327	2.753
YD2	0.970	1.194	1.329	1.452	1.603	1.710	1.815	1.948	2.047	2.364
YF1	0.973	1.186	1.314	1.430	1.571	1.673	1.770	1.895	1.987	2.280
YJ1	0.956	1.231	1.404	1.563	1.761	1.905	2.046	2.228	2.363	2.803
YK2	0.973	1.187	1.316	1.432	1.575	1.677	1.775	1.901	1.993	2.290
YK4	0.971	1.192	1.327	1.449	1.598	1.705	1.808	1.941	2.038	2.351
YK5	0.953	1.238	1.417	1.583	1.791	1.942	2.089	2.280	2.423	2.886
YL1	0.974	1.182	1.307	1.420	1.557	1.655	1.749	1.870	1.959	2.241
YM3	0.893	1.336	1.645	1.947	2.342	2.640	2.939	3.335	3.636	4.642
YN2	0.953	1.239	1.420	1.588	1.797	1.950	2.099	2.292	2.436	2.905
YO6	0.924	1.293	1.538	1.772	2.071	2.294	2.514	2.803	3.021	3.741
YP1	0.935	1.274	1.496	1.704	1.969	2.165	2.358	2.610	2.799	3.421
YQ1	0.960	1.223	1.387	1.537	1.724	1.859	1.991	2.161	2.287	2.696
YR1	0.962	1.218	1.377	1.522	1.703	1.833	1.960	2.123	2.245	2.637
YR2	0.960	1.222	1.385	1.534	1.719	1.854	1.984	2.153	2.278	2.685
YR3	0.968	1.199	1.339	1.467	1.624	1.736	1.845	1.985	2.088	2.421
YS1	0.961	1.220	1.380	1.528	1.711	1.843	1.971	2.137	2.260	2.659
YS3	0.955	1.234	1.409	1.570	1.772	1.919	2.062	2.247	2.385	2.832
ZA1	0.937	1.270	1.487	1.690	1.948	2.139	2.326	2.571	2.754	3.357
ZA2	0.929	1.285	1.519	1.741	2.025	2.234	2.442	2.714	2.918	3.593
ZA3	0.902	1.325	1.617	1.900	2.268	2.545	2.821	3.187	3.464	4.387
ZB1	0.912	1.312	1.584	1.845	2.184	2.437	2.689	3.021	3.272	4.106
ZC2	0.918	1.303	1.561	1.809	2.127	2.365	2.601	2.911	3.145	3.920
ZE1	0.959	1.223	1.388	1.539	1.727	1.863	1.995	2.166	2.293	2.706
ZF1	0.932	1.280	1.509	1.725	2.001	2.205	2.406	2.669	2.867	3.519
ZG1	0.942	1.262	1.468	1.661	1.905	2.085	2.261	2.490	2.662	3.226
ZG2	0.928	1.286	1.523	1.748	2.035	2.247	2.458	2.733	2.940	3.623
ZG3	0.934	1.277	1.501	1.712	1.981	2.180	2.376	2.632	2.825	3.457
ZG4	0.955	1.234	1.410	1.572	1.774	1.922	2.065	2.251	2.390	2.840
ZH1	0.925	1.291	1.534	1.765	2.061	2.281	2.499	2.784	2.999	3.708
ZH2	0.923	1.295	1.544	1.781	2.085	2.311	2.535	2.829	3.051	3.782
ZJ1	0.949	1.248	1.438	1.616	1.838	2.001	2.160	2.367	2.521	3.025
ZK1	0.932	1.280	1.508	1.723	1.998	2.201	2.401	2.663	2.860	3.508
ZK2	0.938	1.270	1.486	1.689	1.947	2.137	2.324	2.568	2.751	3.351
ZL3	0.945	1.256	1.456	1.643	1.878	2.051	2.220	2.440	2.605	3.143
ZM6	0.949	1.248	1.438	1.615	1.837	1.999	2.158	2.365	2.519	3.022
ZM9	0.974	1.183	1.310	1.423	1.562	1.662	1.757	1.879	1.969	2.257
ZN1	0.962	1.217	1.374	1.519	1.698	1.827	1.952	2.114	2.235	2.623

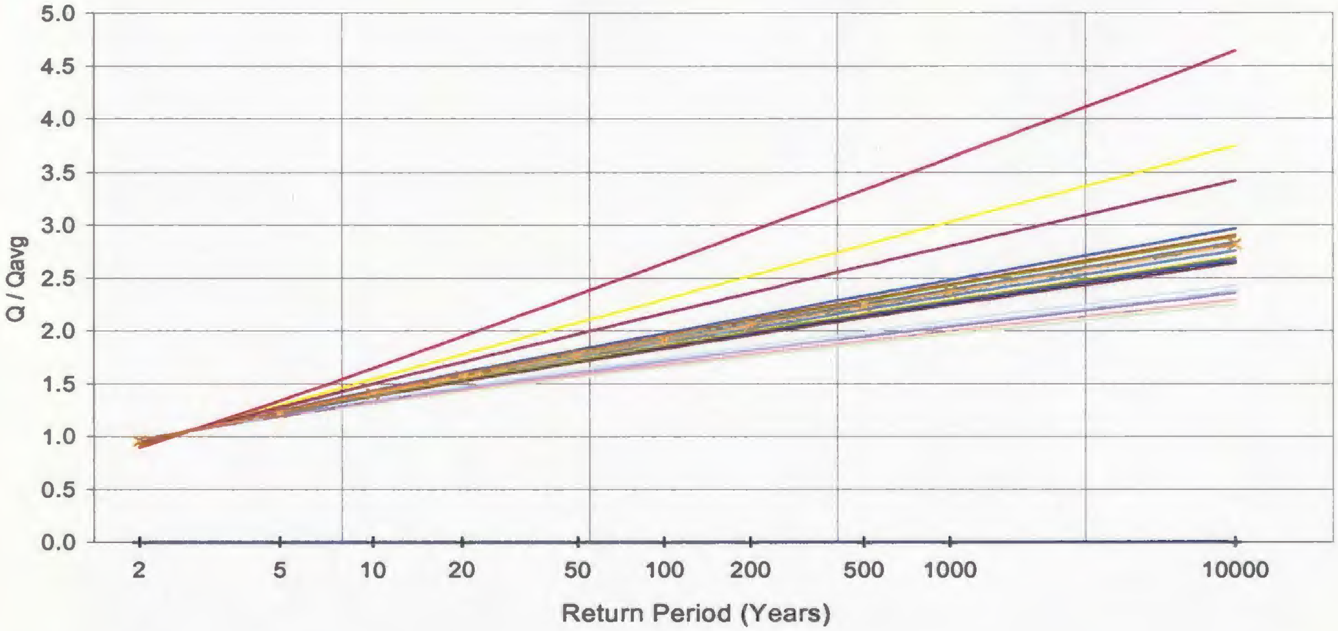


Figure 4-3 Index Flow Plot - North (Y) Region

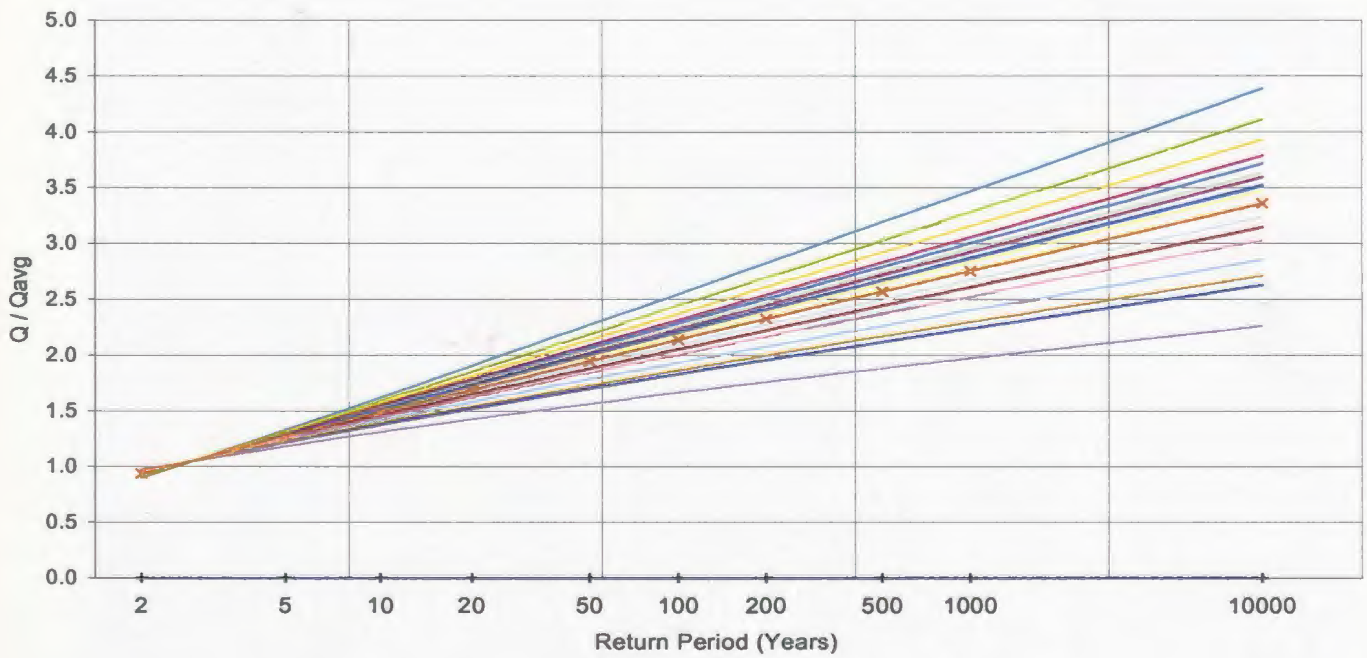


Figure 4-4 Index Flow Plot - South (Z) Region

The Index Flow table of values is used to estimate regional flood frequency factors. Table 4-5 presents the regional flood frequency factors for the north and south regions, based on a weighted average of all gauged basins with 13 or more years of data within each region. The average is weighted based on the number of years of record, under the premise that a longer period of record provides a more accurate flood estimate.

Table 4-5 Regional Flood Frequency Factors

Return Period (Years)	Annual Exceedence Probability	Regional Flood Frequency Factor (Average)	Regional Flood Frequency Factor (Weighted Average)	Lower 90% Confidence Limit	Upper 90% Confidence Limit
North (Y)					
2	0.5	0.96	0.96	0.94	0.97
5	0.2	1.23	1.22	1.16	1.30
10	0.1	1.40	1.39	1.30	1.52
20	0.05	1.56	1.54	1.42	1.73
50	0.02	1.76	1.73	1.57	2.00
100	0.01	1.91	1.87	1.68	2.20
200	0.005	2.05	2.00	1.78	2.40
500	0.002	2.23	2.18	1.92	2.67
1000	0.001	2.37	2.31	2.01	2.88
10 000	0.0001	2.82	2.73	2.32	3.58
South (Z)					
2	0.5	0.94	0.94	0.91	0.96
5	0.2	1.27	1.27	1.19	1.36
10	0.1	1.48	1.49	1.35	1.64
20	0.05	1.69	1.69	1.51	1.91
50	0.02	1.94	1.95	1.70	2.27
100	0.01	2.13	2.14	1.83	2.55
200	0.005	2.32	2.33	1.97	2.84
500	0.002	2.57	2.58	2.13	3.22
1000	0.001	2.75	2.76	2.26	3.53
10 000	0.0001	3.36	3.37	2.64	4.62

Figure 4-5 graphically illustrates the Flood Frequency Factors as a Growth Curve with the upper and lower 90% Confidence Limits. Both the North and South Regions are shown on one plot which shows that the Confidence Intervals are not significantly wide, particularly for the North (Y) Region. This suggests an acceptable level of accuracy in the estimated regional flood frequency factors. As expected, the confidence intervals widen with higher return periods. It is interesting to note that the South Region Growth Curve is steeper than the North Region Growth Curve. This is possibly explained by the fact that on average the mean annual runoff for the north central region of the island is only 700 mm to 900 mm, while that for the south western region of the island is 1300 mm to 2100 mm. Another explanation is that the flood generating mechanism has more variability – steeper curves imply more variability.

The Growth Curve established in this study using the Index Flow Method combined with the Inverse Gaussian Distribution is relatively simple to construct. All that is required is the regional average mean and the regional scale factor. For the index flow method the mean is 1.0. Therefore only one additional parameter, the scale factor, is required. Table 4-6 presents the regional average and regional weighted average Inverse Gaussian distribution scale parameters for both the North (Y) and South (Z) regions.

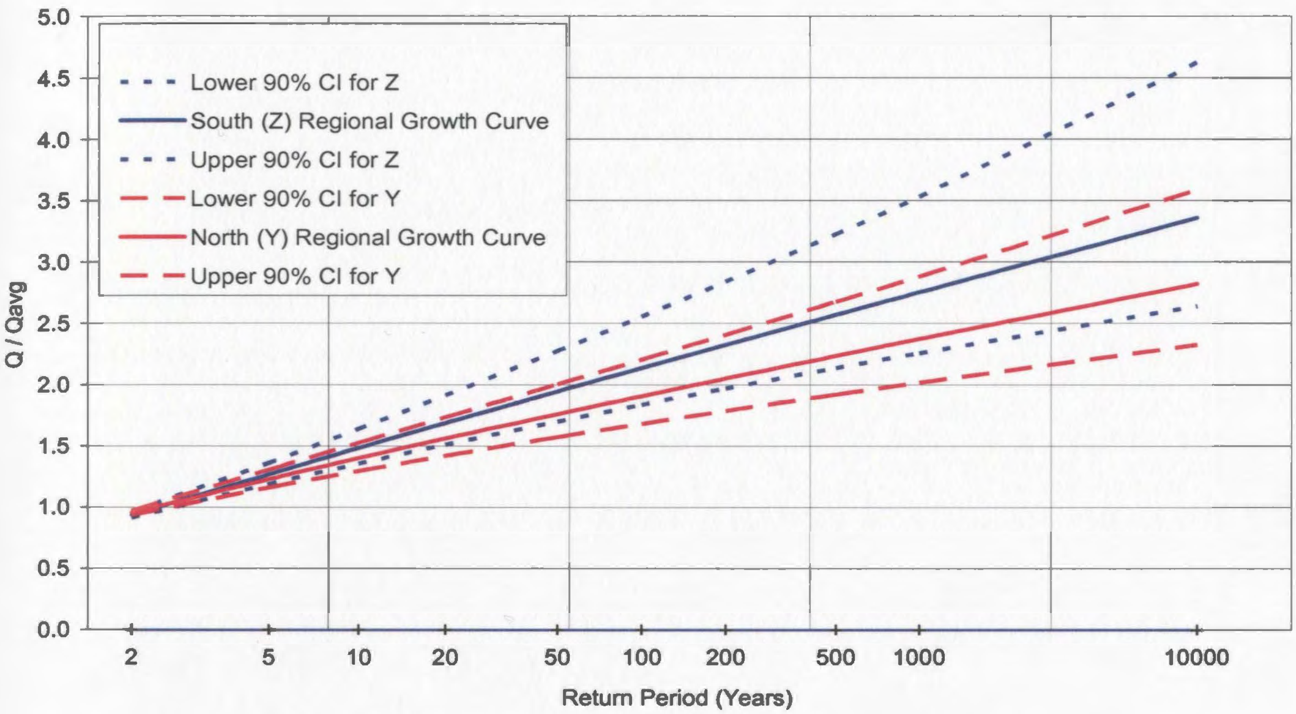


Figure 4-5 Regional Growth Curve - North (Y) and South (Z) Regions

Table 4-6 Inverse Gaussian Regional Scale Parameters

Region	Regional Scale Parameter (Weighted Average)
North (Y)	12.18
South (Z)	7.86

4.2.3 Average Flow Estimate

As discussed above, an estimate of the average annual maximum daily flow is required to use the results of the Index Flow analysis. This average flow for an ungauged site within a region is obtained by deriving a relationship between the average annual maximum daily flow and various physiographic parameters for gauged basins. The drainage area of the basin is the most significant physiographic parameter and is used in this study to derive a relationship to the average annual maximum daily flow. Providing the drainage area of the ungauged basin is known, the average annual maximum daily flow can be estimated based on this relationship.

Simple regression analysis was used to derive the relationships between drainage area and annual maximum daily flows shown in equations (12) and (13) for the north and south regions, respectively. Again, gauged stations with 13 or more years of data were used for developing this relationship.

(12) Y-North Region: $\overline{Q} = 0.5231 \times DA^{0.8586}$

(13) Z-South Region: $\overline{Q} = 1.2519 \times DA^{0.7677}$

These relationships are illustrated graphically in Figures 4-6 and 4-7. The R^2 values for the regression curves are 0.82 and 0.89, respectively for the North and South regions, confirming that the drainage area is indeed the most significant parameter and a significant portion of the relationship is described by this single parameter alone. Using this technique, only the drainage area of the ungauged site is required to obtain an estimate of the average annual maximum daily flow and subsequently the desired return period flow utilizing the Flood Frequency Factors determined in Section 4.4.2.

The residual and normal plots of the regression residuals are shown in Figures 4-8 and 4-9 for the North (Y) and South (Z) Regions, respectively. The plots indicate that the residuals are normal and no trend is apparent. This indicates acceptability of the regression equations estimated for the relationship between the Drainage Area and Mean Annual Daily Maximum Flow.

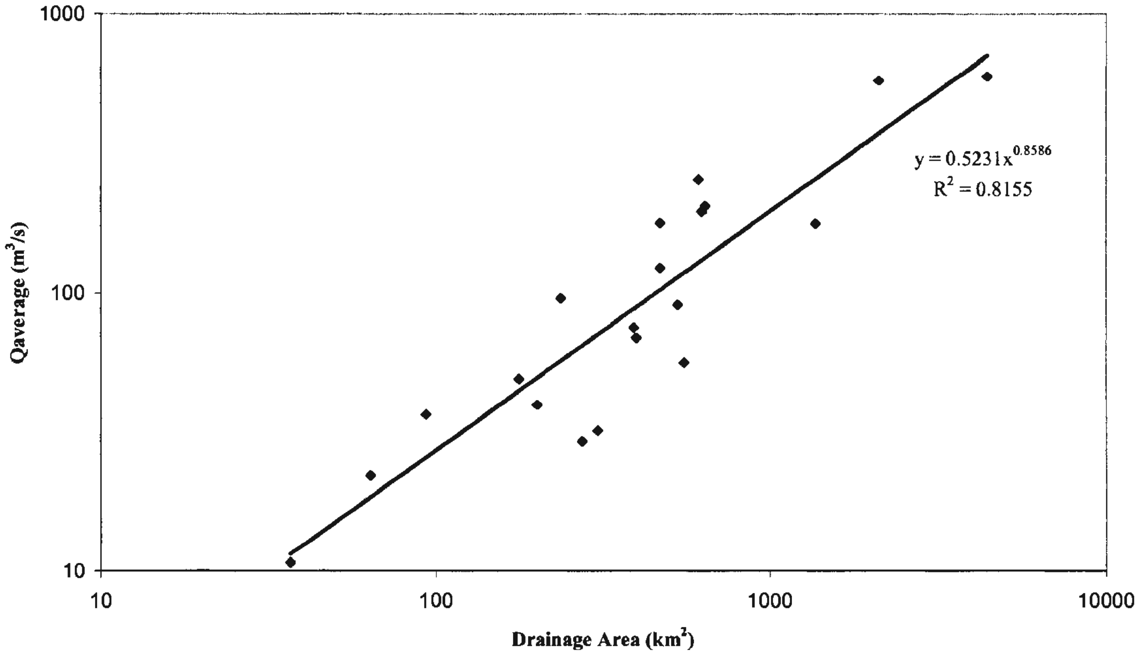


Figure 4-6 Relationship of Mean Annual Daily Maximum Flows to Drainage Area - North (Y) Region

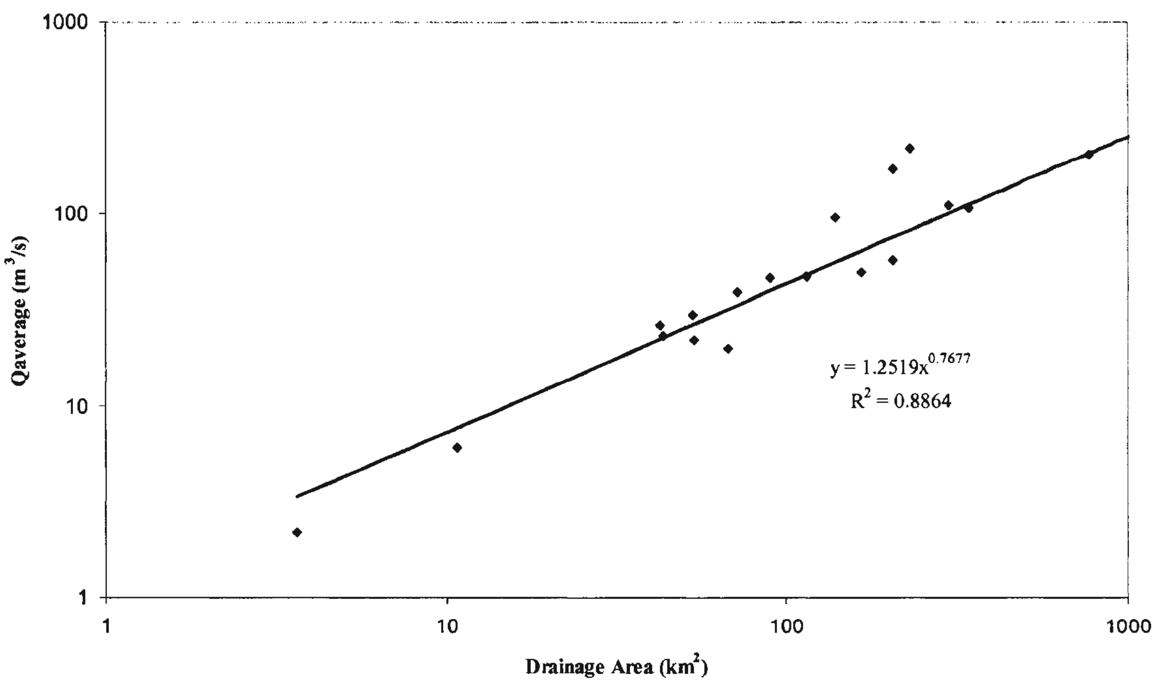
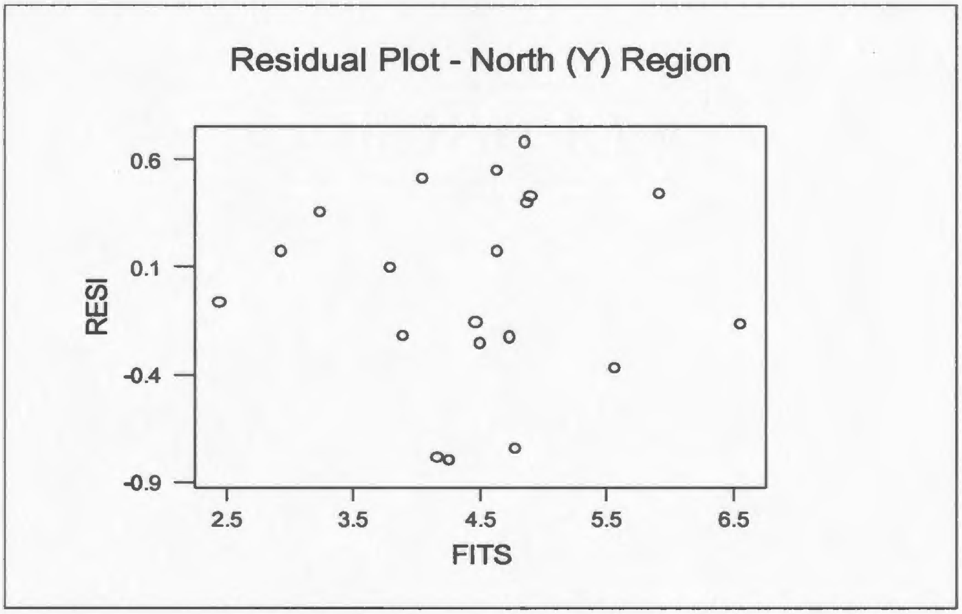


Figure 4-7 Relationship of Mean Annual Daily Maximum Flows to Drainage Area - South (Z) Region



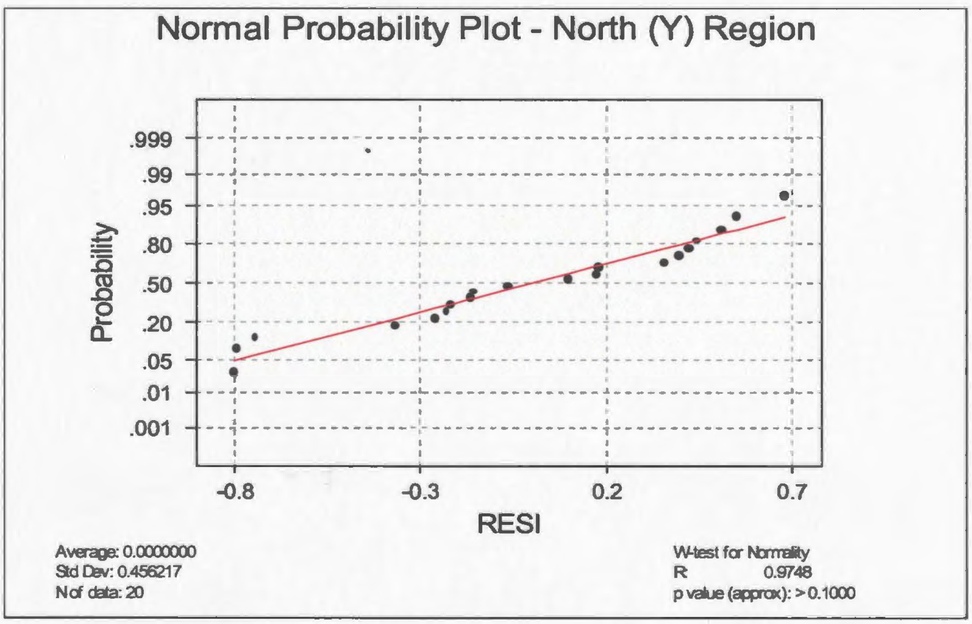
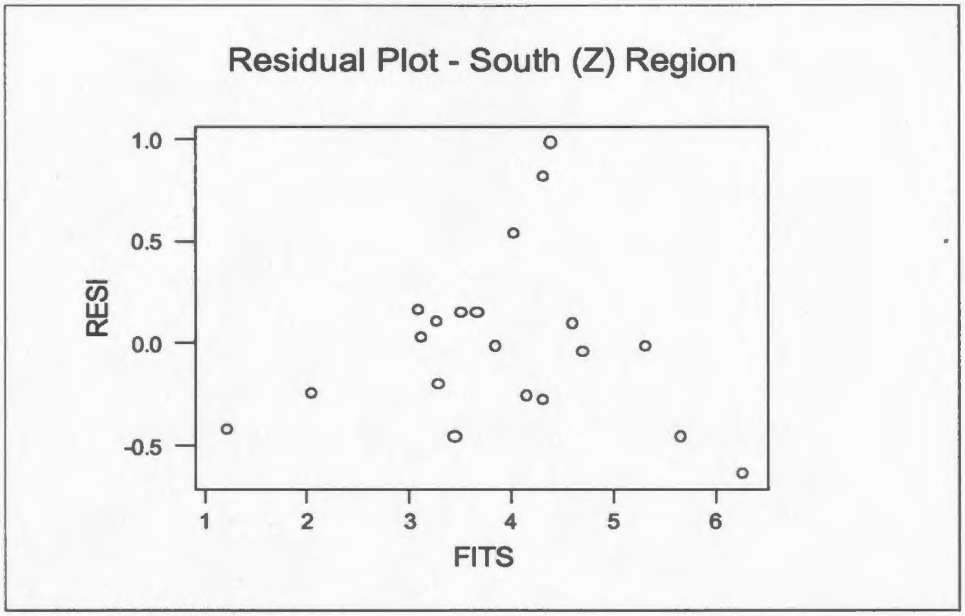


Figure 4-8 Residual and Normal Plots for North (Y) Region Regression Residuals



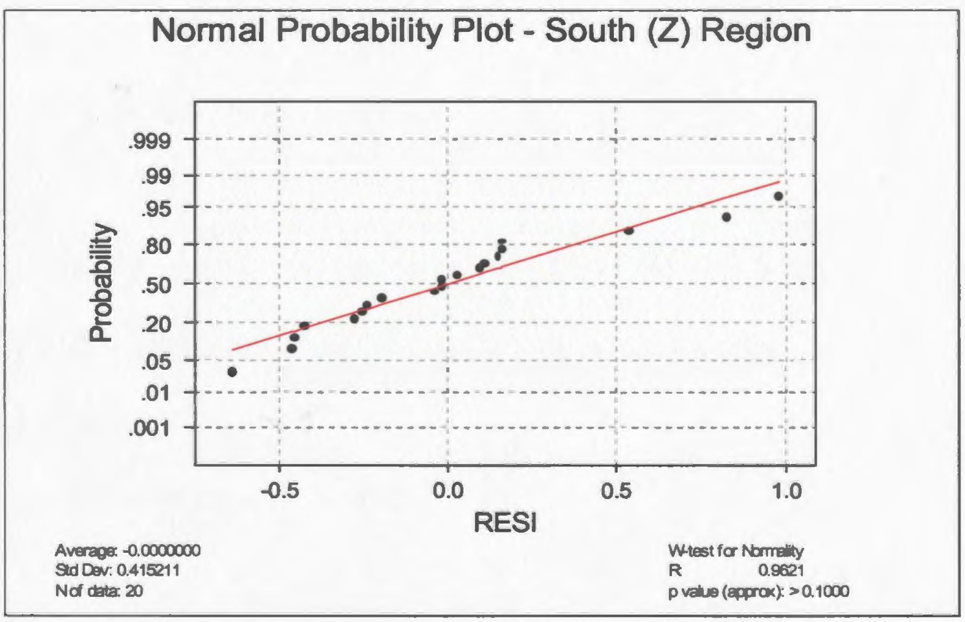


Figure 4-9 Residual and Normal Plots for South (Z) Region Regression Residuals

4.3 Out of Sample Testing

To test the success of the regionalisation technique, Newfoundland basins with 10-12 years of record were used to test the methodology developed. These basins were not used in the development of the Flood Frequency Factors, the regionalisation technique, or to derive the relationship between the drainage area and the average annual maximum daily flows. For the 18 stations with 10-12 years of data, the regionalisation technique described in the previous sections was used to estimate the desired return period flows. The results were compared to single site analysis flow estimates to determine how well the regionalisation technique using the Inverse Gaussian distribution estimates the flow at these stations. In the North (Y) Region, 11 test stations were available while in the South (Z) region 7 test stations were available.

Tables 4-7 and 4-8 provide estimates of return period flows using the regionalisation technique developed in this study and estimates from the single site analysis using the Inverse Gaussian distribution. Table 4-9 presents the single site analysis flow estimates using the top AIC ranking distribution.

Table 4-7 Test Station Flow Estimates using the Proposed Regionalisation Technique

Station	Return Period Flow Estimates Inverse Gaussian Regionalisation Technique									
	2	**5**	**10**	**20**	**50**	**100**	**200**	**500**	**1000**	**10 000**
YE1	25	32	37	40	45	49	53	57	61	72
YH1	10	13	15	16	18	20	21	23	25	29
YK3	79	100	114	127	142	154	165	179	190	225
YK7	29	37	42	46	52	56	60	66	69	82
YL4	17	21	24	27	30	32	34	38	40	47
YL5	6	7	8	9	10	11	12	13	14	16
YO7	24	30	34	38	42	46	49	53	57	67
YO8	160	203	232	257	288	312	333	363	385	455
YO10	17	22	25	28	31	34	36	39	42	49
YQ4	365	464	528	585	657	711	760	828	878	1037
YS5	343	436	496	550	618	668	714	779	825	975
ZJ2	32	43	51	57	66	73	79	88	94	114
ZK3	19	26	30	34	39	43	47	52	55	68
ZK4	42	56	66	75	86	95	103	114	122	149
ZL4	16	21	25	28	32	35	39	43	46	56
ZL5	8	10	12	14	16	17	19	21	22	27
ZM16	10	14	17	19	22	24	26	29	31	38
ZN2	10	13	15	17	20	22	24	26	28	35

Table 4-8 Test Station Flow Estimates using the Inverse Gaussian Single Site Analysis Technique

Station	Return Period Flow Estimates Single Site Analysis – Inverse Gaussian Distribution									
	2	**5**	**10**	**20**	**50**	**100**	**200**	**500**	**1000**	**10 000**
YE1	36	45	50	55	61	65	70	75	79	92
YH1	5	7	9	10	12	13	15	17	18	23
YK3	61	82	95	108	124	136	148	163	175	212
YK7	22	28	32	36	40	44	47	51	54	64
YL4	26	37	45	52	61	69	76	85	92	116
YL5	10	15	18	22	26	30	33	38	41	53
YO7	24	29	33	36	40	43	45	49	52	60
YO8	213	279	321	360	409	445	481	526	560	671
YO10	11	14	16	18	20	22	24	26	28	34
YQ4	571	797	947	1089	1272	1407	1542	1718	1850	2288
YS5	230	292	331	366	410	442	473	513	542	638
ZJ2	12	13	14	14	15	16	16	17	17	19
ZK3	29	39	45	51	58	64	69	76	81	98
ZK4	75	92	103	113	125	134	142	153	161	186
ZL4	11	14	16	18	21	23	24	27	28	34
ZL5	4	5	5	6	6	6	7	7	8	9
ZM16	10	13	14	16	18	20	22	24	25	30
ZN2	7	9	10	12	13	15	16	17	18	22

Table 4-9 Test Station Flow Estimates using the Top AIC Ranking Distribution

Station	Return Period Flow Estimates Single Site Analysis – Top AIC Distribution									
	2	**5**	**10**	**20**	**50**	**100**	**200**	**500**	**1000**	**10 000**
YE1	36	45	50	55	61	65	70	75	79	92
YH1	5	7	9	10	12	13	15	17	18	23
YK3	66	83	90	96	100	103	105	107	108	109
YK7	22	28	32	36	40	44	47	51	54	64
YL4	26	36	44	51	61	69	76	86	94	118
YL5	8	13	20	33	67	117	209	455	821	5920
YO7	21	28	39	58	110	189	335	733	1338	1010[illegible]
YO8	213	279	321	360	409	445	481	526	560	671
YO10	11	14	16	18	20	22	24	26	28	34
YQ4	571	797	947	1089	1272	1407	1542	1718	1850	2288
YS5	230	292	331	366	410	442	473	513	542	638
ZJ2	11	13	14	15	16	17	18	19	20	23
ZK3	29	39	45	51	58	64	69	76	81	98
ZK4	74	92	104	116	131	143	154	169	181	218
ZL4	11	14	16	19	24	28	32	40	46	76
ZL5	4	5	5	6	6	7	7	8	9	10
ZM16	10	13	14	16	18	20	22	24	25	30
ZN2	7	9	10	12	13	15	16	17	18	22

Combining and summarizing the results of the above three tables, Table 4-10 illustrates the median absolute percent difference between the Regional estimate and the single site frequency analysis estimate for each return period from 2 to 10 000. Table 4-11 shows the same information, with distinction between the two regions.

Table 4-10 Median Absolute Percent Difference between Regional Inverse Gaussian Technique and Top Single Site Analysis Distribution

Return Period	Absolute Percent Difference
2	36
5	42
10	45
20	46
50	49
100	50
200	51
500	51
1000	50
10000	50
Mean % Difference	47

Table 4-11 Median Absolute Percent Difference between Regional Inverse Gaussian Technique and Top Single Site Analysis, by Region

Return Period	Absolute Percent Difference	
	North (Y) Region	South (Z) Region
2	33	45
5	42	47
10	44	49
20	46	46
50	51	36
100	50	34
200	51	33
500	52	33
1000	52	32
10000	53	32
Mean % Difference	47	39

These differences may appear high. However, when compared to other studies they are of the same order of magnitude. Table 4-12 sites the Median Absolute Percent Difference between Frequency Analysis Estimates and Regression Equation Estimates for Independent sites presented in the Regional Flood Frequency Analyses for the Island of Newfoundland (Government of Newfoundland, 1999).

Table 4-12 Median Absolute Percentage Difference Between Frequency Analysis Estimates and Regression Equation Estimates for the Independent Data Sets (Government of Newfoundland and Labrador, 1999)

Region	Government of Newfoundland RFFA Analysis			Current Analysis
	1984	1990	1999	
NW	69.5	52.7	68.1	
NE	52.0	49.8	37.3	
SE	36.1	35.7	25.3	
SW	32.8	66.8	42.4	
Average	**47.6**	**51.3**	**43.3**	**43.3**
North (EC Y)				45.2
South (EC Z)				41.3

Note: Return Periods 2, 5, 10, 20, 50, 100 and 200 are used in the analysis, following that in the Government of Newfoundland (1999) Analysis.

For the 1999 RFFA analysis, the results range from 25.3% in the SE Region to 68.1% in the NW Region. The results of the current analysis fall well within that range, indicating acceptability of the Median Absolute Percent Difference.

Shown in Table 4-13 are the Growth Curve Frequency Factors developed by Pokhrel (2002) along with those developed in this study for both the North and South Environment Canada Y and Z subregions.

Table 4-13 Comparison of Flood Factors Results

Return Period (Years)	Flood Frequency Factors				Ratio POKHREL / DIGNARD	
	Y DIGNARD	Y POKHREL	Z DIGNARD	Z POKHREL	Y	Z
2	0.96	0.94	0.94	0.93	0.98	0.99
10	1.40	1.42	1.48	1.54	1.01	1.04
20	1.56	1.60	1.69	1.76	1.03	1.04
50	1.76	1.84	1.94	2.05	1.05	1.05
100	1.91	2.01	2.13	2.27	1.05	1.06
200	2.05	2.19	2.32	2.49	1.07	1.07
				Maximum	**1.07**	**1.07**
				Minimum	**0.98**	**0.98**
				Average	**1.03**	**1.03**

Note: DIGNARD – Maximum Annual Daily Flows

POKHREL – Maximum Annual Instantaneous Flows

Therefore, Ratio $_{\text{POKHREL/DIGNARD}}$ = Ratio $_{\text{INSTANTANEOUS/DAILY}}$

It is very interesting to note that these two completely independent studies, with very different approaches lead to very similar results. Pokhrel's (2002) study, using instantaneous data along with the well established Method of L-Moments, are on average just slightly higher than the results of this study using maximum annual daily flows with off-the-shelf theory for the Inverse Gaussian distribution. Although difficult to say with certainty, the difference of approximately 10%, is possibly due to the use of instantaneous flow data over daily flow data.

Overall, the results of using the Inverse Gaussian distribution for regional flood frequency analysis of Newfoundland rivers are very promising. Similarities with other studies, despite very different techniques, provides additional credibility to the methodology and application discussed in this thesis. The Inverse Gaussian distribution and the flood frequency analysis methodology presented here are ready for application by engineers and hydrologists to flood flow analysis for Newfoundland rivers.

5.0 Recommended Methodology and Application

The preceding sections proved the Inverse Gaussian distribution is suitable for flood frequency analysis. This section summarizes the regionalisation technique and recommends a methodology for applying the results of a regional flood frequency analysis to an ungauged basin to determine flood estimates of a desired return period.

5.1 Regionalisation Method Using the Inverse Gaussian Distribution

The regionalisation technique used in this study combines the properties of the Inverse Gaussian distribution with an Index Flow analysis.

Section 2 described the three steps required to apply the methodology applied in this study to another data set. In summary the three steps are as follows:

1. Test for homogeneity. Once logical groupings of the data set have been identified, each group or region must be tested for homogeneity. The

procedure for this test using the Inverse Gaussian distribution is described in Section 2.2.1.

2. Index flow analysis. For a regional flood frequency analysis, flood frequency factors are required for each homogeneous region. These can be displayed in table format or through the use of a graph. The plot is useful in providing a graphical picture of how the regional estimate compares with the individual stations used to estimate the regional curve, while the table provides the flood frequency factors at each return period without requiring interpolation from the plot. The regional flood frequency factors are calculated by averaging the normalized, or index flows for each return period. Normalization of the data is accomplished by dividing each flow ordinate by the average annual maximum daily flow for the station. The resulting 'factor' is the multiplier that is used with the mean annual maximum daily flow to estimate the desired return period flood flow. This is described in more detail in Section 2.2.2.

3. Regional estimate of Q_{avg}. For each region a method to estimate the mean annual maximum daily flow for an ungauged basin is required, using only measurable basin parameters. The relationship of flow to the basin drainage area is determined by regression analysis. This is discussed in Section 2.2.3.

5.2 Application of Regionalisation Results

The following summarizes how the results of this thesis can be used by water resource engineers to estimate flood frequency for rivers within the two regions identified by this study. This approach is especially useful for rivers where there is no flow record available, and where the flow record is too short.

To obtain estimates of maximum annual daily flows with a specified return period at an ungauged basin, the following steps should be followed.

1. Determine the drainage area of the ungauged basin.

2. Estimate the average maximum annual daily flow from Figure 4-6 or 4-7, depending on whether the ungauged basin falls within the Y (North) or Z (South) region, respectively. Equations (12) or (13) could also be used.

3. Calculate the flow for the estimated return period of interest by multiplying by the appropriate ratio from Table 4-5.

The technique described in this study is useful for providing an estimate of the flood flow at an ungauged basin for a desired return period. The results of this technique are particularly useful for providing a first estimate of a desired return period flow at an ungauged basin as well as for verifying that an alternate approach provides reasonable flow estimates.

6.0 Conclusions and Recommendations

6.1 Conclusions

The following conclusions arise from this study:

- The main conclusion of this study is that for Newfoundland hydrological data, the Inverse Gaussian distribution is very well suited to annual maximum daily flows, but performs very poorly with annual minimum daily flows.

- The 2-parameter Inverse Gaussian distribution is a flexible statistical distribution that is capable of representing a variety of data sets, from very symmetrical to highly skewed data.

- Using the AIC as an indicator of suitability, for high flow analysis the Inverse Gaussian is better than other 2-parameter distributions as well as the 3-parameter distributions considered in this analysis.

- The Inverse Gaussian regional flood factors developed in this study can be used to provide or verify estimates of flood flows of a specified return period at an ungauged watershed on the island of Newfoundland. Given the uncertainty in any statistical analysis results, and with all regionally calculated flood flow estimates, care should be taken and engineering judgement used when applying the results of this analysis to engineering projects and designs.

6.2 Recommendations

The recommendations from this study are as follows.

- Apply the Inverse Gaussian distribution to streamflow data in areas other than the island of Newfoundland. Such application of the distribution to both high and low flows for other areas would determine if the suitability for high flows, or the inappropriateness for low flows, is limited to Newfoundland streamflow data.

- Develop a method of estimating the inverse of the Inverse Gaussian distribution, such as an approximation of the Z-statistic used for estimating the inverse of the Normal (Gaussian) distribution. This would be extremely useful in future application of the Inverse Gaussian distribution to Newfoundland streamflow data as well as to other suitable data sets. With the inverse of the Inverse Gaussian distribution, an equation for the regional growth curve could be developed. The region would have a mean of 1 (due to using Q/Qavg) and a Regional λ which follows the Inverse Gaussian distribution. The estimated Regional λ is easily determined by the weighted average of all λ's within the region.

- Update the analysis in the future to determine how additional data impacts the results of the regionalisation using the Inverse Gaussian distribution.

7.0 References

Baker, Calvin, Personal Communication, Environment Canada, January 2003.

Balakrishnan, N and Chen, William W.S., 1997, "Tables for Order Statistics from Inverse Gaussian Distributions with Applications", CRC Press Inc.

Chhikara, Raj. S. and Folks, Leroy J., 1989, "The Inverse Gaussian Distribution - Theory, Methodology, and Applications", Marcel Dekker Inc., New York and Basel

Chow, K.C. Ander and Watt, W. Edgar, "Use of Akaike information criterion for selection of flood frequency distribution", Canadian Journal of Civil Engineering, 1992, Volume 19, pp 616-626

Chow, V.T. (Ed), 1964, "Handbook of Hydrology", McGram Hill, New York, USA.

Condie, R. et al, "Comparison of Regional Flood Frequency Methods in Southern Ontario using Analysis of Variance Techniques", Regional Frequency Analysis – Proceedings of the International Symposium on Flood Frequency Analysis, May 1986, pp 213-222

Cunnane Conleth, "Review of Statistical Models for Flood Frequency Estimation", Regional Frequency Analysis – Proceedings of the International Symposium on Flood Frequency Analysis, May 1986, pp 49-95

Devore, Jay L., 1991, "Probability and Statistics for Engineering and the Sciences" Third Edition, Brooks/Cole Publishing Company, Pacific Grove, California, USA

Environment Canada, 1993, "Consolidated Frequency Analysis", Version 3.1

Environment Canada, 1988, "Consolidated Frequency Analysis", 1988 Version

Environment Canada, 1996, "Hydat for Windows"

Farquharson, F.A.K. et al, "Comparison of Flood Frequency Curves for Many Different Regions of the World", Regional Frequency Analysis – Proceedings of the International Symposium on Flood Frequency Analysis, May 1986, pp 223-256

Government of Newfoundland and Labrador, 1984, "Regional Flood Frequency Analysis for the Island of Newfoundland", Newfoundland Flood Damage Reduction Program, Department of Environment and Lands

Government of Newfoundland and Labrador, 1990, "Regional Flood Frequency Analysis for the Island of Newfoundland", Government of Newfoundland and Labrador, Department of Environment and Lands, Water Resources Division.

Government of Newfoundland and Labrador, 1999, "Regional Flood Frequency Analysis for the Island of Newfoundland", Government of Newfoundland and Labrador, Department of Environment and Lands, Water Resources Division.

Hebson, C.S., Cunnane, C., "Assessment of use of At-Site and Regional Flood Data for Flood Frequency Estimation", Regional Frequency Analysis – Proceedings of the International Symposium on Flood Frequency Analysis, May 1986, pp 433-448

Hosking, J.R.M. and Wallis, J.R., 1997, "Regional Frequency Analysis – An Approach Based on L-Moments", Cambridge University Press, USA

Jorgensen, Bent, 1982, "Lecture Notes in Statistics", Statistical Properties of the Generalized Inverse Gaussian Distribution, Springer-Verlag New York Inc

Johnson, Norman and Kotz, Samuel, 1970, "Distributions in Statistics: Continuous Univariate Distributions - 1", Houghton Mifflin Company, USA

Kite, G.W., 1977, "Frequency and Risk Analysis in Hydrology" A Water Resources Publications

Koziol, James A., 1989, "Handbook of Percentage Points of the Inverse Gaussian Distribution", CRC Press Inc.

Lawal, Semiu A. and Watt, W. Edgar, "Frequency analysis of low flows using the Akaike information criterion", Canadian Journal of Civil Engineering, 1996, Volume 23, pp 1180-1189

Lawal, Semiu A. and Watt, W. Edgar, "Non-Zero Lower Limit in Low Flow Frequency Analysis", Water Resources Bulletin, 1996, Volume 32, pp 1159-1166

Minitab Inc., 1994, "Minitab", Release 10.2, Minitab Inc., Pennsylvania, USA

Nathan, R.J. and McMahon, T.A., "Identification of Homogeneous Regions for the Purposes of Regionalisation", Journal of Hydrology, 1990, pp 217-238

Palisade Corporation, 1993-96, "BestFit for Windows", Version 2.0d, Palisade Corporation

Pokhrel, Jhapendra, 2002, "Regional Flood Frequency Analysis for the Island of Newfoundland, Canada using L-Moments", Master's Thesis, Faculty of Engineering and Applied Science, Memorial University of Newfoundland, Canada

Potter, Kenneth and Walker, John, "An Empirical Study of Flood Measurement Error", Water Resources Research, 1985, Volume 21, Number 3, pp 403-406

Quick, Michael C., "Reliability of Flood Discharge Estimates", Canadian Journal of Civil Engineering, August 1991, Volume 18, Number 4, pp 624-630

Ribeiro, Joseph and Rousselle, Jean, "Robust Simple Scaling Analysis of Flood Peak Series", Canadian Journal of Civil Engineering, August 1996, Volume 23, Number 6, pp 1139-1145

Richter, S.H., 1994, "Relationships of Flow and Basin Variables on the Island of Newfoundland, Canada with a Regional Application", Master's Thesis, Faculty of Engineering and Applied Science, Memorial University of Newfoundland, Canada

Quick, Michael C., "Reliability of Flood Discharge Estimates: Reply", Canadian Journal of Civil Engineering, December 1992, Volume 19, Number 6, pp 1088-1090

Sceviour, Edward and Lye, Leonard M. Dr., "Hydrometric Data Analysis for the Purposes of Regionalisation"

Sellars, C. David and Rodman Richard F., "Reliability of Flood Discharge Estimates: Discussion", Canadian Journal of Civil Engineering, December 1992, Volume 19, Number 6, pp 1087-1088

Sheshadri, V., 1993, "The Inverse Gaussian Distribution - A Case Study in Exponential Families", Oxford Science Publications

Smith, C.D., "Reliability of Flood Discharge Estimates: Discussion", Canadian Journal of Civil Engineering, December 1992, Volume 19, Number 6, pp 1085-1087

Van Thanh Van Nguyen et al, "A Scaling Approach to Regional Estimation of Extreme Hydrologic Variables", Society for Civil Engineering Annual Conference, May 1997, Volume 3, pp. 81-90

Wasan, M.T., "First Passage Time Distribution of Brownian Motion with Positive Drift", Department of Mathematics, Queen's University, Kingston, Ontario.

Zanakis, Stelis H., "A Simulation Study of Some Simple Estimators for the Three-Parameter Weibull Distribution", Journal of Computer Simulations, 1979, Volume 9, pp 101-116

Appendix A

Single Site Analysis AIC Results – High Flow Analysis

Table A-1	AIC Values for Fitted Distributions - High Flow Analysis								
	WSC Station	AIC Value					Maximum AIC	Minimum AIC	Maximum Difference
		3PLN	GEV	EV	IG	LN			
1	YA1	183.6	184.1	185.2	185.1	185.2	185.2	183.6	1.6
2	YC1	400.2	400.3	399.4	399.2	399.3	400.3	399.2	1.0
3	YD1	184.7	184.8	182.8	182.8	182.9	184.8	182.8	2.0
4	YD2	115.4	115.4	113.5	113.7	113.7	115.4	113.5	1.9
5	YE1	86.4	86.3	84.5	84.3	84.3	86.4	84.3	2.1
6	YF1	158.6	159.7	157.3	157.4	157.5	159.7	157.3	2.4
7	YH1	48.8	48.5	49.2	47.3	47.3	49.2	47.3	1.9
8	YJ1	291.4	291.7	289.9	289.3	289.6	291.7	289.3	2.4
9	YK2	204.5	205.3	203.4	203.7	203.8	205.3	203.4	1.9
10	YK3		102.2	103.5	102.6	102.6	103.5	102.2	1.4
11	YK4	212.2	211.8	211.1	210.1	210.2	212.2	210.1	2.1
12	YK5		199.9	207.5	202.9	202.4	207.5	199.9	7.5
13	YK7	78.6	78.5	76.8	76.6	76.6	78.6	76.6	2.0
14	YL1	822.7	822.6	822.7	820.6	820.7	822.7	820.6	2.1
15	YL4	97.1	97.0	96.4	95.3	95.2	97.1	95.2	1.8
16	YL5	66.8	59.1	65.8	62.7	63.0	66.8	59.1	7.7
17	YM3	130.1	130.1	128.5	128.9	128.8	130.1	128.5	1.7
18	YN2	156.6	156.7	154.7	154.6	154.7	156.7	154.6	2.1
19	YO10	57.6	58.1	57.0	56.7	56.8	58.1	56.7	1.4
20	YO6	123.0	122.5	127.3	123.9	123.8	127.3	122.5	4.8
21	YO7	76.8	72.0	74.5	74.7	74.8	76.8	72.0	4.8
22	YO8	130.2	130.1	128.3	128.0	128.1	130.2	128.0	2.1
23	YP1	95.2	95.1	93.2	93.3	93.2	95.2	93.2	2.0
24	YQ1	591.6	591.9	592.3	590.8	590.5	592.3	590.5	1.8
25	YQ4	171.1	170.6	168.5	167.8	168.0	171.1	167.8	3.3
26	YR1	256.6	256.8	255.0	254.5	254.7	256.8	254.5	2.3
27	YR2	161.3	161.6	160.5	160.7	160.8	161.6	160.5	1.1
28	YR3		117.4	118.4	117.1	117.2	118.4	117.1	1.3
29	YS1	333.8	333.6	332.1	332.0	331.8	333.8	331.8	1.9
30	YS3	139.3	141.8	141.6	141.5	141.7	141.8	139.3	2.6
31	YS5		116.3	116.7	115.9	115.9	116.7	115.9	0.9
32	ZA1	165.8	165.5	164.3	163.6	163.8	165.8	163.6	2.2
33	ZA2	110.9	111.0	109.0	108.8	108.9	111.0	108.8	2.2
34	ZA3	136.2	136.7	137.3	135.5	135.7	137.3	135.5	1.8
35	ZB1	373.5	374.3	376.5	373.4	373.8	376.5	373.4	3.1
36	ZC2	157.5	157.7	155.8	155.4	155.5	157.7	155.4	2.4
37	ZE1		245.9	247.7	246.1	246.2	247.7	245.9	1.9
38	ZF1	486.7	479.0	484.0	480.1	479.6	486.7	479.0	7.7
39	ZG1	314.9	315.6	316.2	315.1	315.3	316.2	314.9	1.3
40	ZG2	159.7	159.5	158.6	157.9	157.9	159.7	157.9	1.9
41	ZG3	131.4	131.2	130.4	130.0	130.0	131.4	130.0	1.5
42	ZG4	103.2	100.8	101.7	100.4	100.4	103.2	100.4	2.8
43	ZH1	485.6	485.9	484.2	483.3	483.6	485.9	483.3	2.6
44	ZH2	191.0	191.0	189.6	189.4	189.4	191.0	189.4	1.6
45	ZJ1	122.2	122.1	120.4	120.3	120.2	122.2	120.2	2.0
46	ZJ2	48.4	48.4	46.7	47.9	47.9	48.4	46.7	1.7
47	ZK1	457.9	456.6	461.1	457.9	458.0	461.1	456.6	4.5
48	ZK2	138.9	139.0	137.1	136.8	136.9	139.0	136.8	2.2
49	ZK3	94.5	94.9	93.3	93.0	93.0	94.9	93.0	1.9
50	ZK4	110.1	110.3	108.3	108.4	108.4	110.3	108.3	2.0
51	ZL3		69.0	72.5	70.4	70.4	72.5	69.0	3.5
52	ZL4	67.3	67.0	68.3	67.3	67.3	68.3	67.0	1.3
53	ZL5	30.7	30.9	28.9	29.0	29.1	30.9	28.9	2.0
54	ZM16	67.0	67.2	65.2	65.0	65.2	67.2	65.0	2.2
55	ZM6	88.6	87.7	87.9	87.5	87.1	88.6	87.1	1.5
56	ZM9	100.7	100.9	99.7	100.3	100.3	100.9	99.7	1.2
57	ZN1	201.6	201.1	199.5	199.0	199.1	201.6	199.0	2.6
58	ZN2	50.1	51.1	48.7	48.5	48.6	51.1	48.5	2.6
							Maximum Difference =		7.679
							Minimum Difference =		0.852
							Average Difference =		2.409

Appendix B

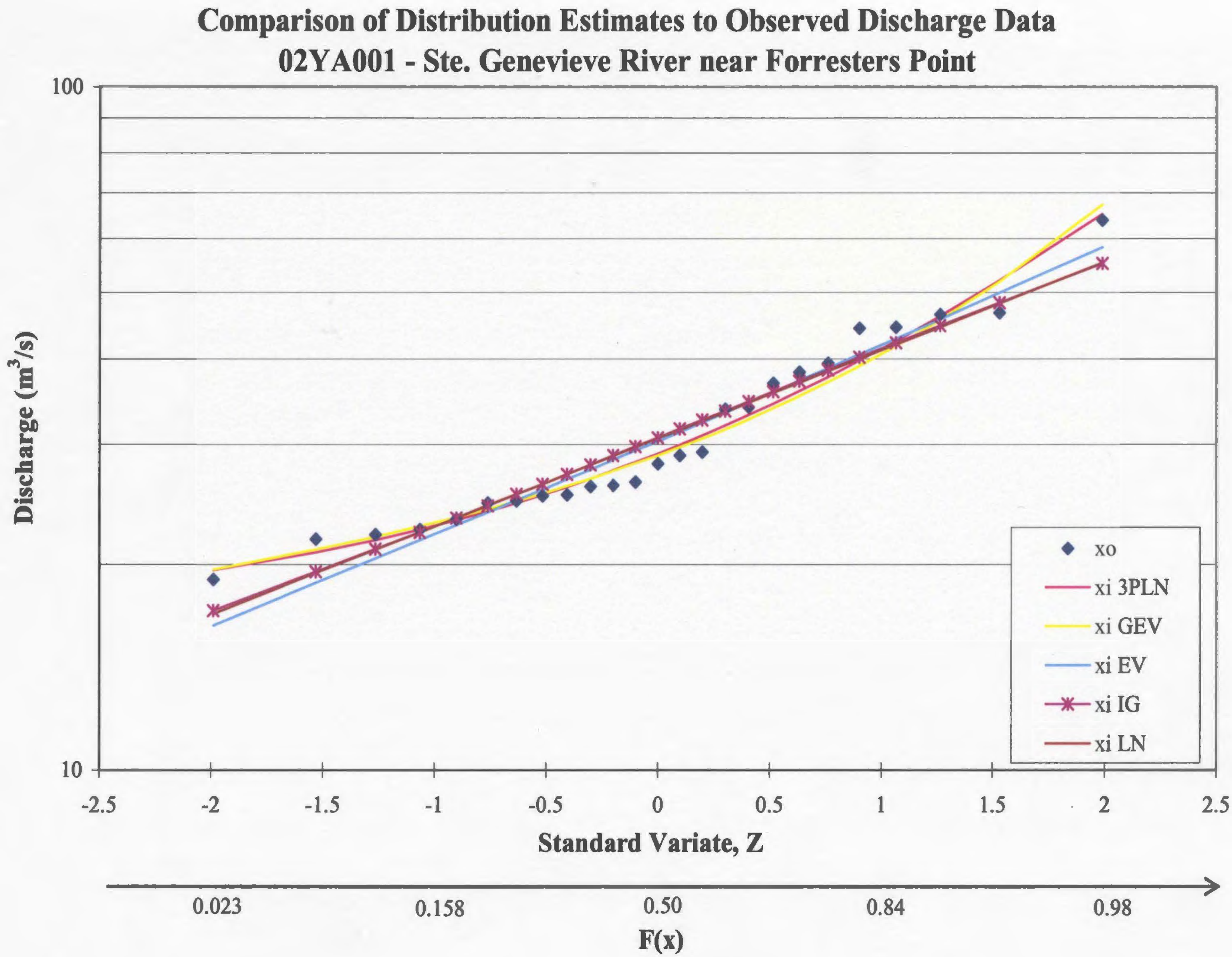
Single Site Analysis AIC Results – Low Flow Analysis

Table B-1 AIC Values for Fitted Distributions - Low Flow Analysis

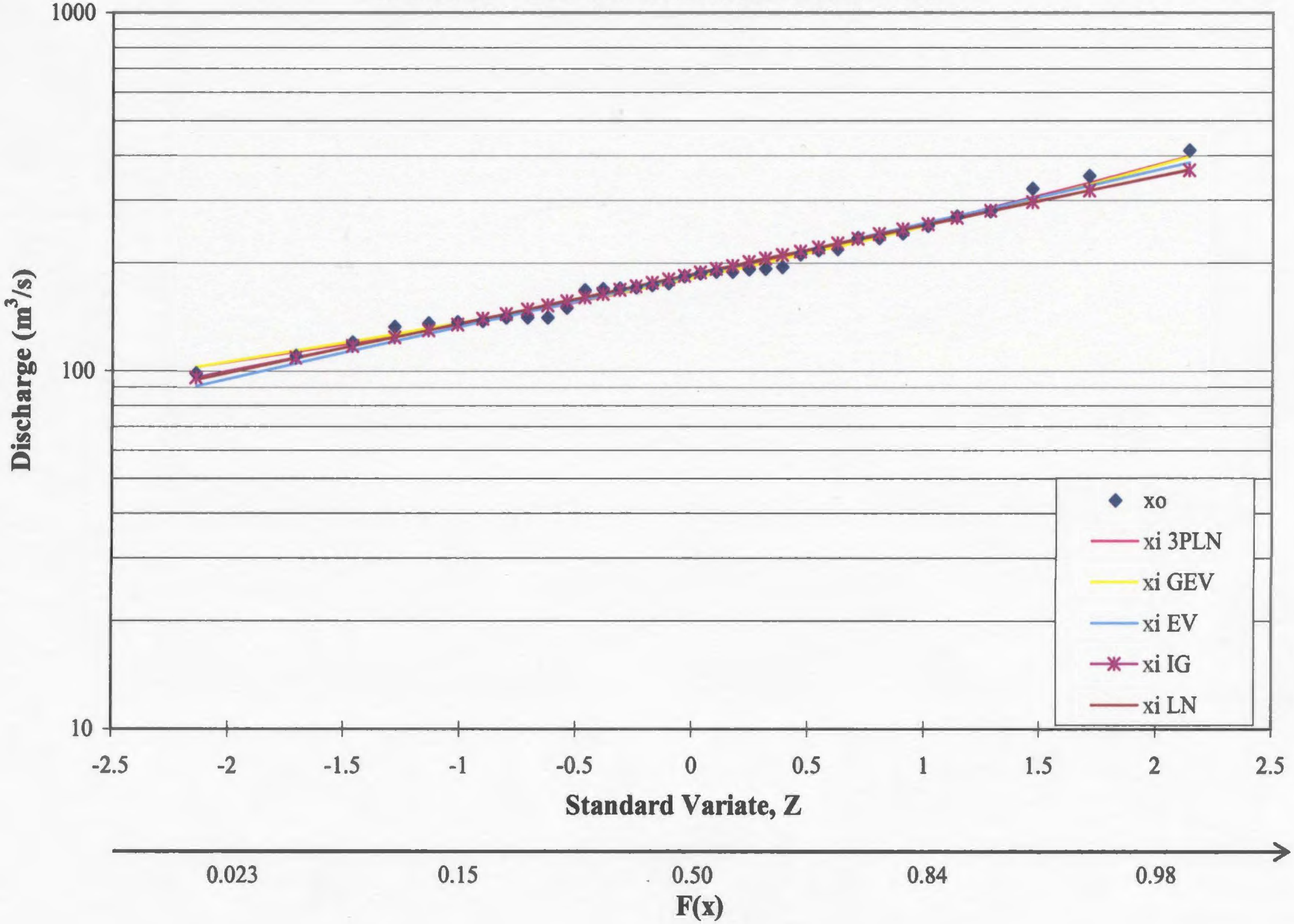
	WSC Station	AIC Value W2	W3	EV	IG	LN	Maximum AIC	Minimum AIC	Maximum Difference
1	YA1	72.3	72.2	69.0	69.0	69.0	72.3	69.0	3.4
2	YC1	121.7	124.0	125.5	127.2	125.9	127.2	121.7	5.5
3	YD1	15.2	14.6	15.0	13.3	13.4	15.2	13.3	1.9
4	YD2	5.3	7.2	7.0	9.9	7.5	9.9	5.3	4.6
5	YJ1	102.3	103.8	102.1	102.4	102.2	103.8	102.1	1.7
6	YK2	52.4	63.1	60.4	57.2	56.8	63.1	52.4	10.6
7	YK5	47.2	50.4	47.5	48.2	48.0	50.4	47.2	3.2
8	YL1	353.6	358.1	349.2	354.5	352.5	358.1	349.2	8.9
9	YM3	-41.8	-39.8	-38.6	-31.1	-37.9	-31.1	-41.8	10.7
10	YQ1	332.6	341.2	339.6	348.4	344.9	348.4	332.6	15.8
11	YR1	62.2	66.3	66.2	76.2	71.7	76.2	62.2	14.0
12	YR2	24.7	26.4	27.7	31.7	28.7	31.7	24.7	7.0
13	YS1	161.2	165.2	164.4	165.6	165.3	165.6	161.2	4.4
14	YS3	-94.4	-90.5	-94.5	-92.5	-93.6	-90.5	-94.5	4.0
15	ZA1	10.6	15.6	20.5	15.3	15.3	20.5	10.6	9.9
16	ZB1	30.4	28.7	28.6	27.3	27.7	30.4	27.3	3.1
17	ZE1	161.5	163.4	162.9	164.4	163.9	164.4	161.5	2.9
18	ZF1	238.9	246.3	248.5	257.0	252.3	257.0	238.9	18.1
19	ZG1	66.2	73.6	71.4	96.4	84.8	96.4	66.2	30.2
20	ZG2	24.7	30.2	30.0	30.1	29.8	30.2	24.7	5.5
21	ZG3	-8.4	-6.6	-8.1	-7.1	-7.7	-6.6	-8.4	1.8
22	ZH1	143.8	152.0	144.4	183.2	164.5	183.2	143.8	39.4
23	ZH2	-66.1	-60.6	-65.4	-53.9	-59.9	-53.9	-66.1	12.2
24	ZJ1	-38.3	-36.3	-35.9	-30.9	-34.6	-30.9	-38.3	7.4
25	ZK1	85.3	88.3	86.0	89.7	88.4	89.7	85.3	4.4
26	ZK2	-1.1	0.0	-2.0	-1.1	-1.4	0.0	-2.0	2.1
27	ZL3	-78.1	-75.6	-77.2	-78.9	-79.6	-75.6	-79.6	4.0
28	ZM9	-11.1	-7.3	-6.5	-6.7	-7.0	-6.5	-11.1	4.6
29	ZN1	-25.3	-21.7	-19.0	-20.9	-21.4	-19.0	-25.3	6.3
							Maximum Difference =		39.388
							Minimum Difference =		1.718
							Average Difference =		8.537

Appendix C

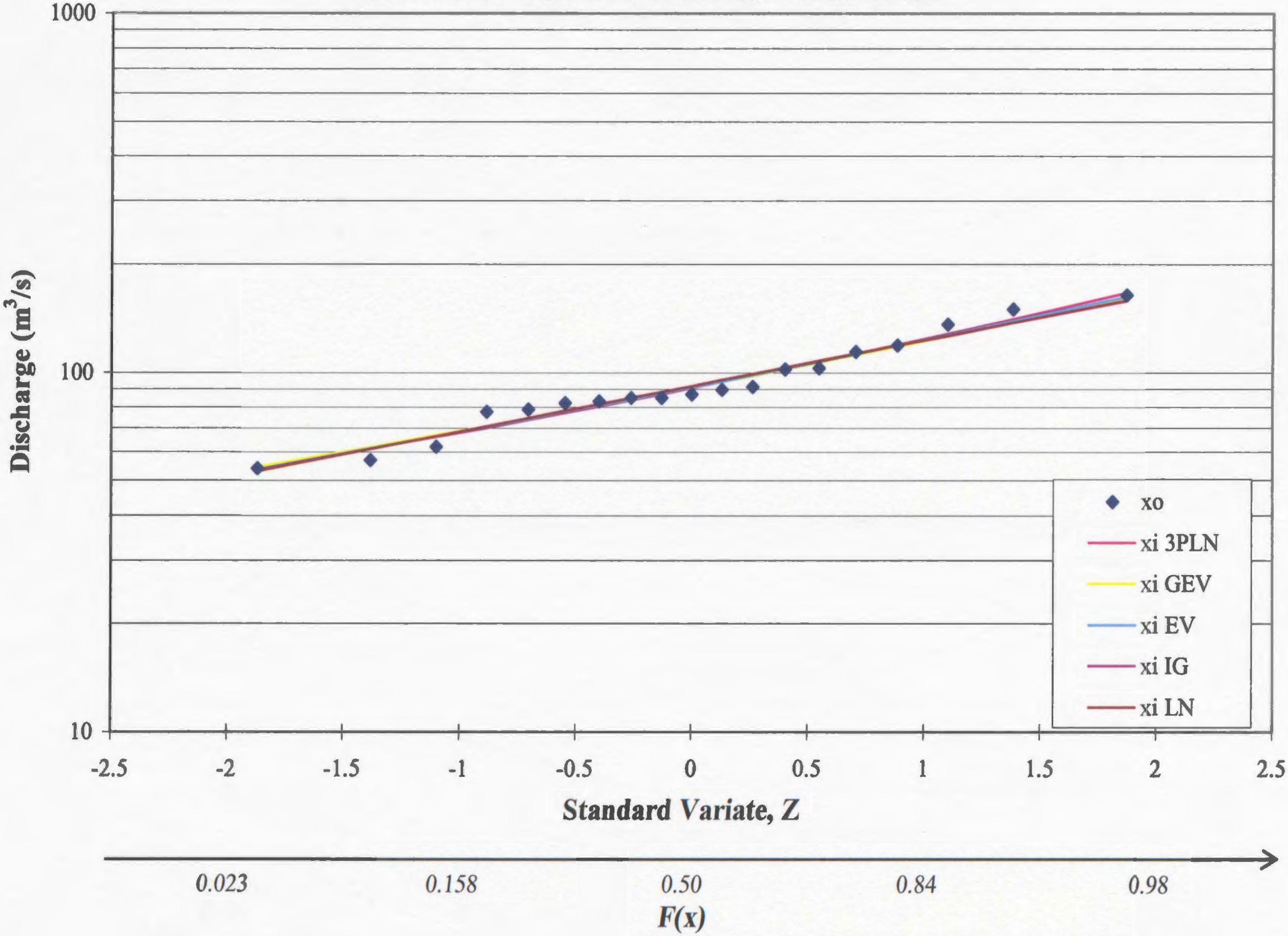
Single Site Analysis Results – High Flow Analysis Data Plots

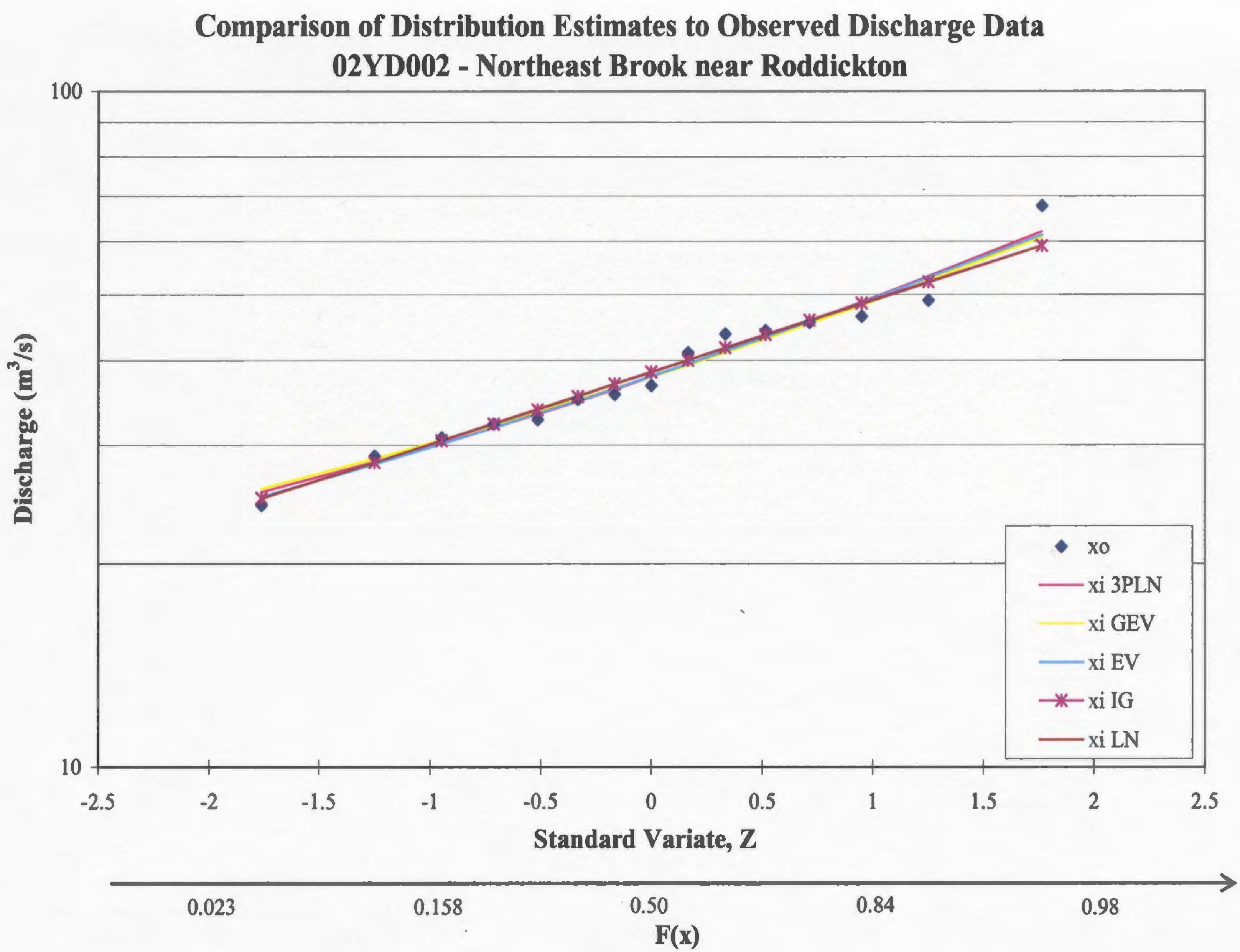

Comparison of Distribution Estimates to Observed Discharge Data
02YA001 - Ste. Genevieve River near Forresters Point
100
10
Discharge (m3/s)
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN

Comparison of Distribution Estimates to Observed Discharge Data
02YC001 - Torrent River at Bristol's Pool



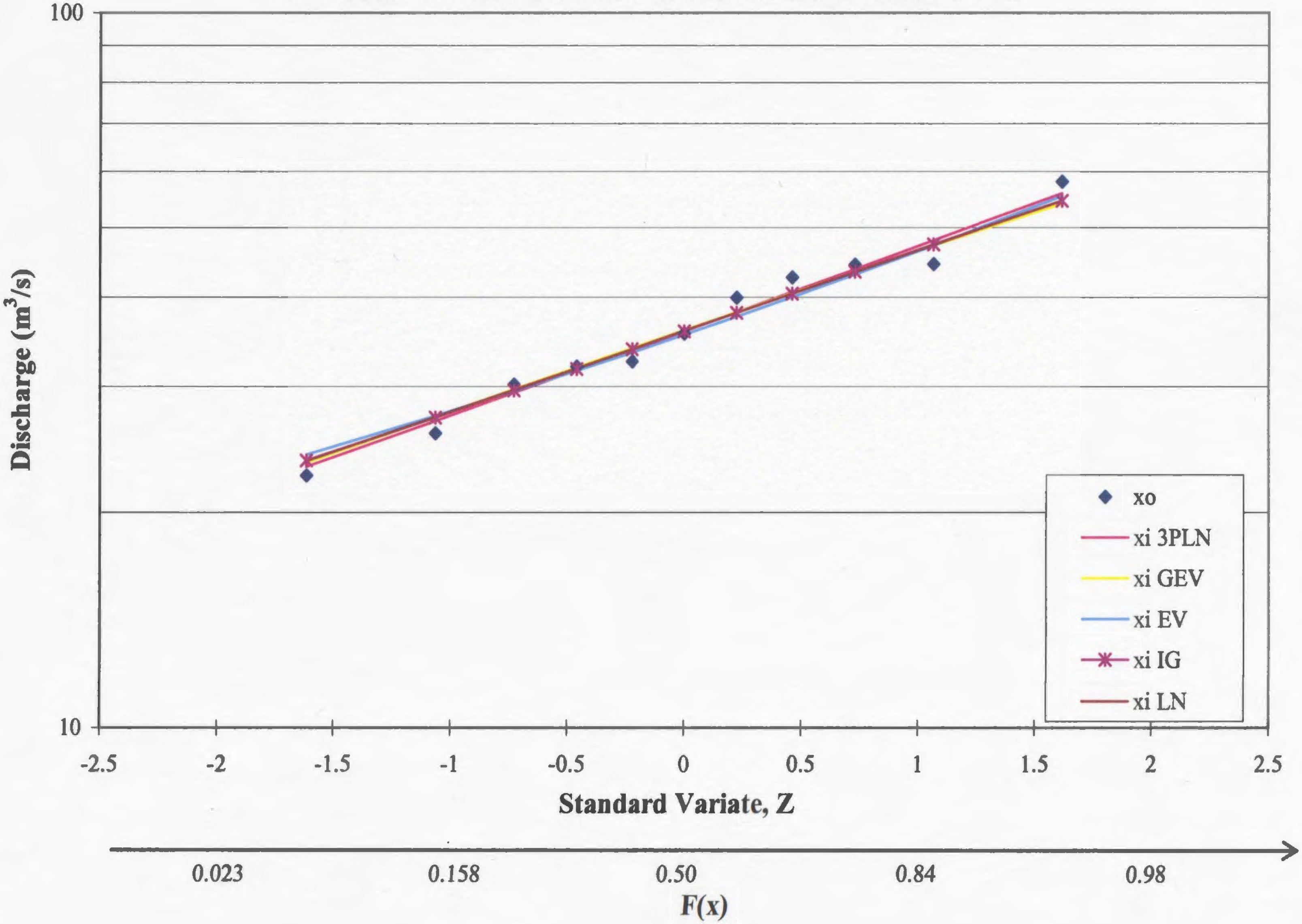
Comparison of Distribution Estimates to Observed Discharge Data
02YD001 - Beaver Brook near Roddickton

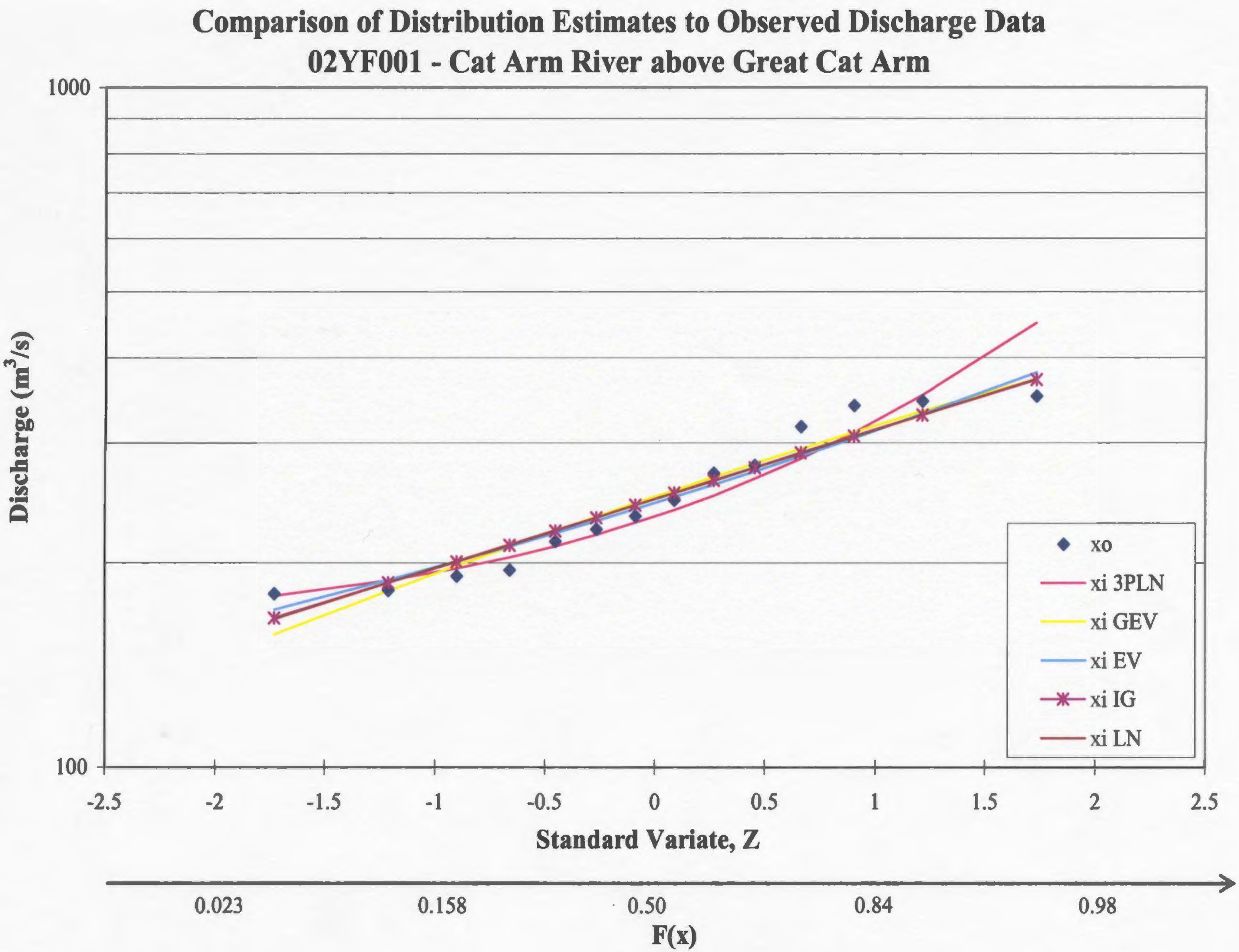




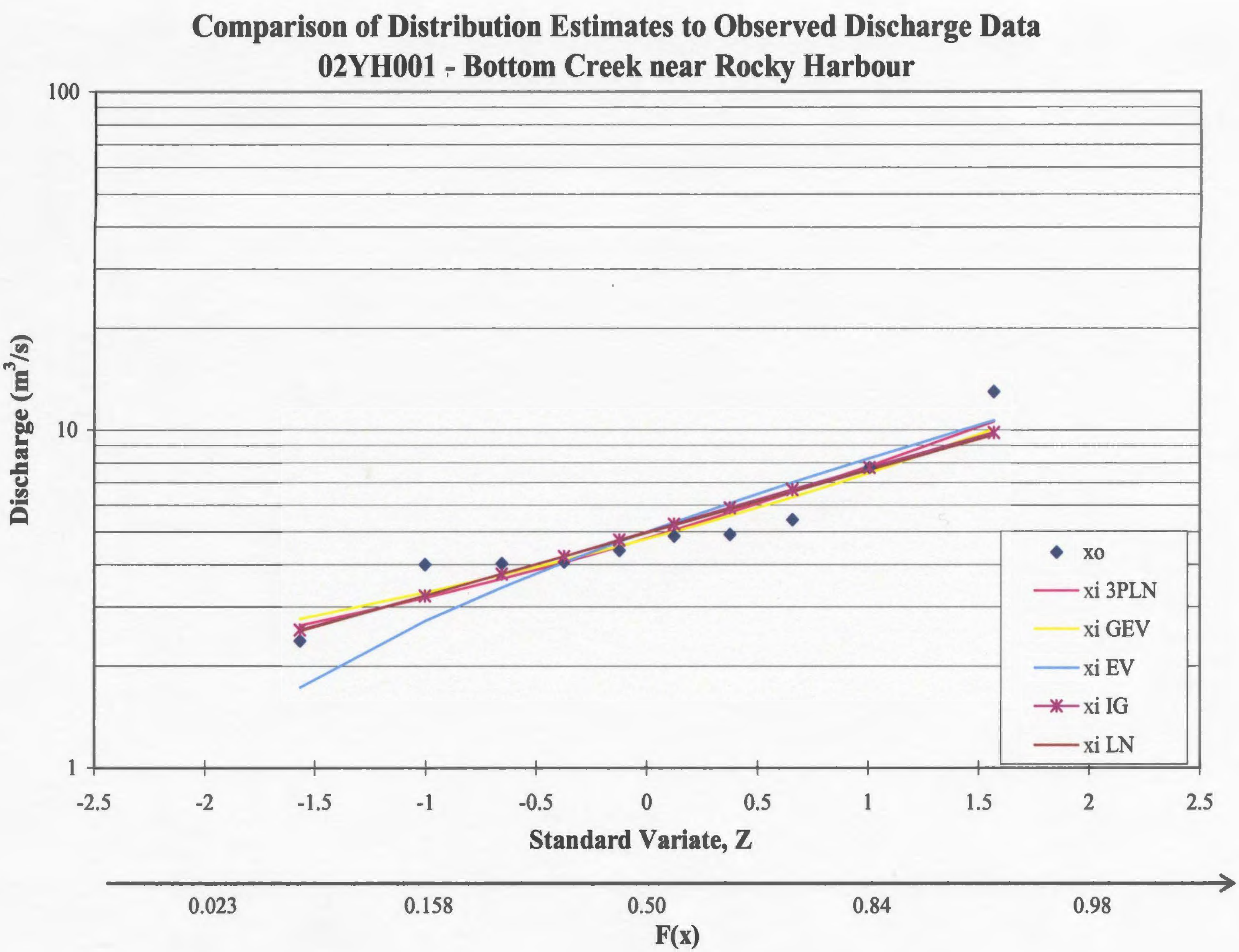
Comparison of Distribution Estimates to Observed Discharge Data
02YD002 - Northeast Brook near Roddickton
Discharge (m³/s)
100
10
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN

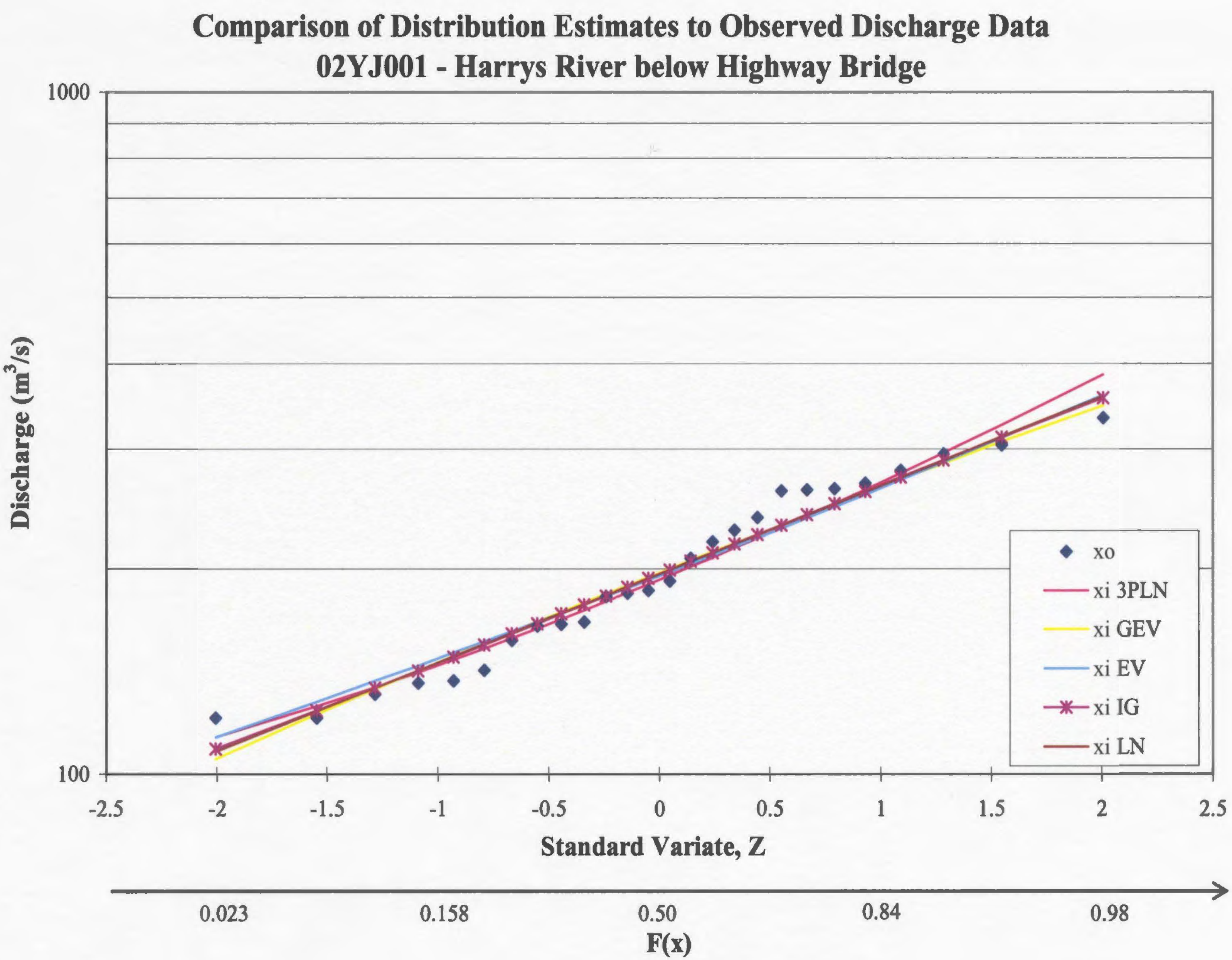
Comparison of Distribution Estimates to Observed Discharge Data
02YE001 - Greavett Brook above Portland Creek Pond



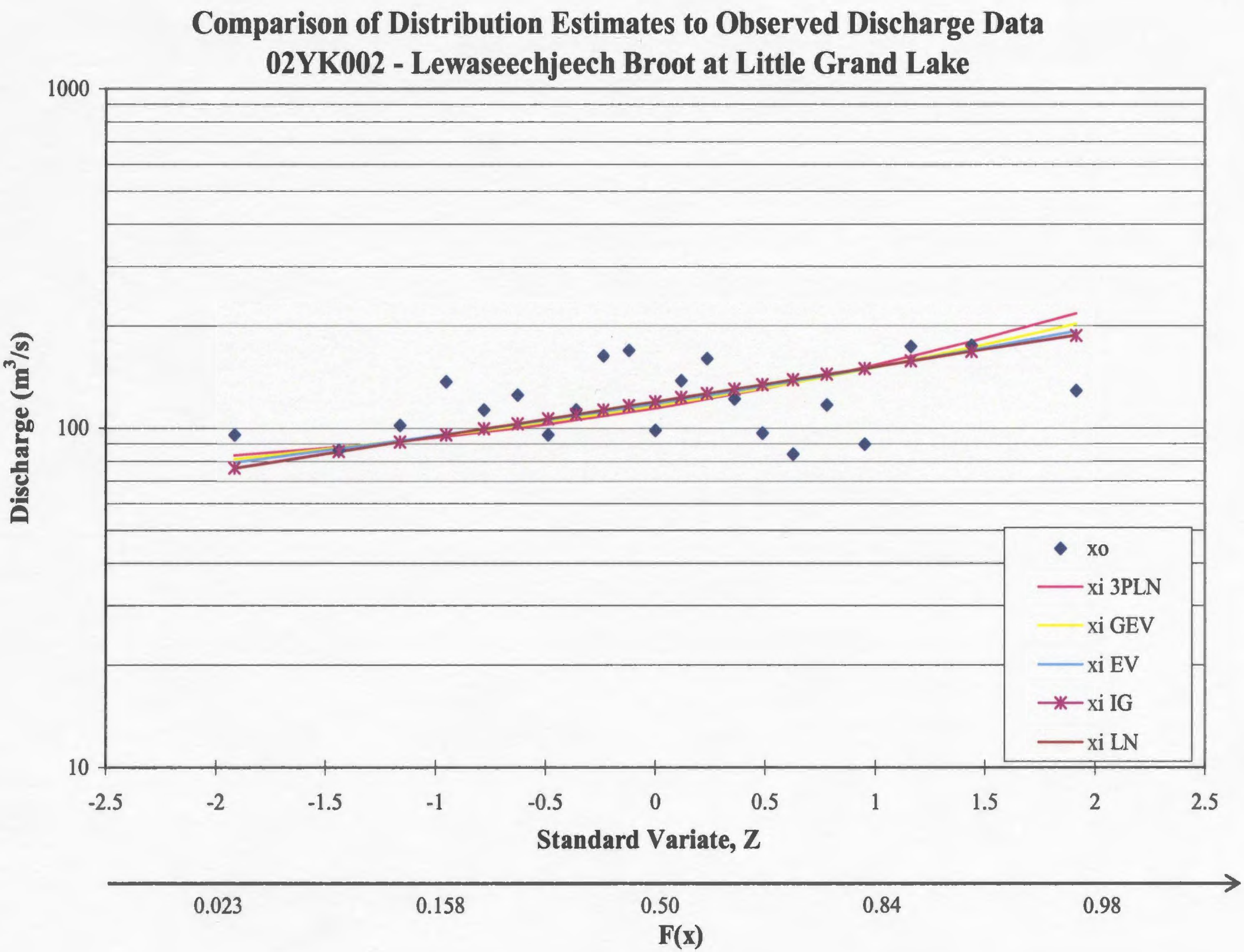


Comparison of Distribution Estimates to Observed Discharge Data
02YF001 - Cat Arm River above Great Cat Arm
1000
100
Discharge (m³/s)
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN

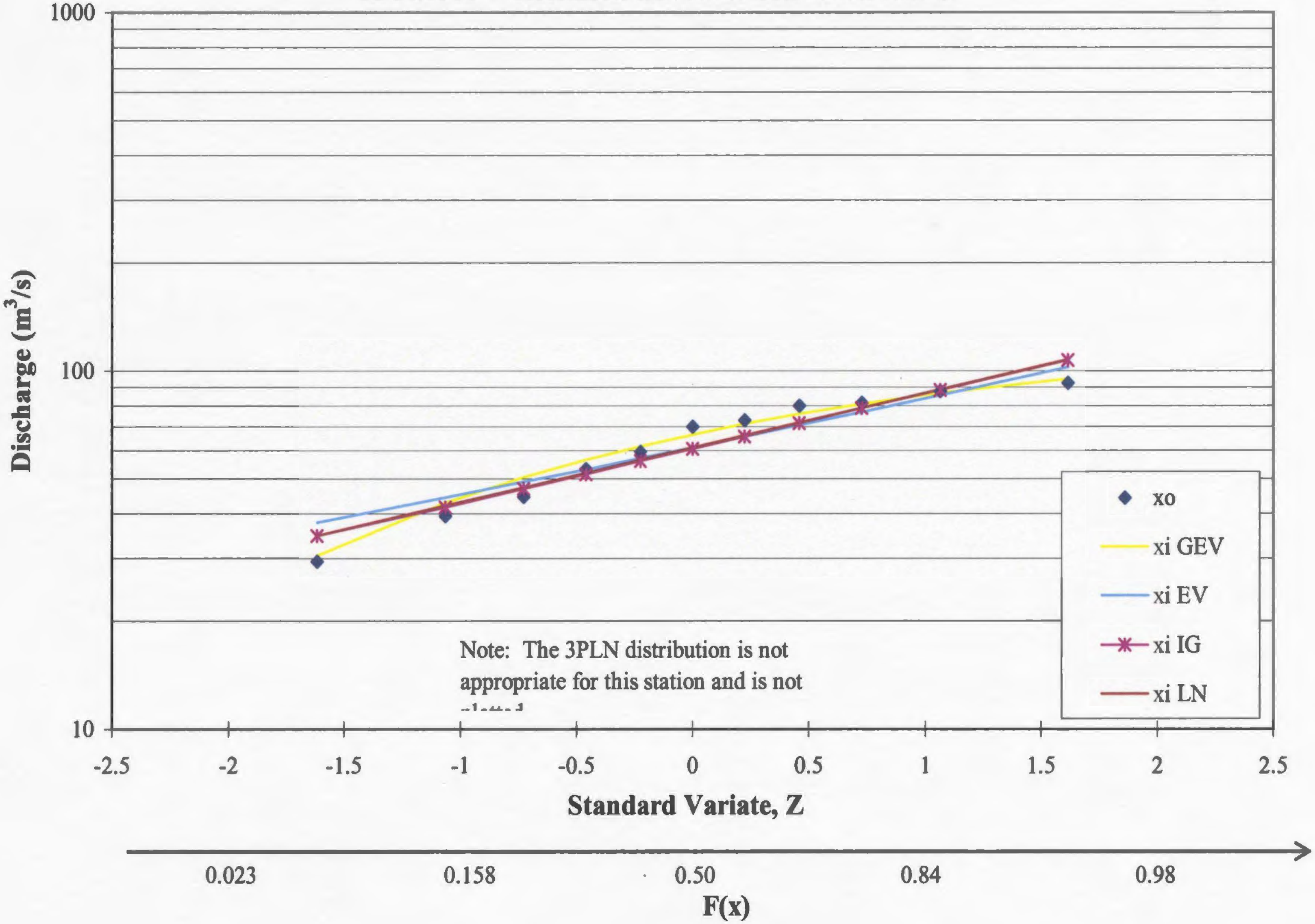

Comparison of Distribution Estimates to Observed Discharge Data
02YH001 - Bottom Creek near Rocky Harbour
Discharge (m^3/s)
100
10
1
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN

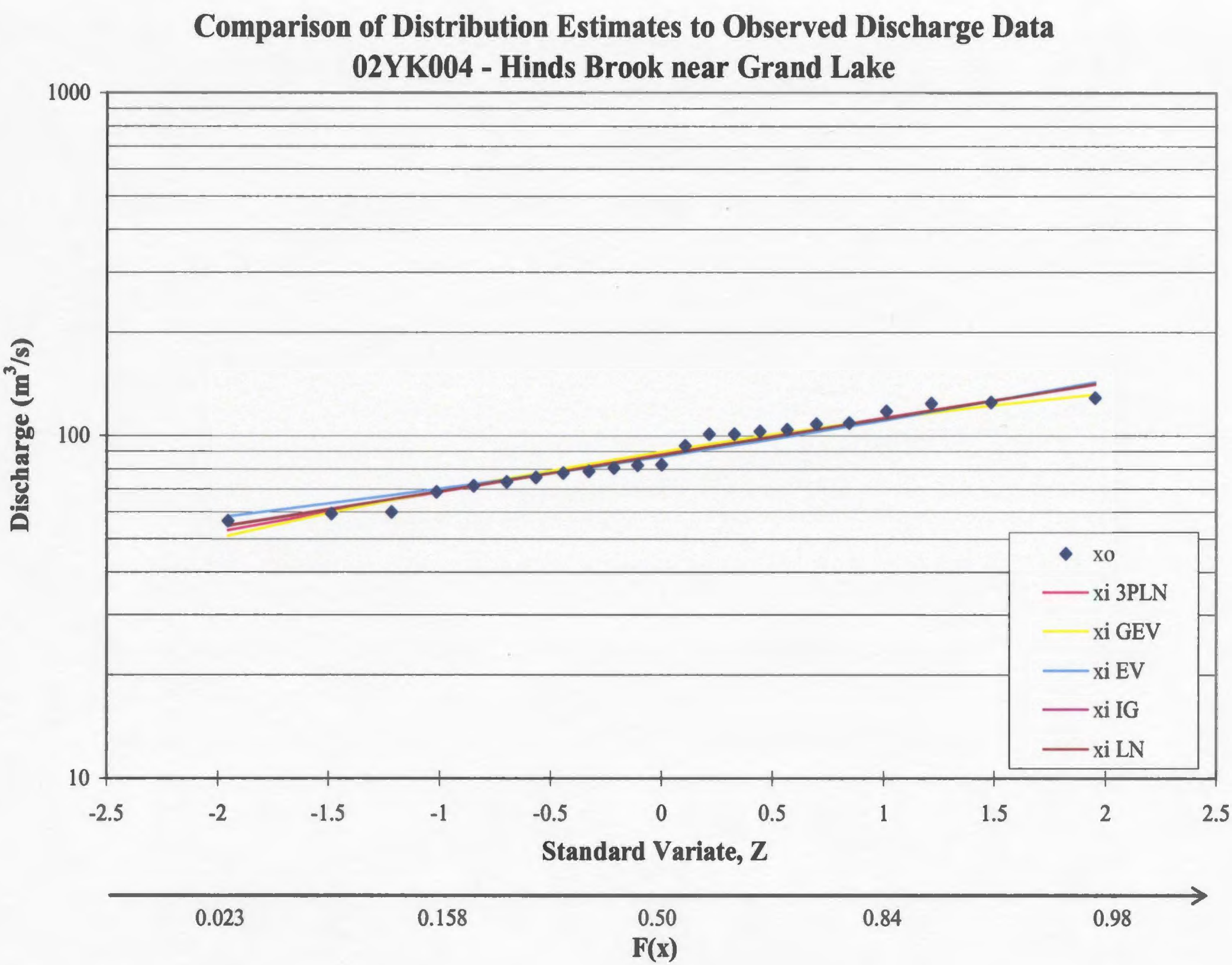


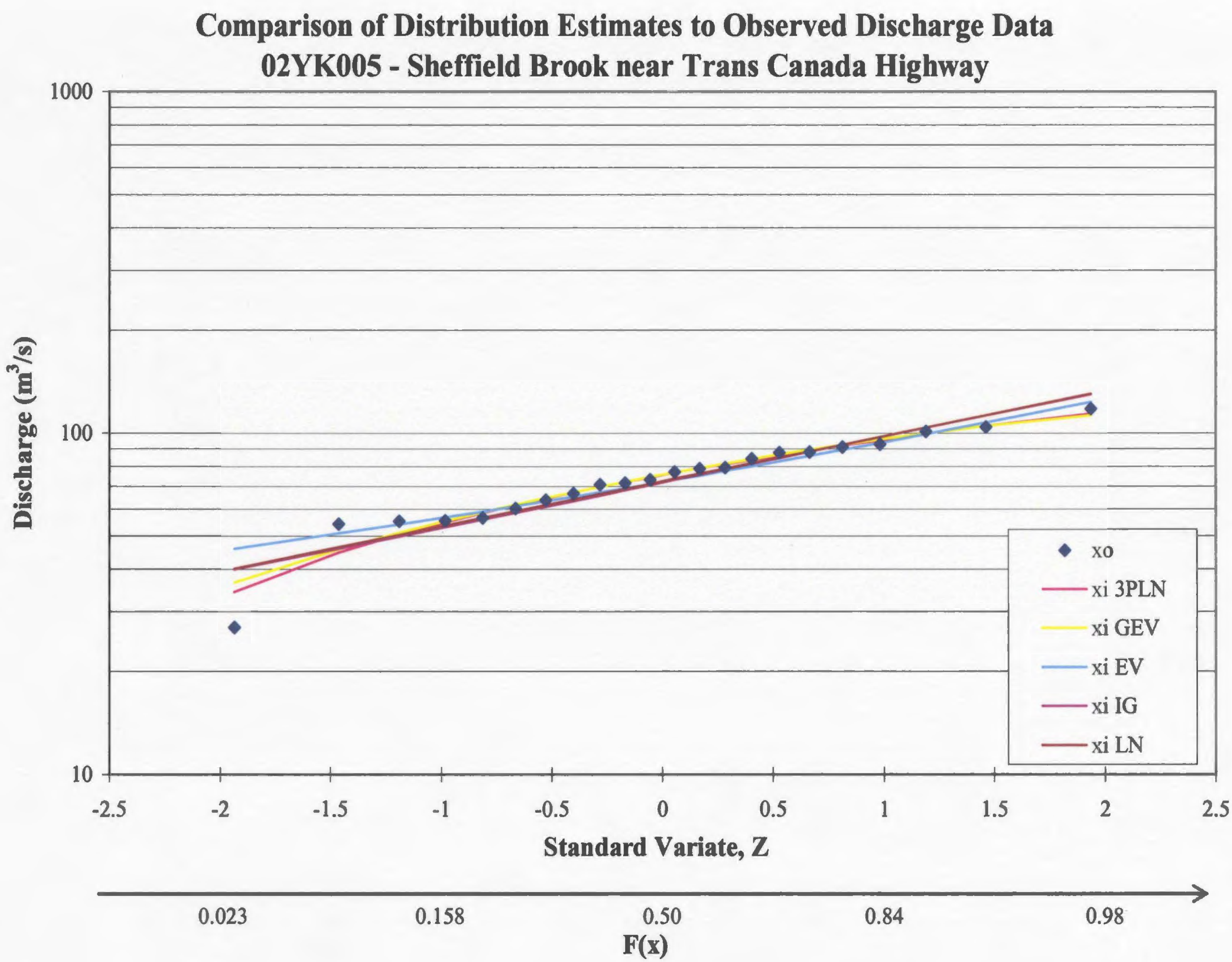
Comparison of Distribution Estimates to Observed Discharge Data
02YJ001 - Harrys River below Highway Bridge
Discharge (m3/s)
1000
100
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN


Comparison of Distribution Estimates to Observed Discharge Data
02YK002 - Lewaseechjeech Broot at Little Grand Lake
1000
100
10
Discharge (m3/s)
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN

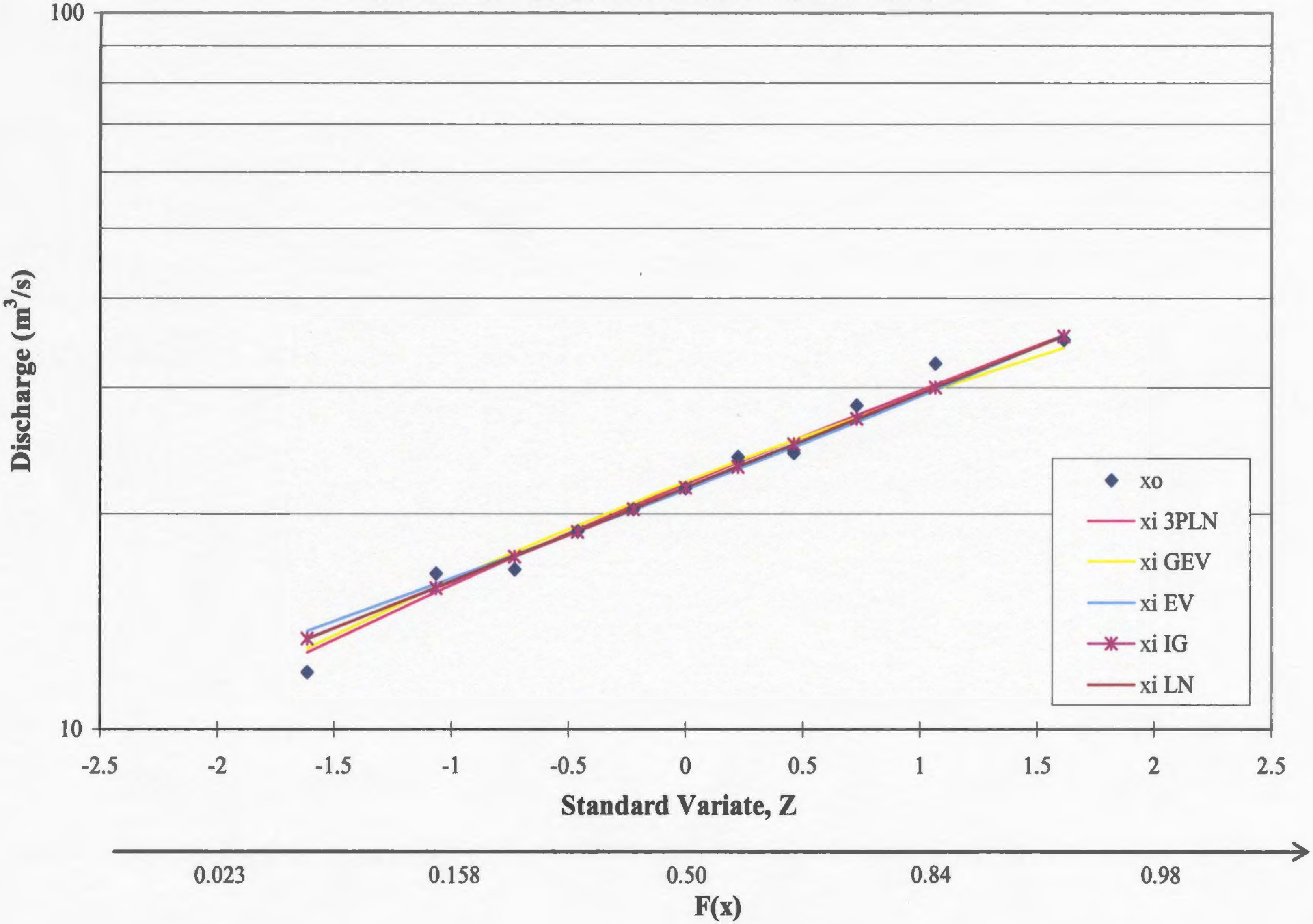
Comparison of Distribution Estimates to Observed Discharge Data
02YK003 - Sheffield River at Sheffield Lake



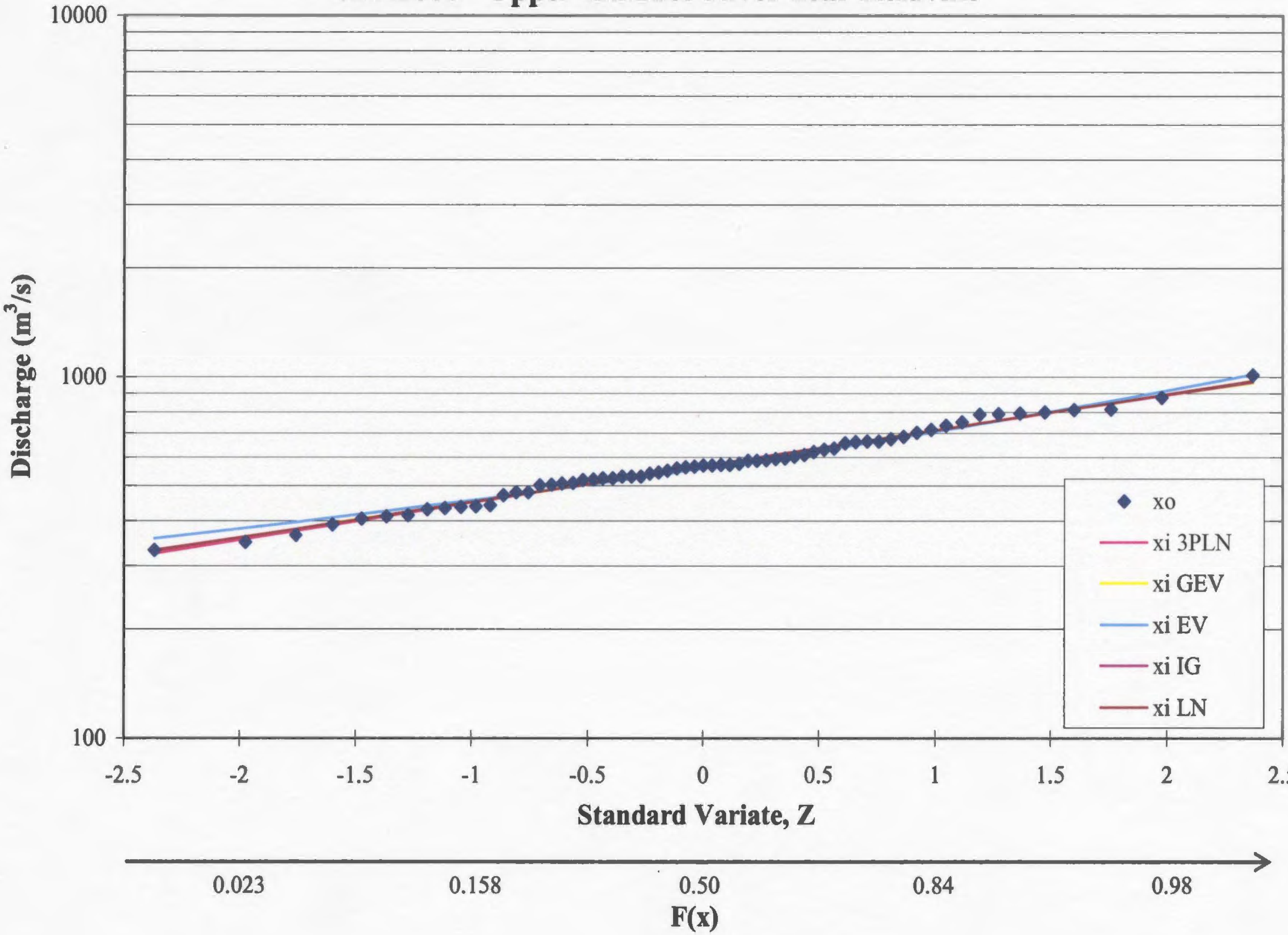

Comparison of Distribution Estimates to Observed Discharge Data
02YK004 - Hinds Brook near Grand Lake
Discharge (m^3/s)
1000
100
10
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN


Comparison of Distribution Estimates to Observed Discharge Data
02YK005 - Sheffield Brook near Trans Canada Highway
Discharge (m³/s)
1000
100
10
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN

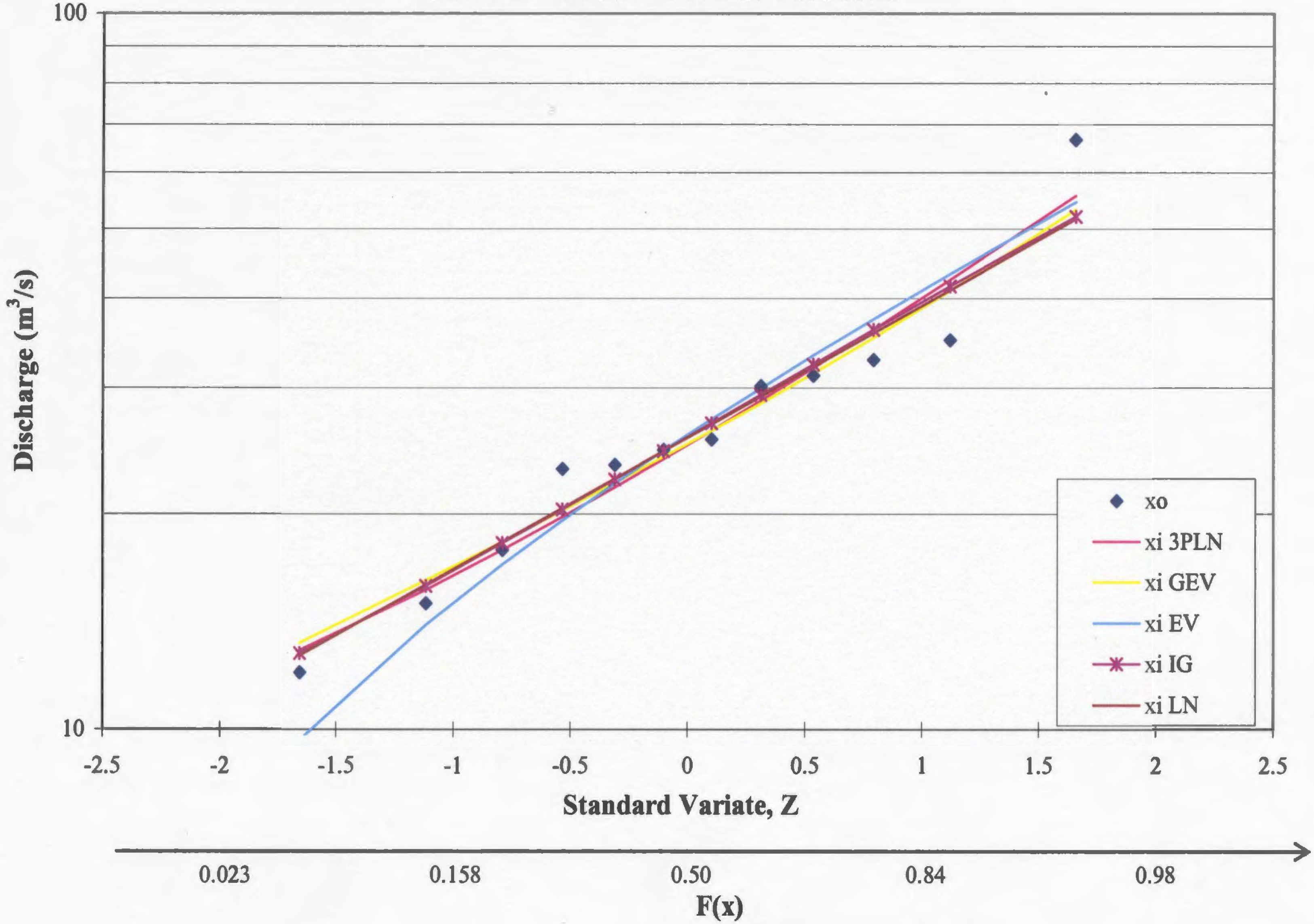
Comparison of Distribution Estimates to Observed Discharge Data
02YK007 - Glide Brook below Glide Lake



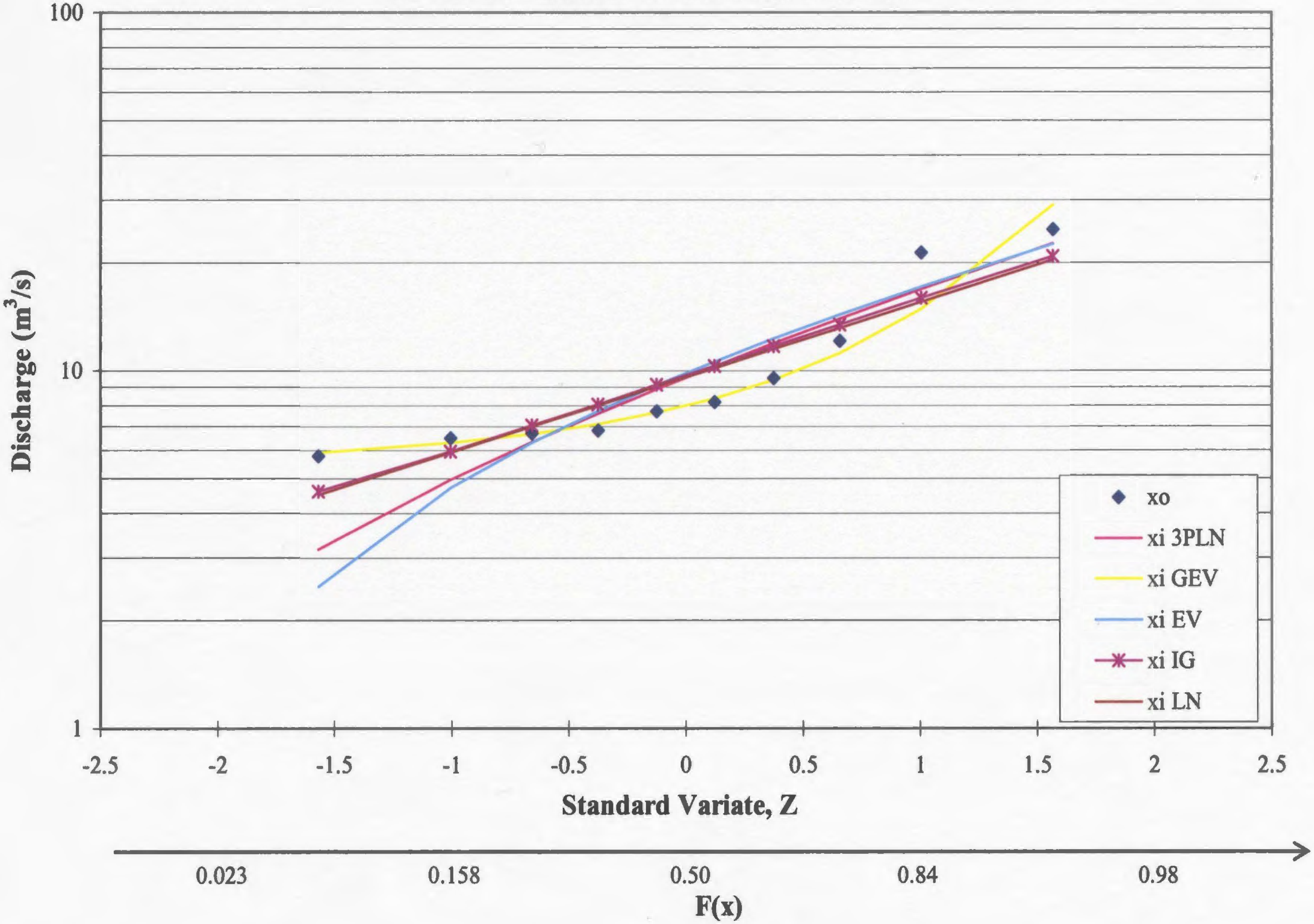
Comparison of Distribution Estimates to Observed Discharge Data
02YL001 - Upper Humber River near Reidville



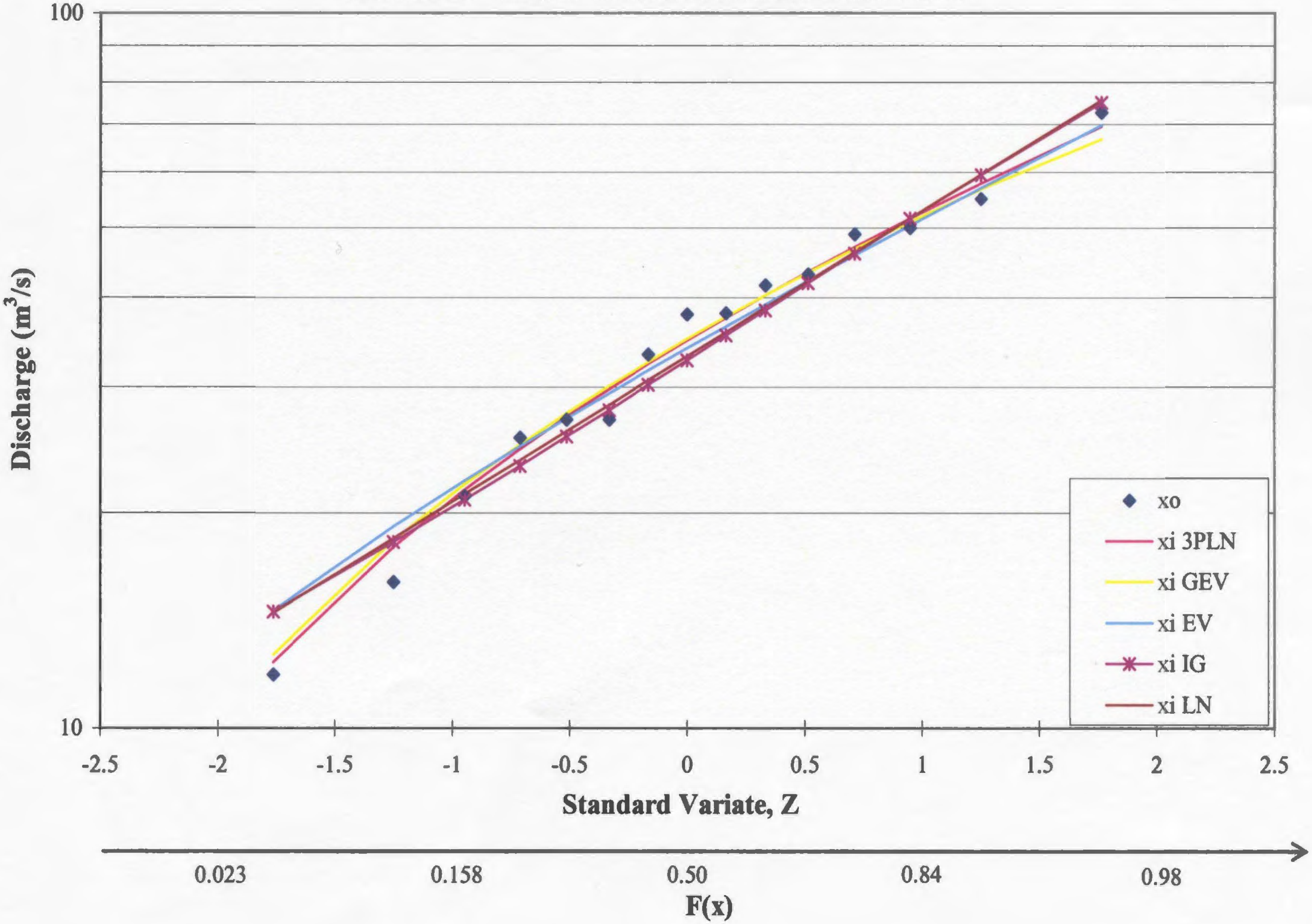
Comparison of Distribution Estimates to Observed Discharge Data
02YL004 - South Brook at Pasadena

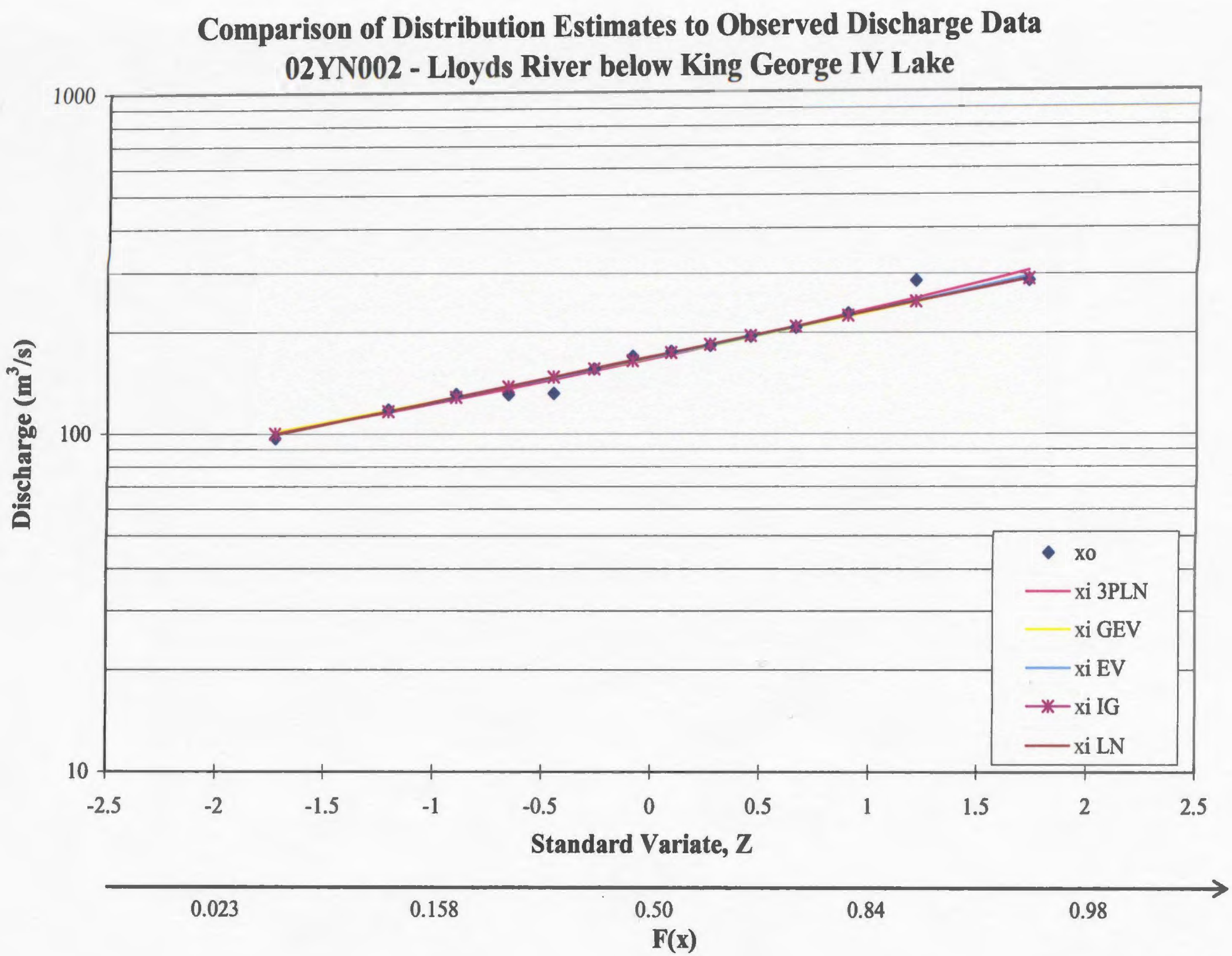


Comparison of Distribution Estimates to Observed Discharge Data
02YL005 - Rattler Brook near McIvers

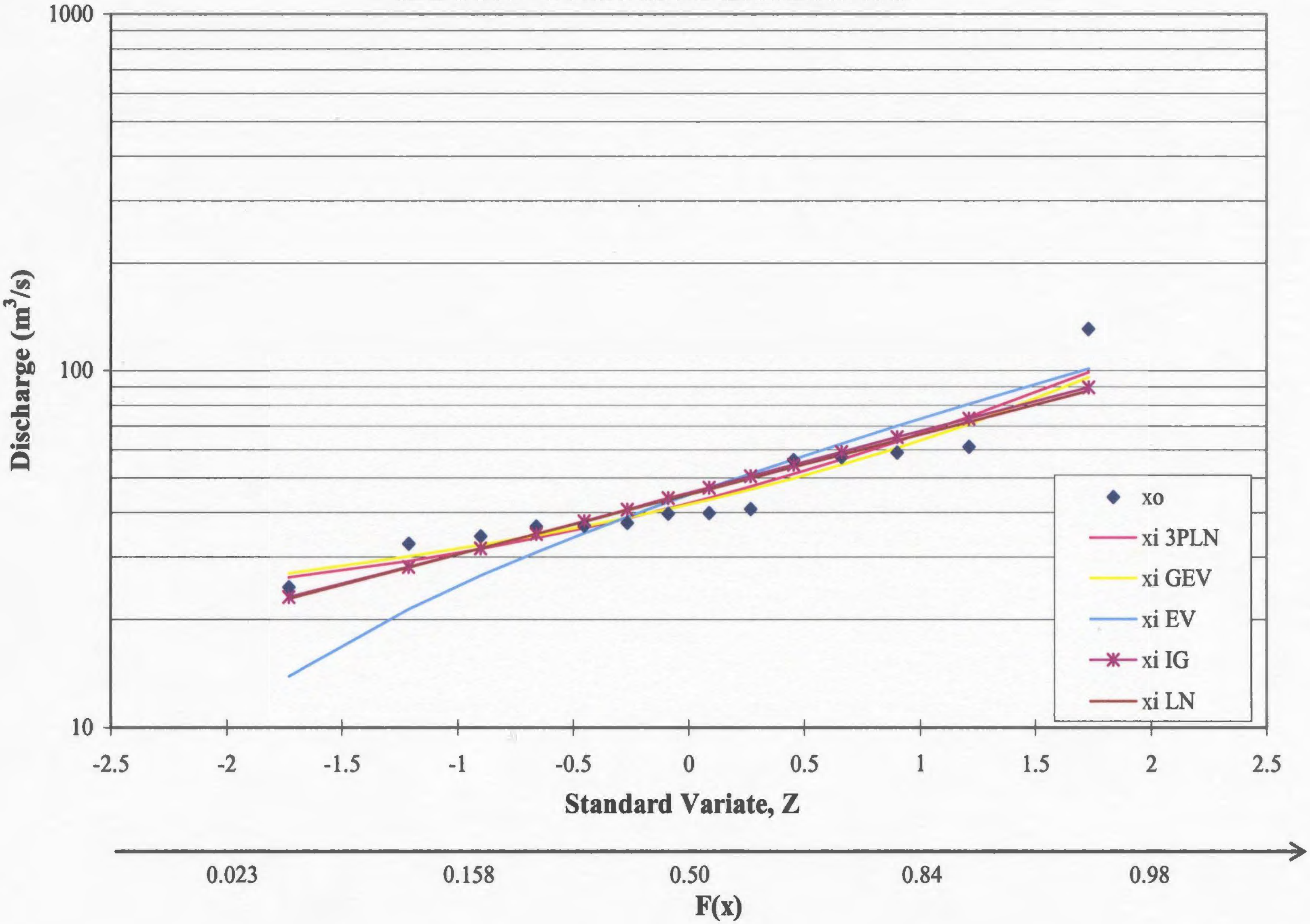


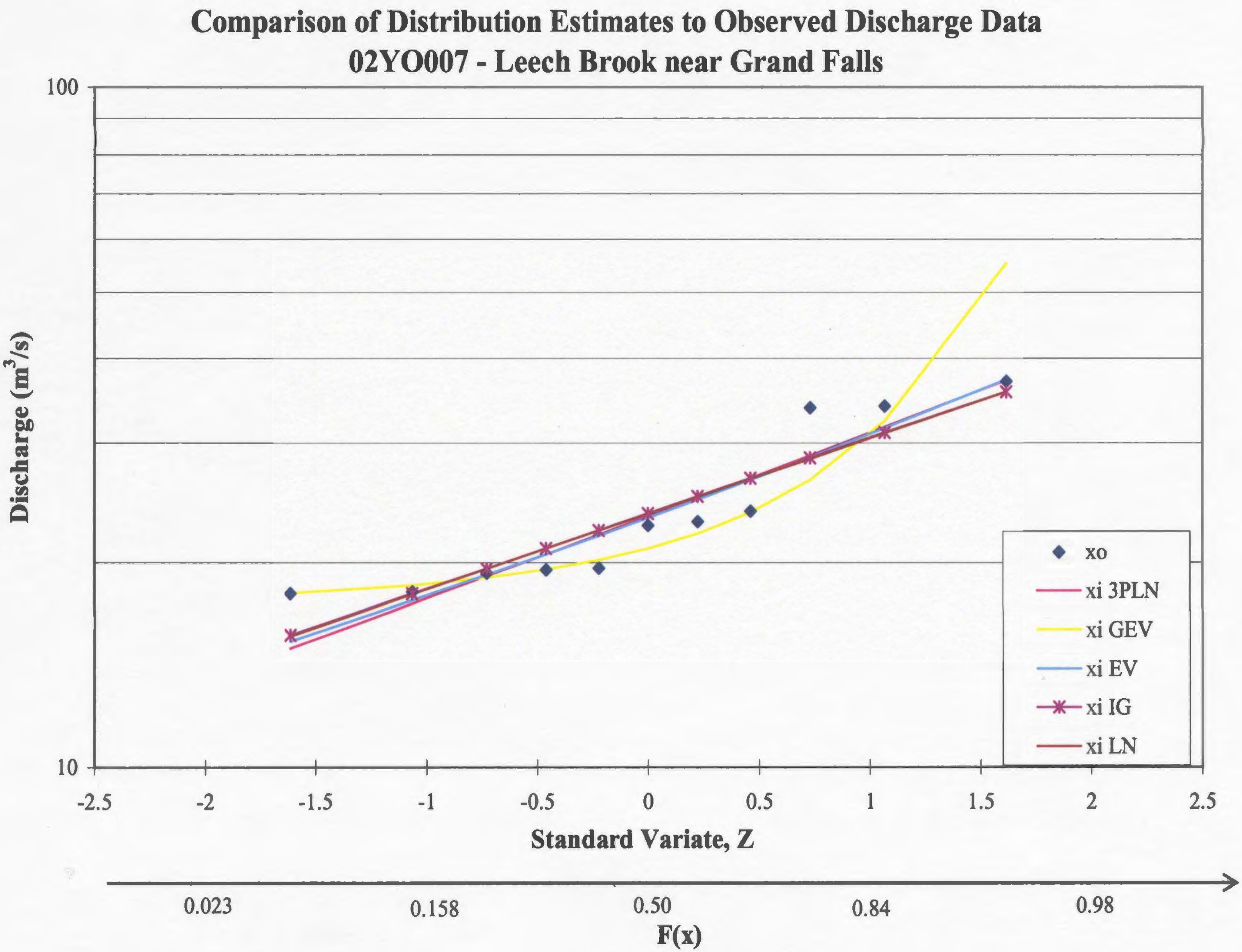
Comparison of Distribution Estimates to Observed Discharge Data
02YM003 - South West Brook near Baie Verte

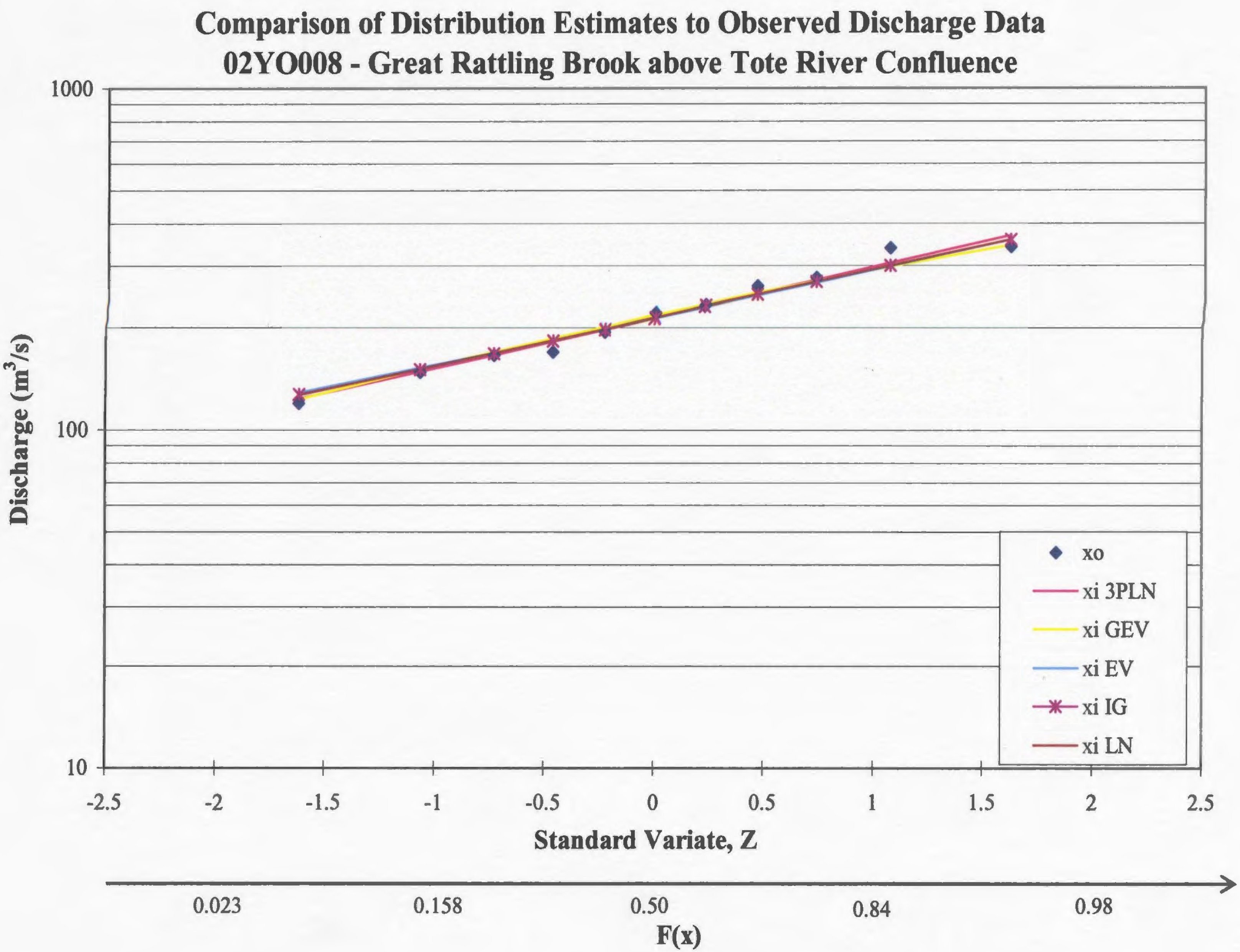


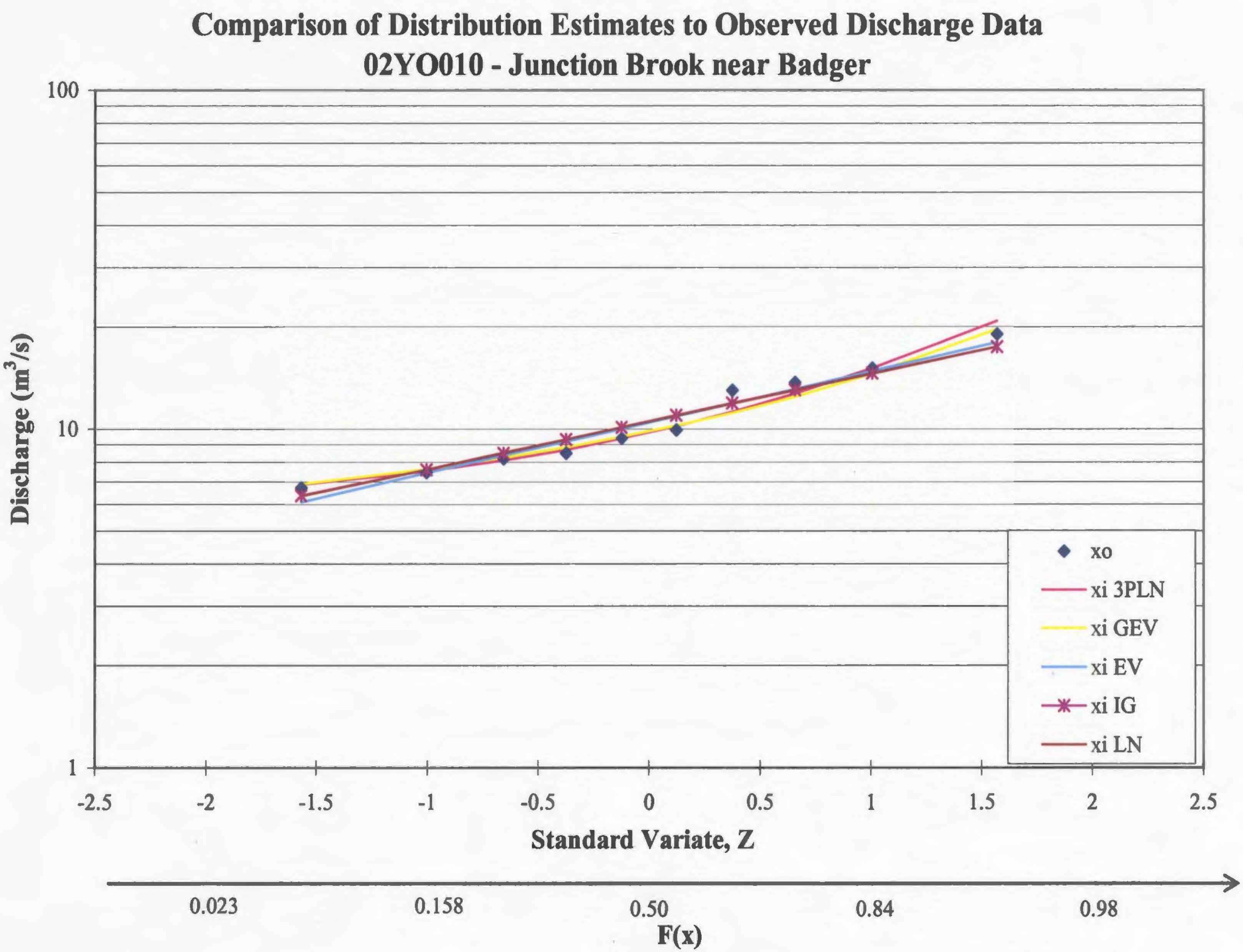

Comparison of Distribution Estimates to Observed Discharge Data
02YN002 - Lloyds River below King George IV Lake
1000
100
10
Discharge (m3/s)
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN

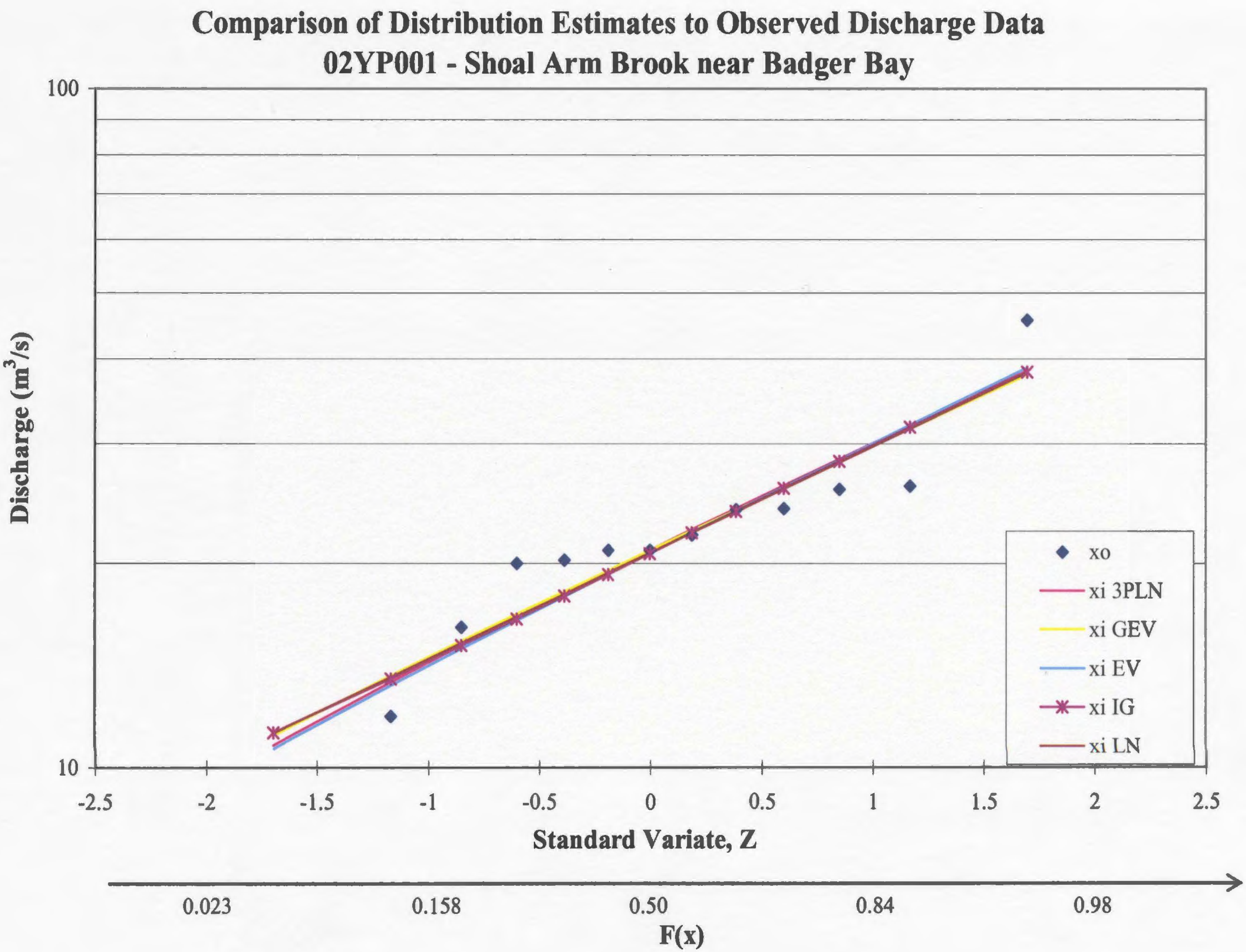
Comparison of Distribution Estimates to Observed Discharge Data
02YO006 - Peters River neat Botwood



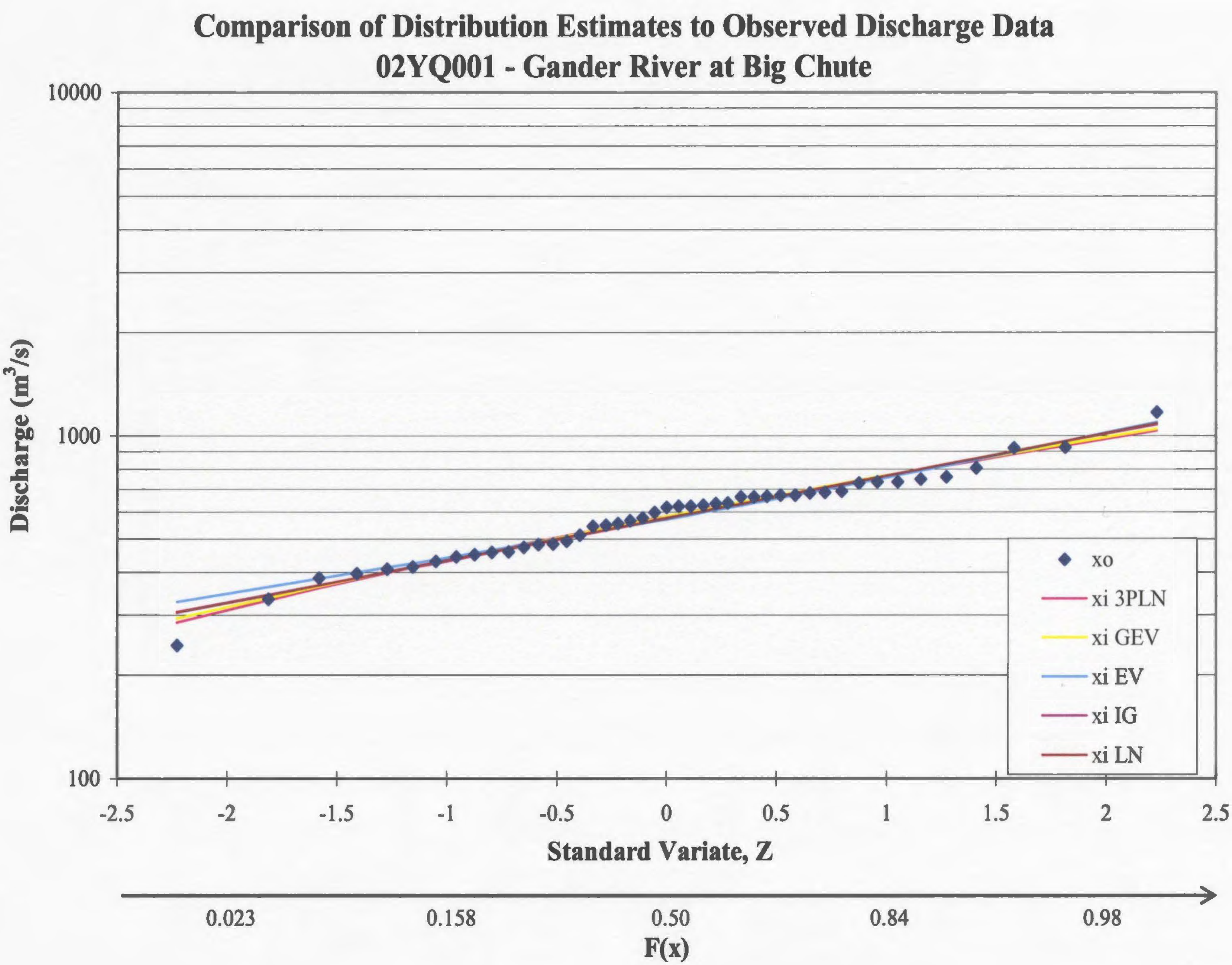

Comparison of Distribution Estimates to Observed Discharge Data
02YO007 - Leech Brook near Grand Falls
100
10
Discharge (m3/s)
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN


Comparison of Distribution Estimates to Observed Discharge Data
02YO008 - Great Rattling Brook above Tote River Confluence
Discharge (m3/s)
1000
100
10
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN

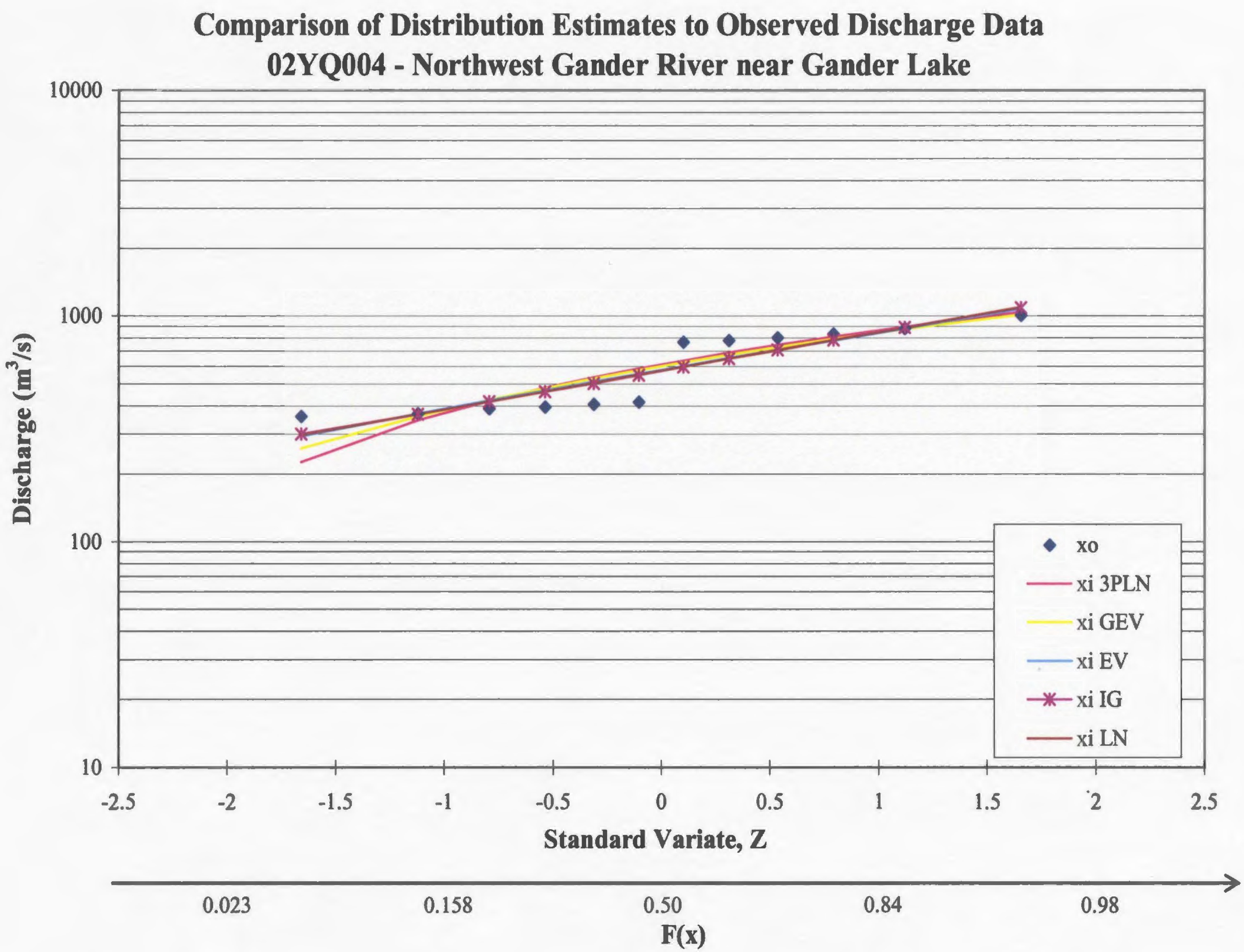

Comparison of Distribution Estimates to Observed Discharge Data
02YO010 - Junction Brook near Badger
Discharge (m3/s)
100
10
1
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN

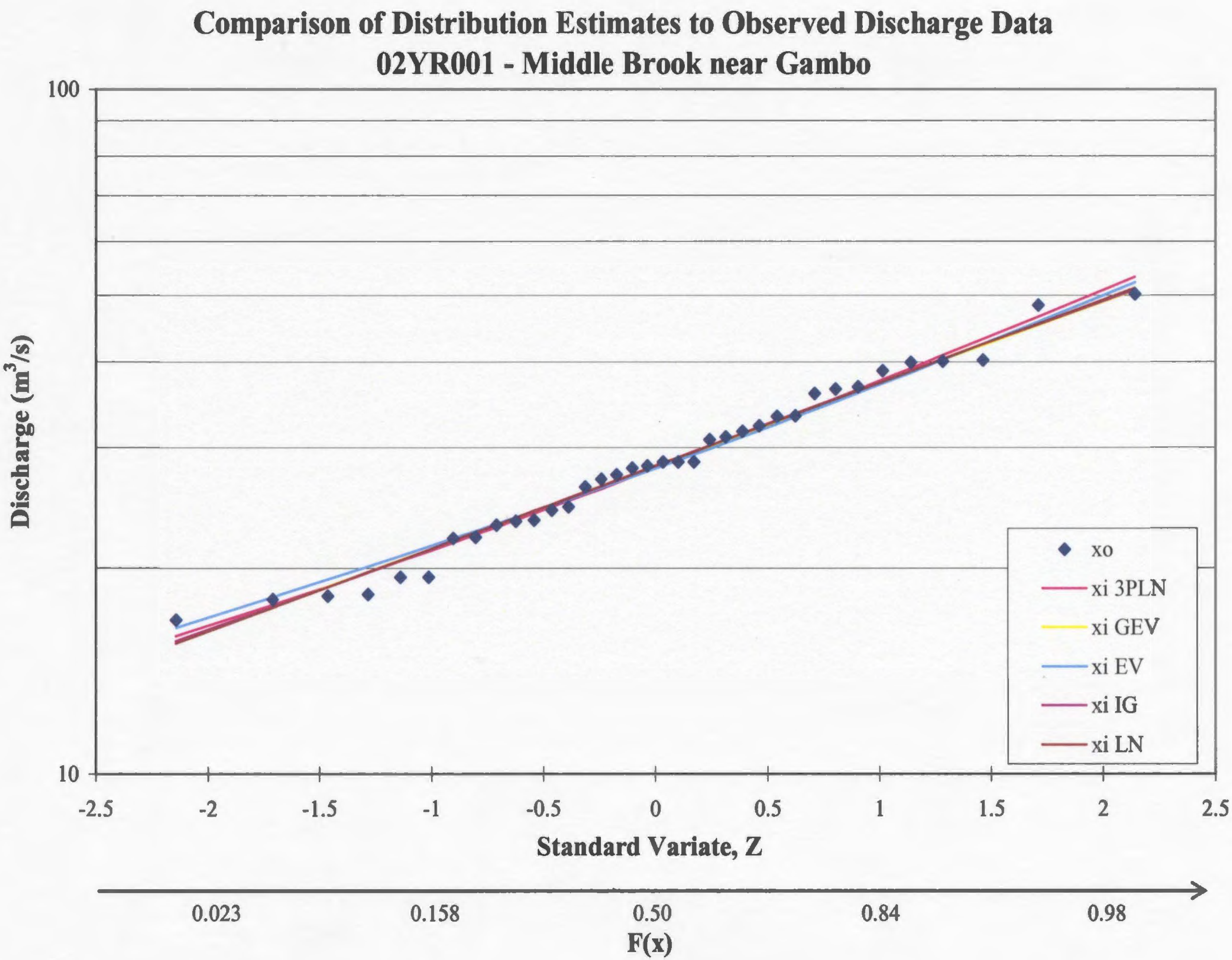


Comparison of Distribution Estimates to Observed Discharge Data
02YP001 - Shoal Arm Brook near Badger Bay
100
10
Discharge (m³/s)
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN

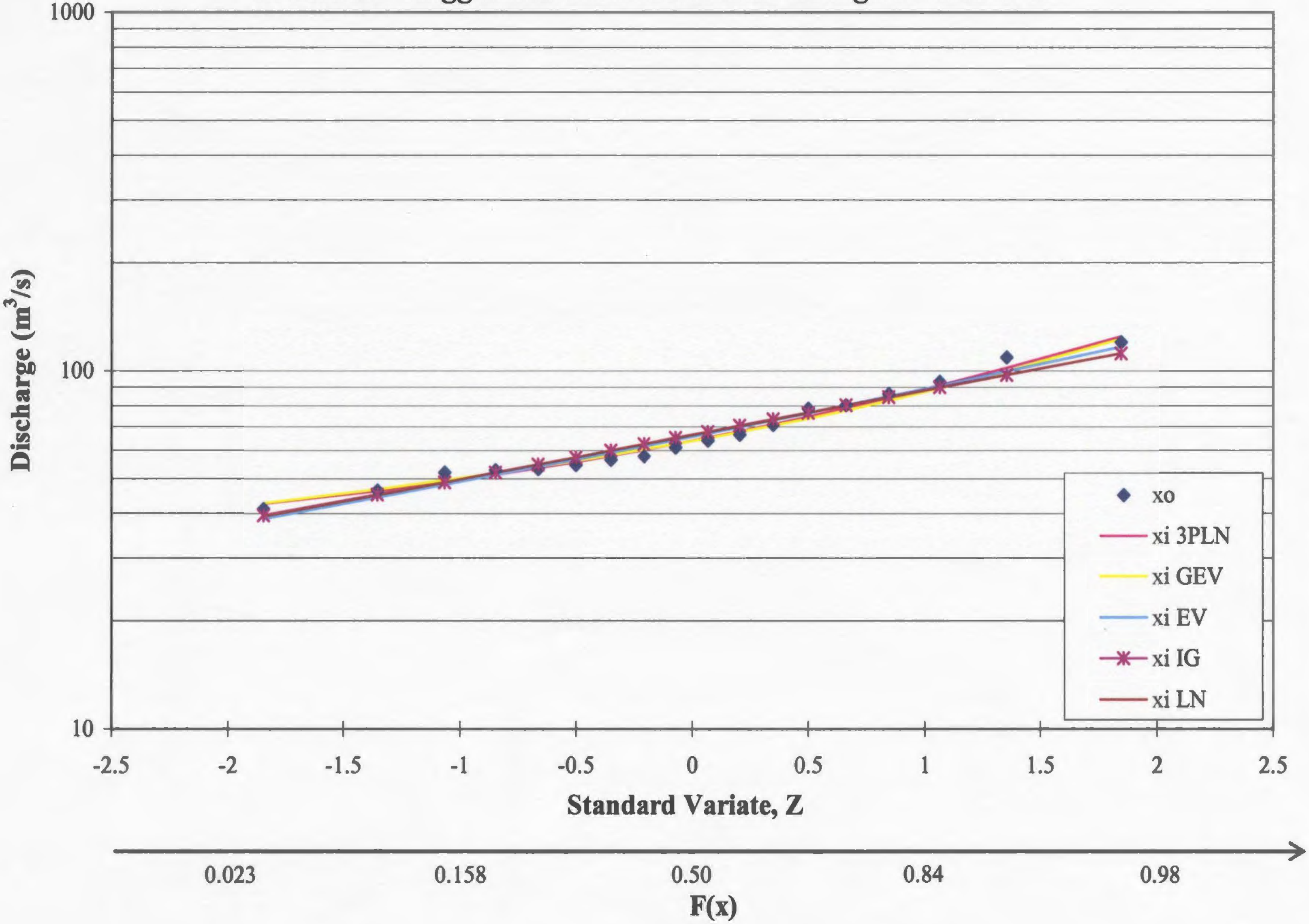


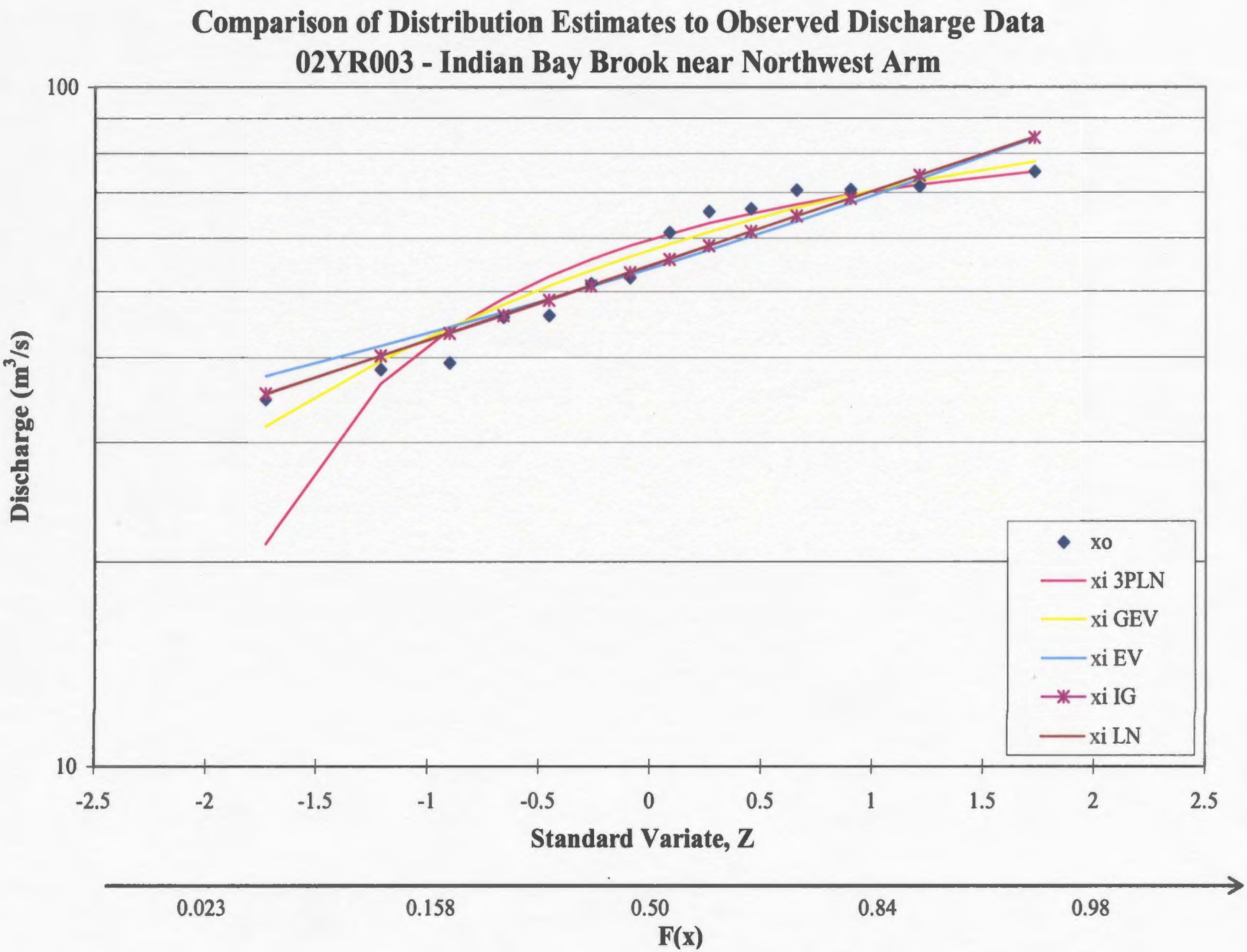
Comparison of Distribution Estimates to Observed Discharge Data
02YQ001 - Gander River at Big Chute
10000
1000
100
Discharge (m3/s)
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN

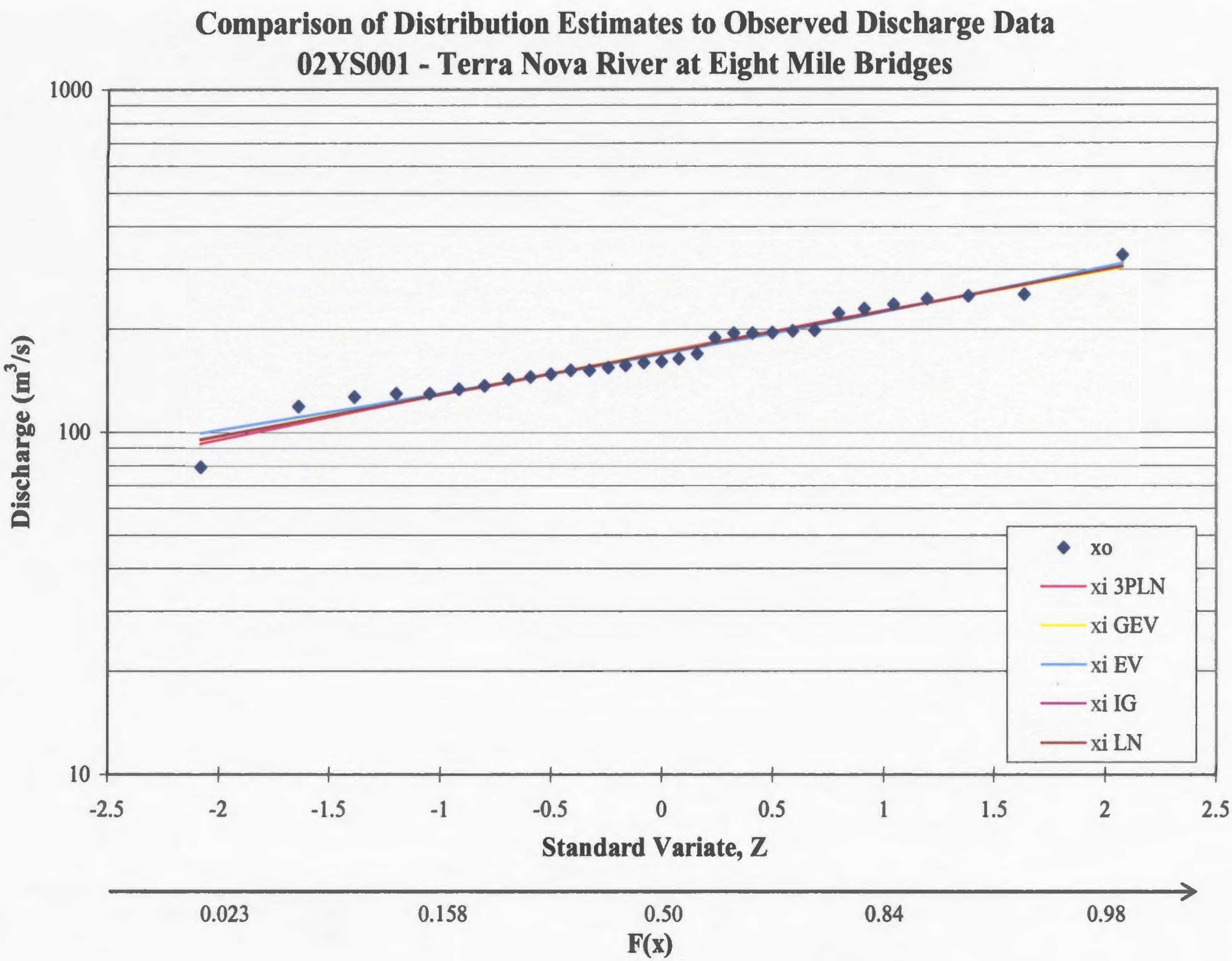

Comparison of Distribution Estimates to Observed Discharge Data
02YQ004 - Northwest Gander River near Gander Lake
Discharge (m3/s)
10000
1000
100
10
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN

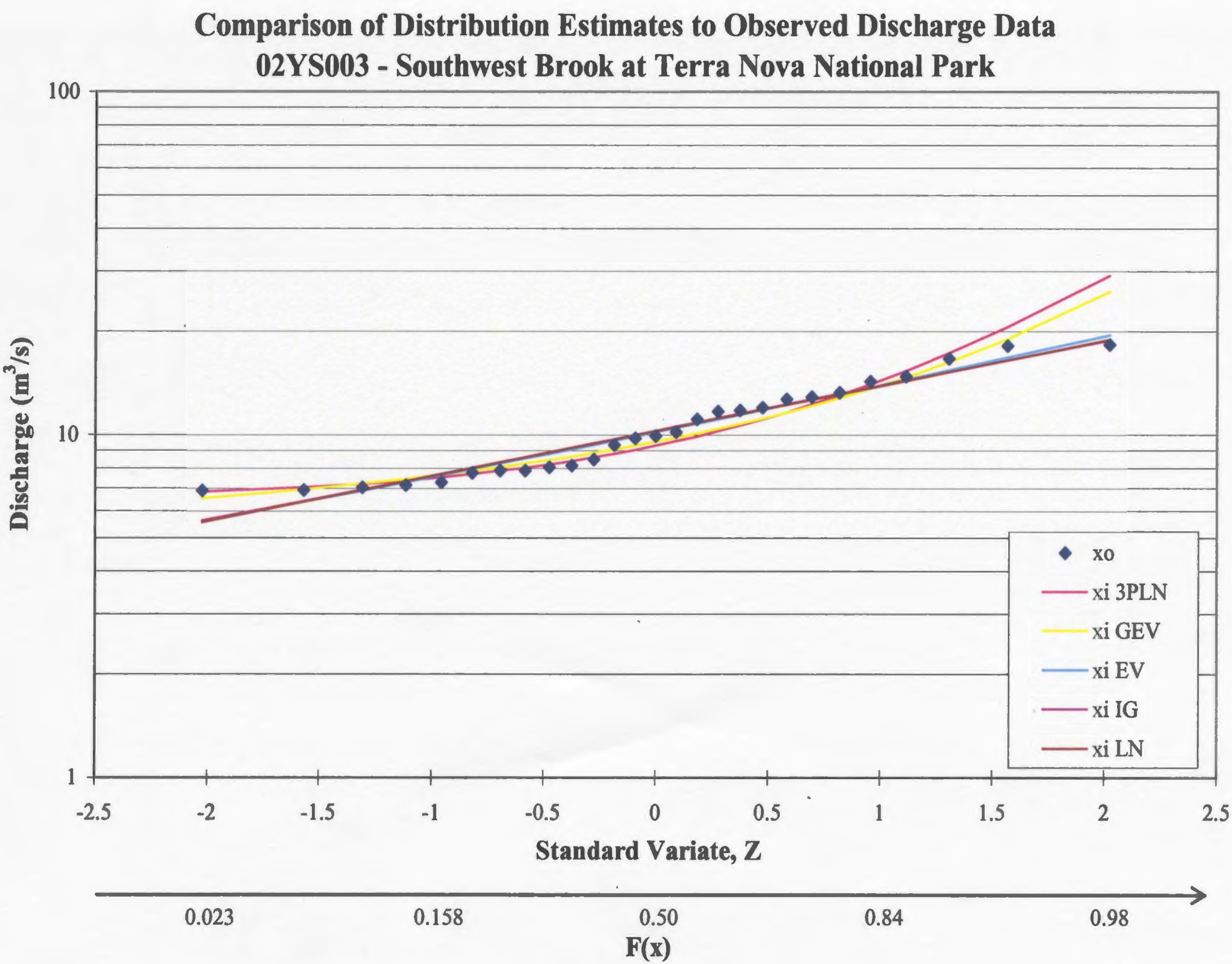

Comparison of Distribution Estimates to Observed Discharge Data
02YR001 - Middle Brook near Gambo
Discharge (m³/s)
100
10
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN
0.023
0.158
0.50
0.84
0.98
F(x)

Comparison of Distribution Estimates to Observed Discharge Data
02YR002 - Ragged Harbour River near Musgrave Harbour

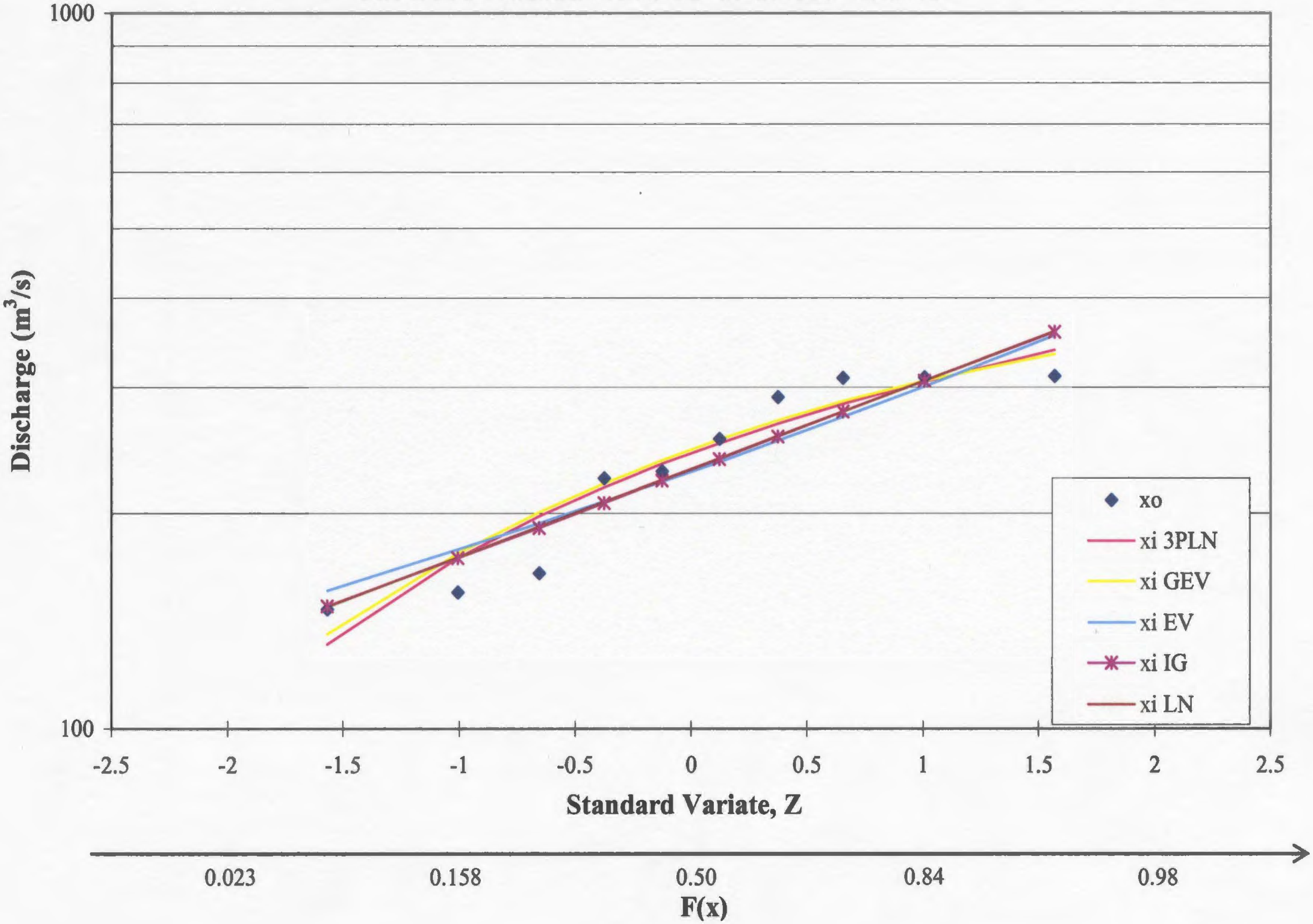



Comparison of Distribution Estimates to Observed Discharge Data
02YR003 - Indian Bay Brook near Northwest Arm
100
10
Discharge (m³/s)
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN

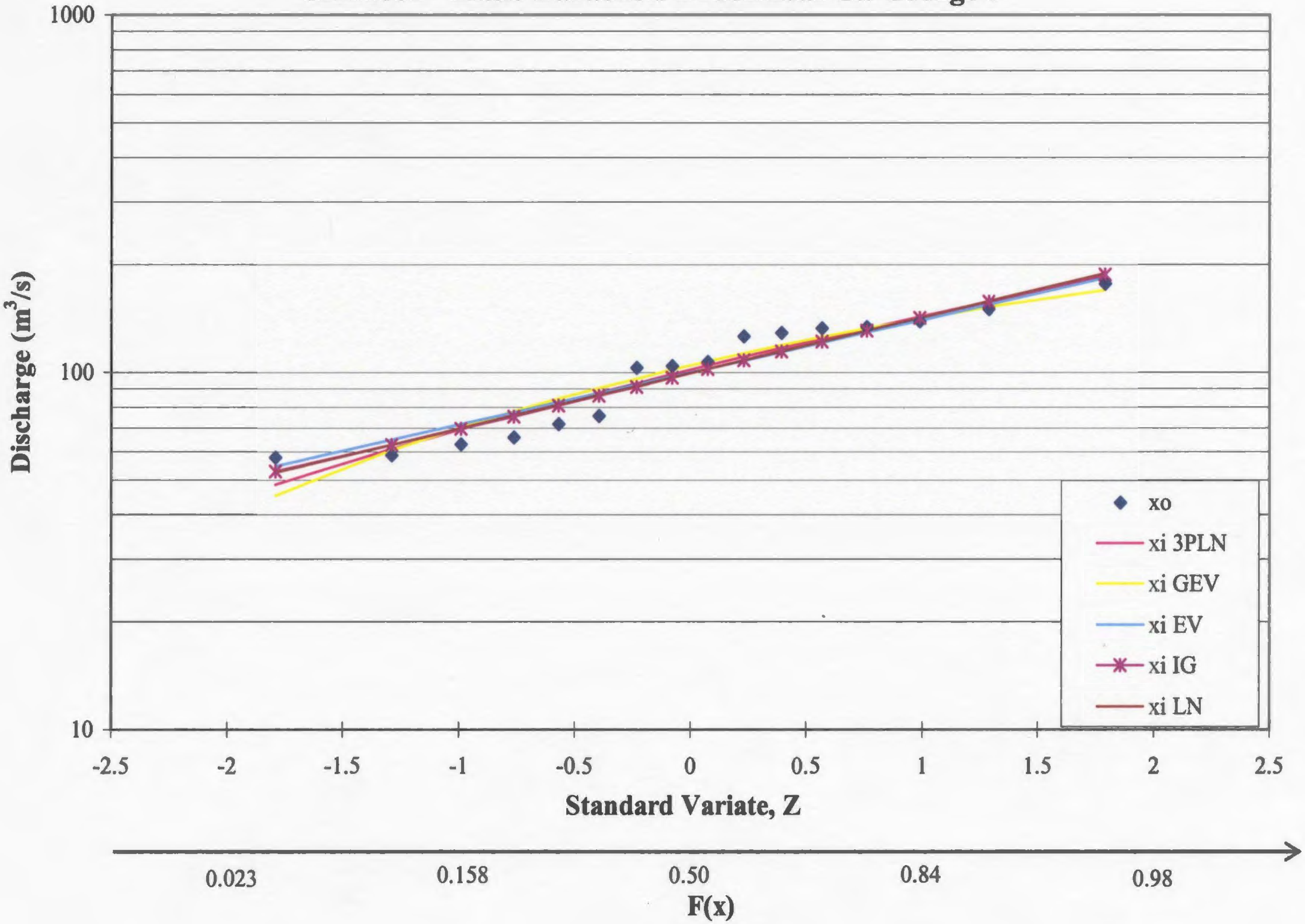

Comparison of Distribution Estimates to Observed Discharge Data
02YS001 - Terra Nova River at Eight Mile Bridges
Discharge (m3/s)
1000
100
10
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN


Comparison of Distribution Estimates to Observed Discharge Data
02YS003 - Southwest Brook at Terra Nova National Park
Discharge (m^3/s)
100
10
1
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN

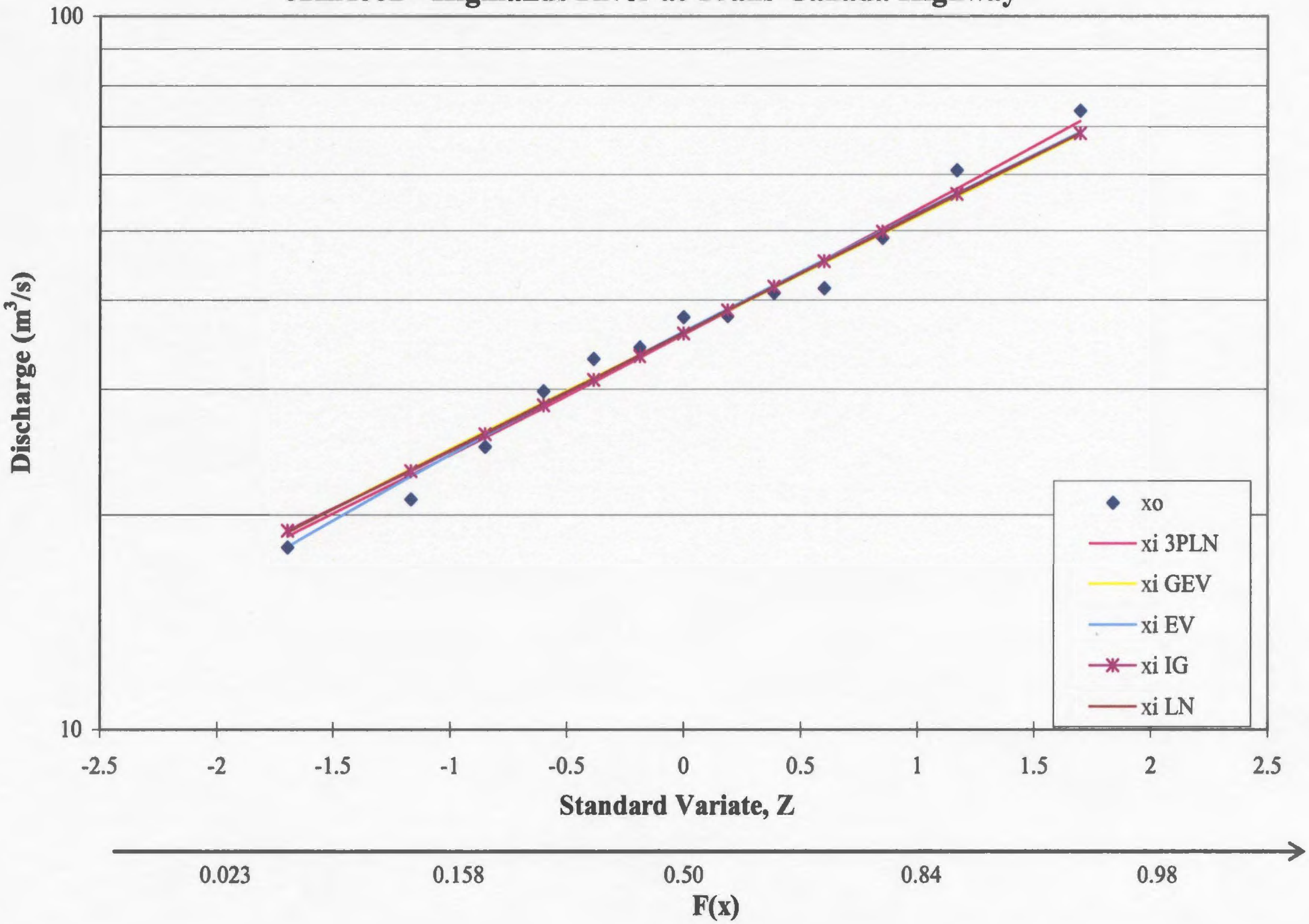
Comparison of Distribution Estimates to Observed Discharge Data
02YS005 - Terra Nova River at Glovertown



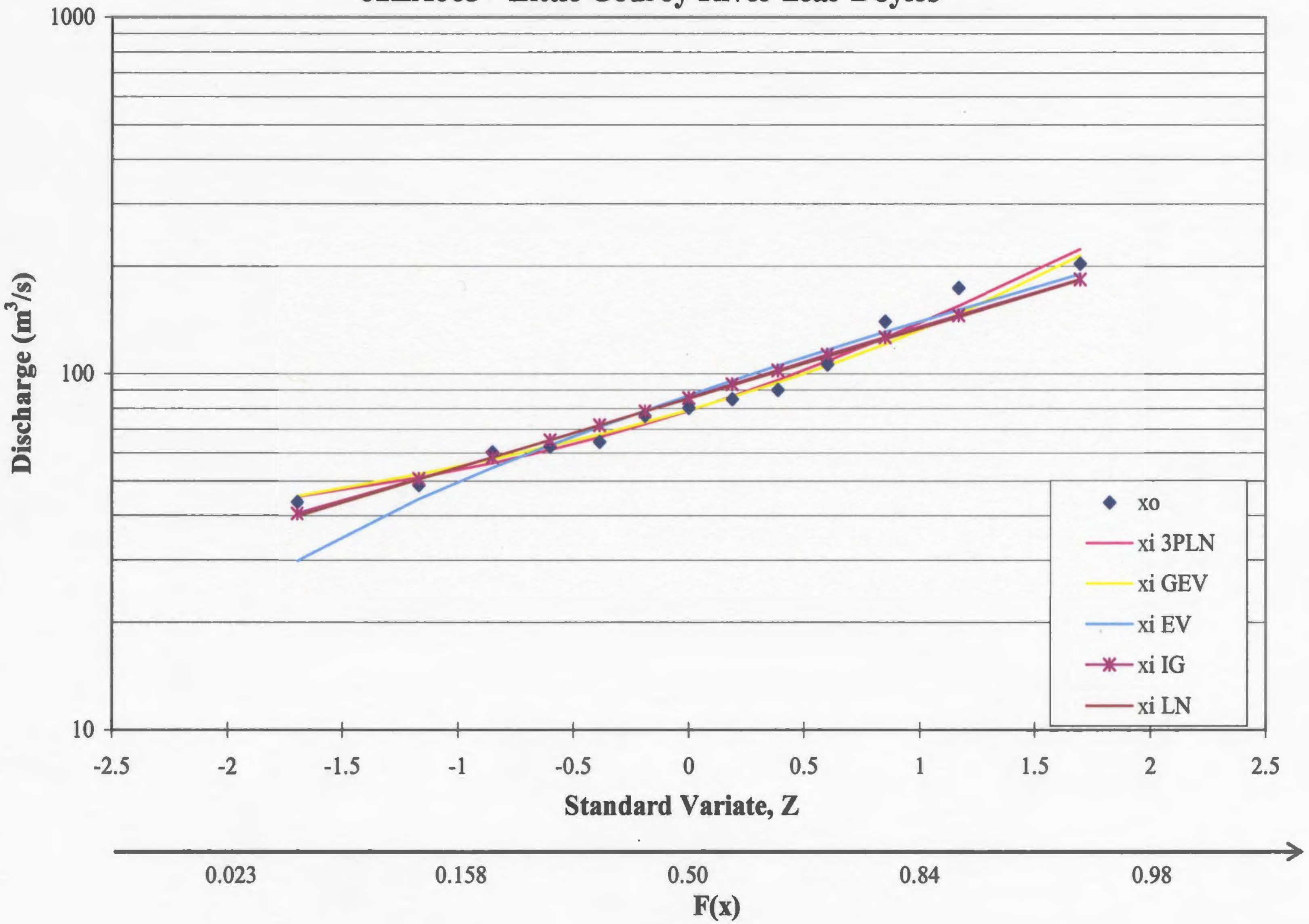
Comparison of Distribution Estimates to Observed Discharge Data
02ZA001 - Little Barachois Brook near St. George's



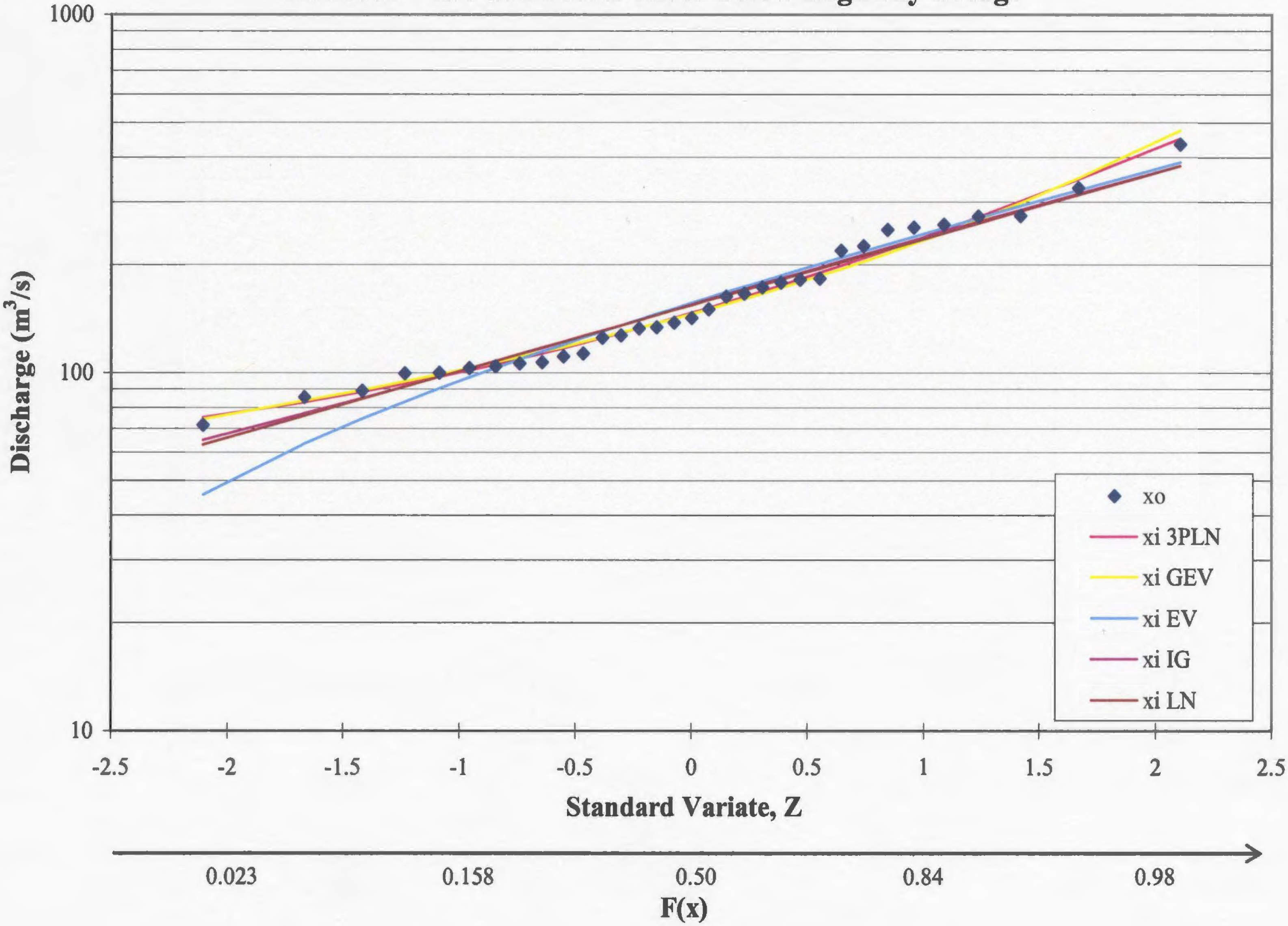
Comparison of Distribution Estimates to Observed Discharge Data
02ZA002 - Highlands River at Trans-Canada Highway



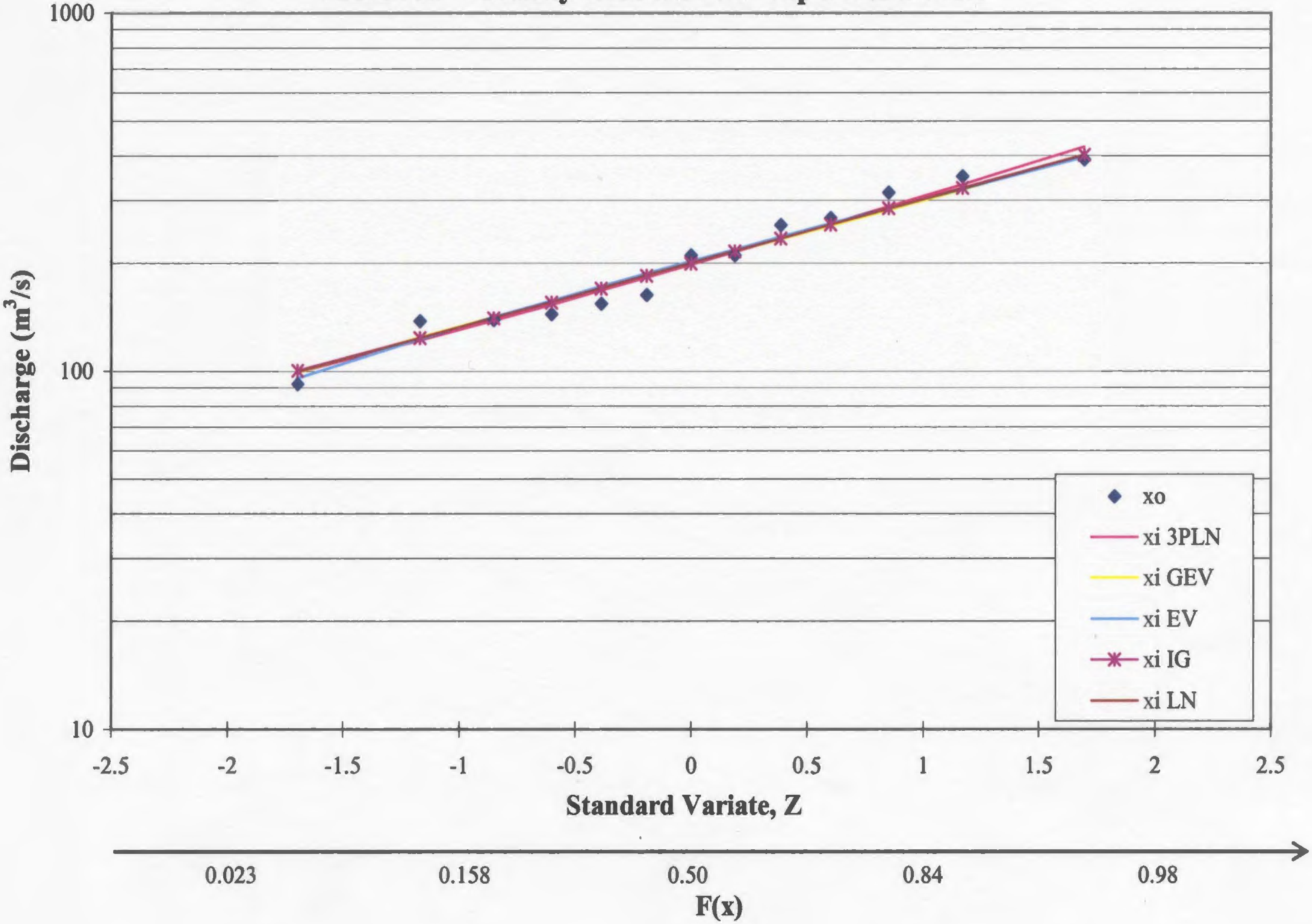
Comparison of Distribution Estimates to Observed Discharge Data
02ZA003 - Little Codroy River near Doyles

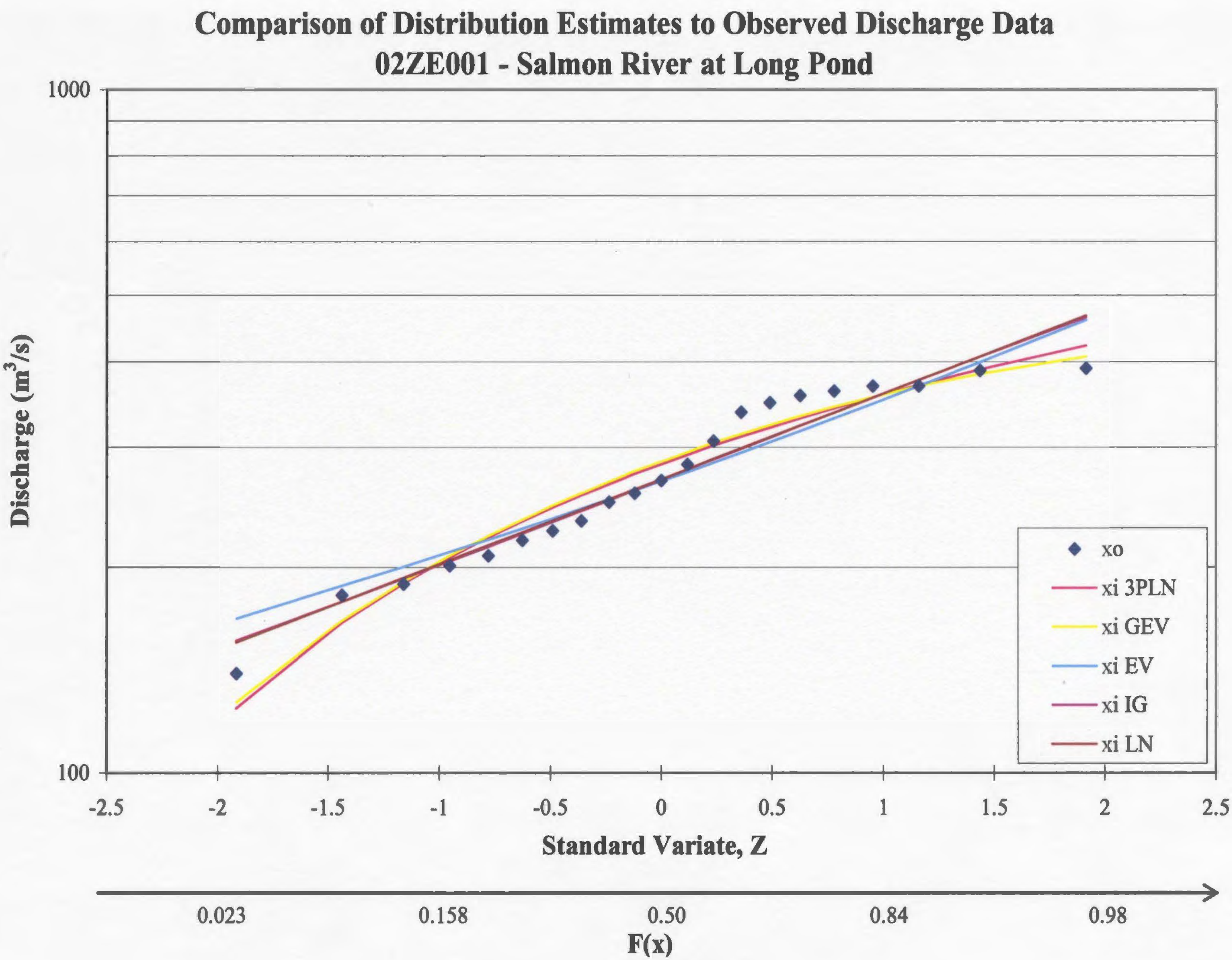


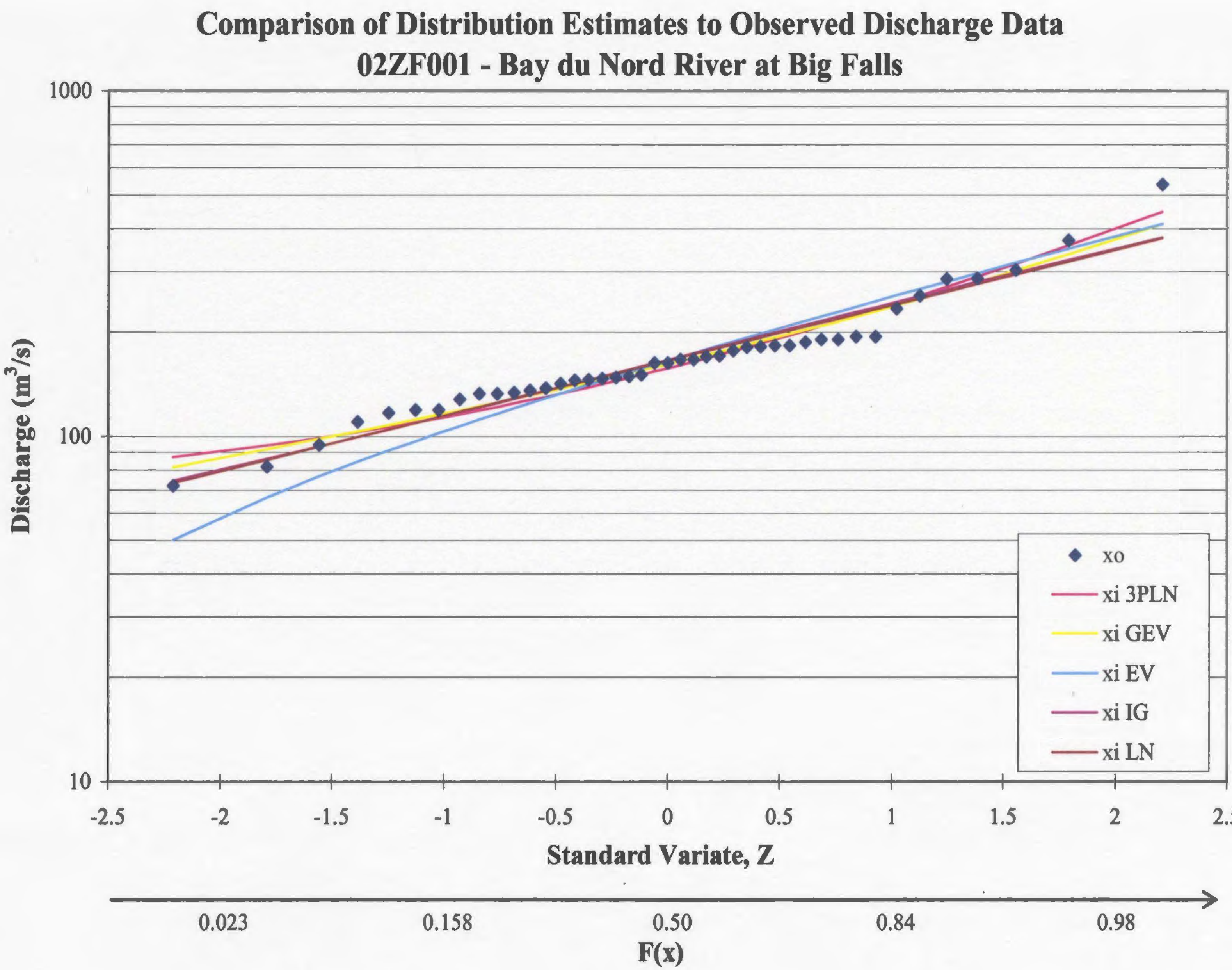
Comparison of Distribution Estimates to Observed Discharge Data
02ZB001 - Isle aux Morts River below Highway Bridge



Comparison of Distribution Estimates to Observed Discharge Data
02ZC002 - Grandy Brook below Top Pond Brook

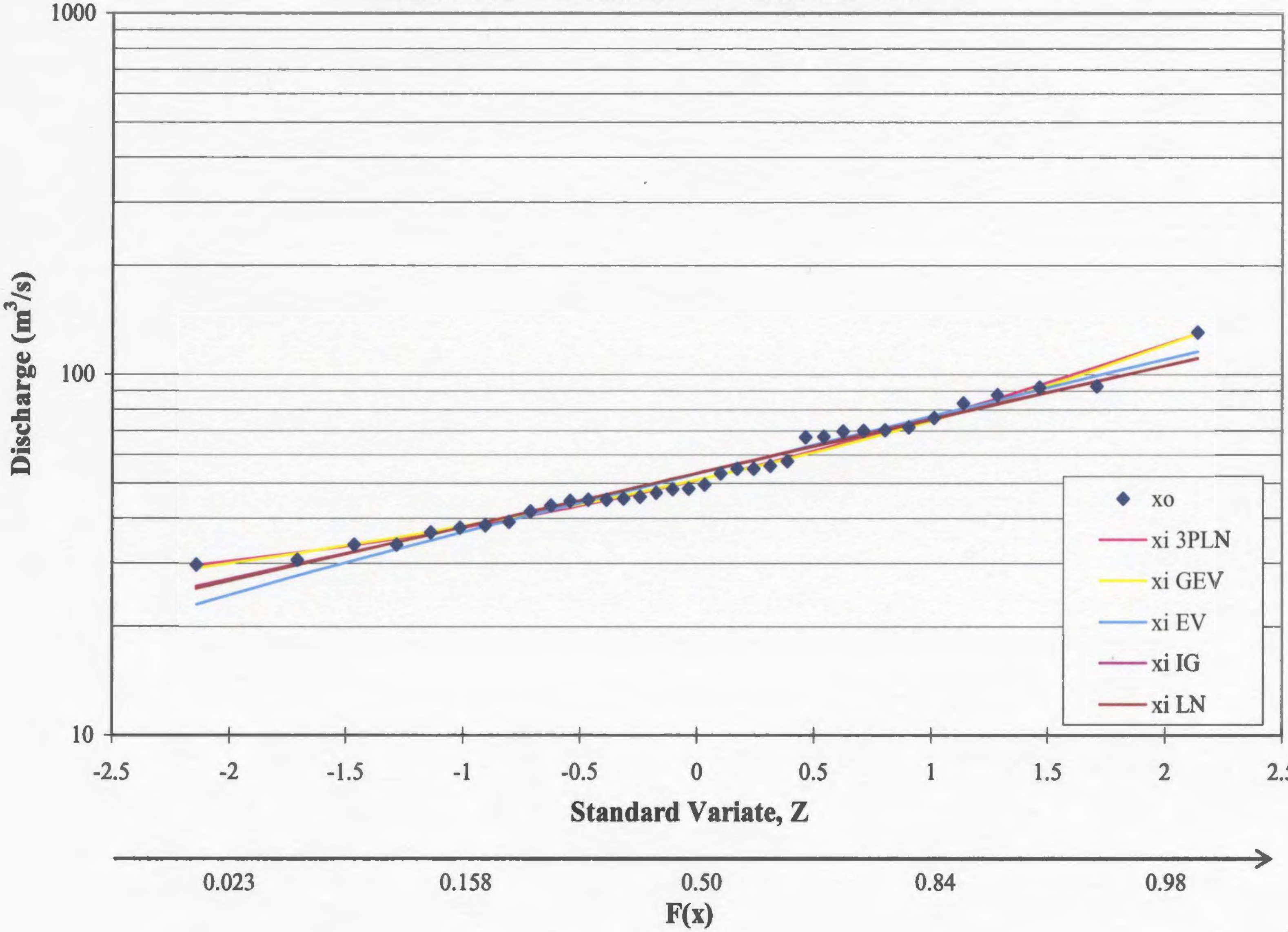


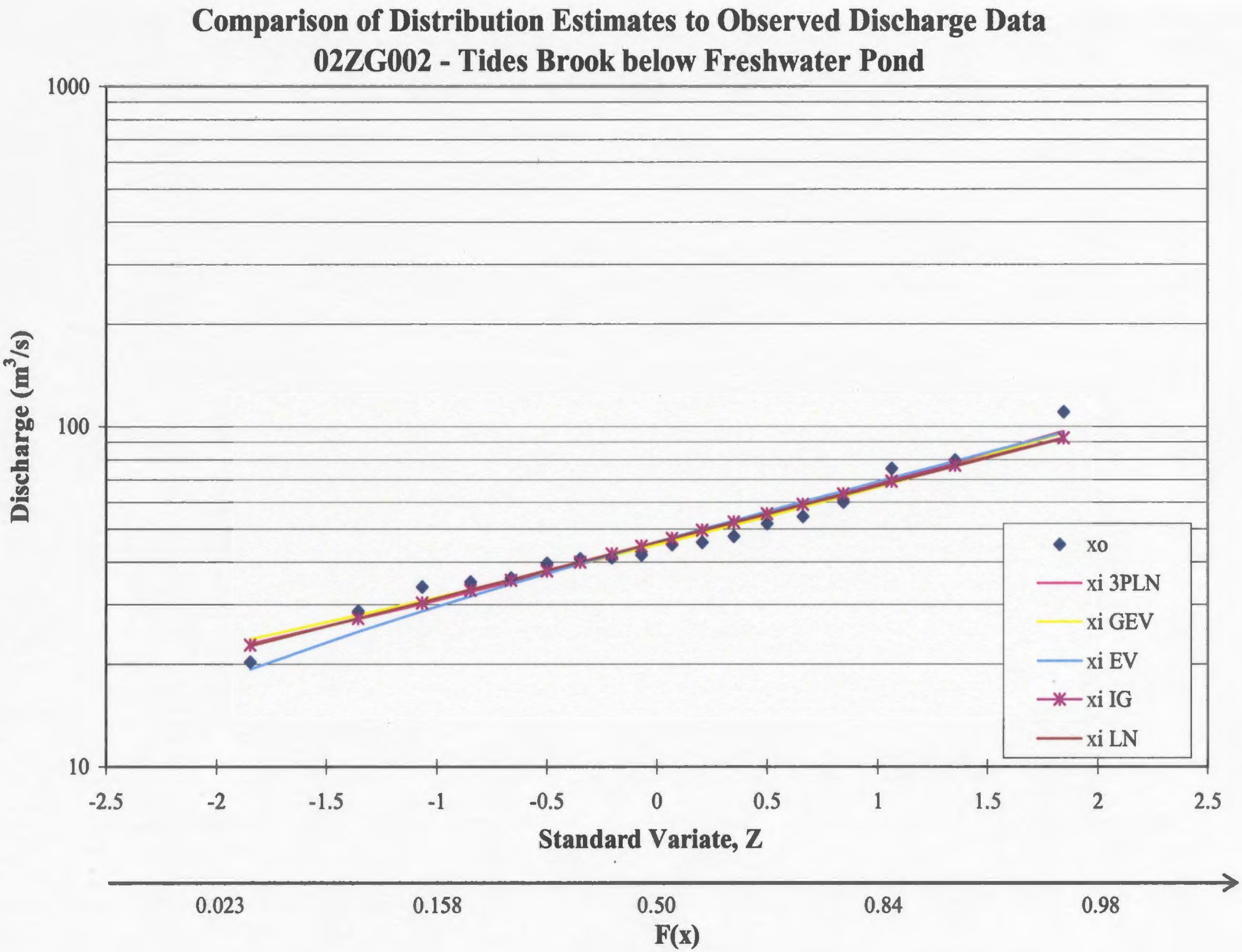

Comparison of Distribution Estimates to Observed Discharge Data
02ZE001 - Salmon River at Long Pond
1000
100
Discharge (m3/s)
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN



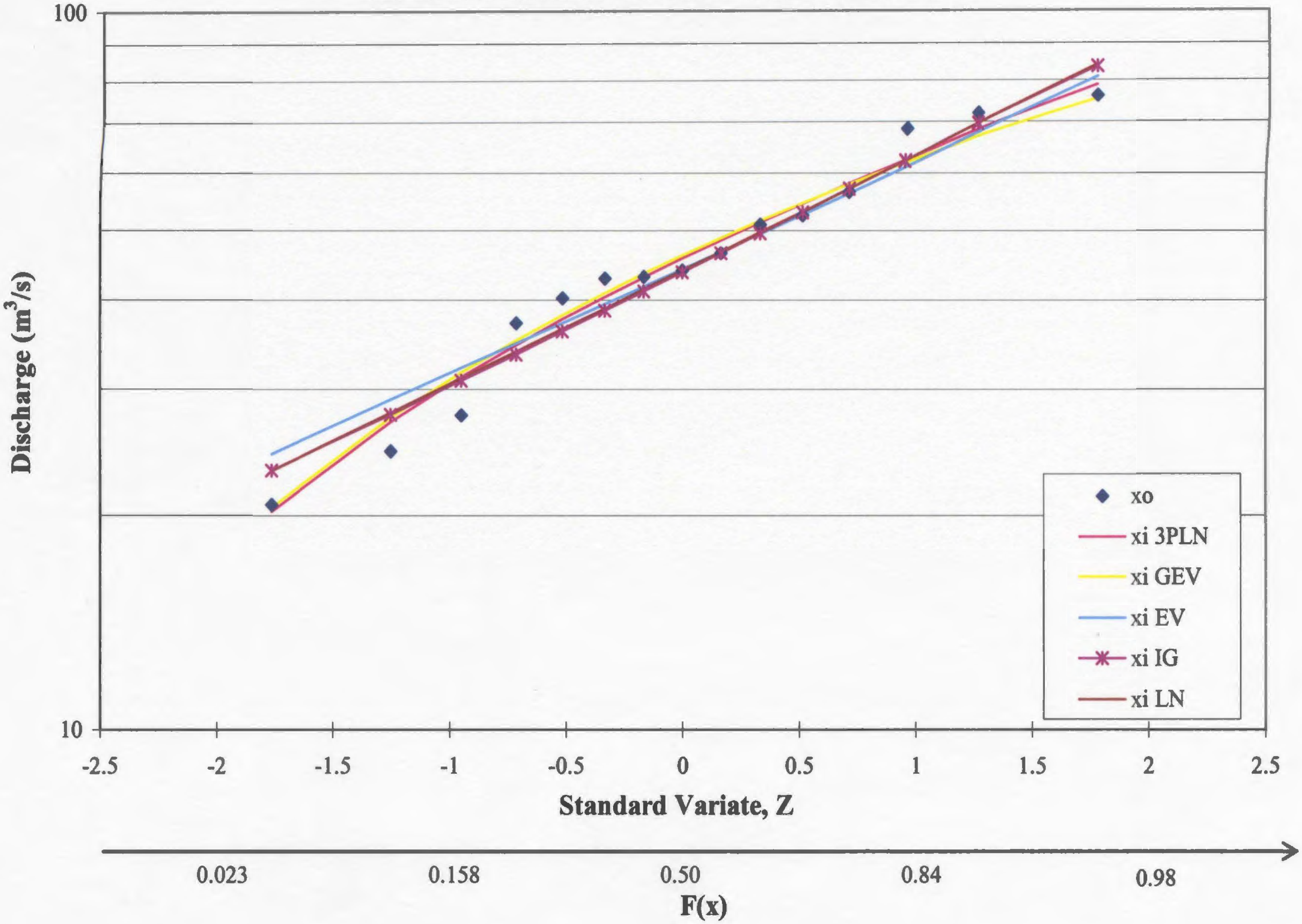
Comparison of Distribution Estimates to Observed Discharge Data
02ZF001 - Bay du Nord River at Big Falls
1000
100
10
Discharge (m³/s)
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN

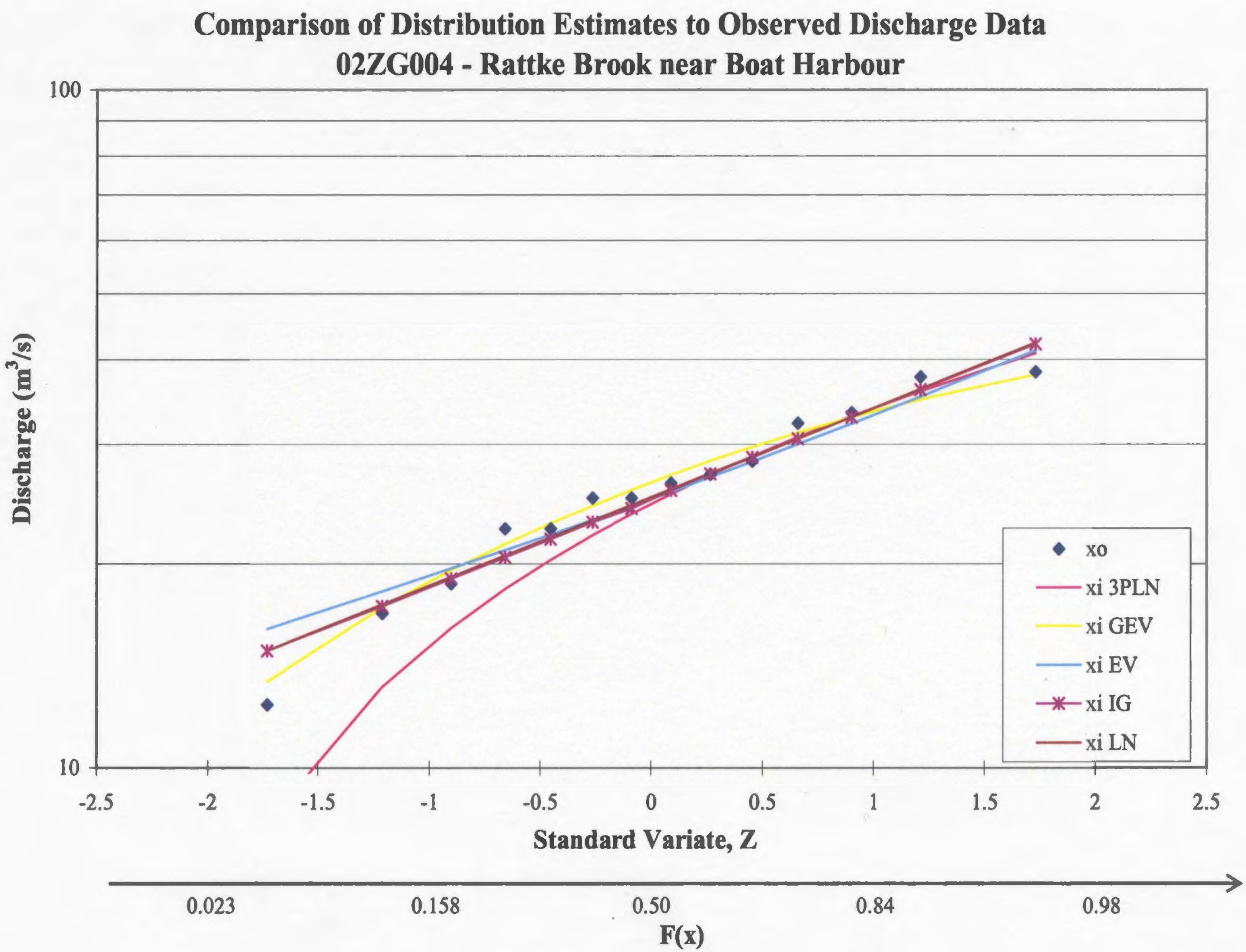
Comparison of Distribution Estimates to Observed Discharge Data
02ZG001 - Garnish River near Garnish



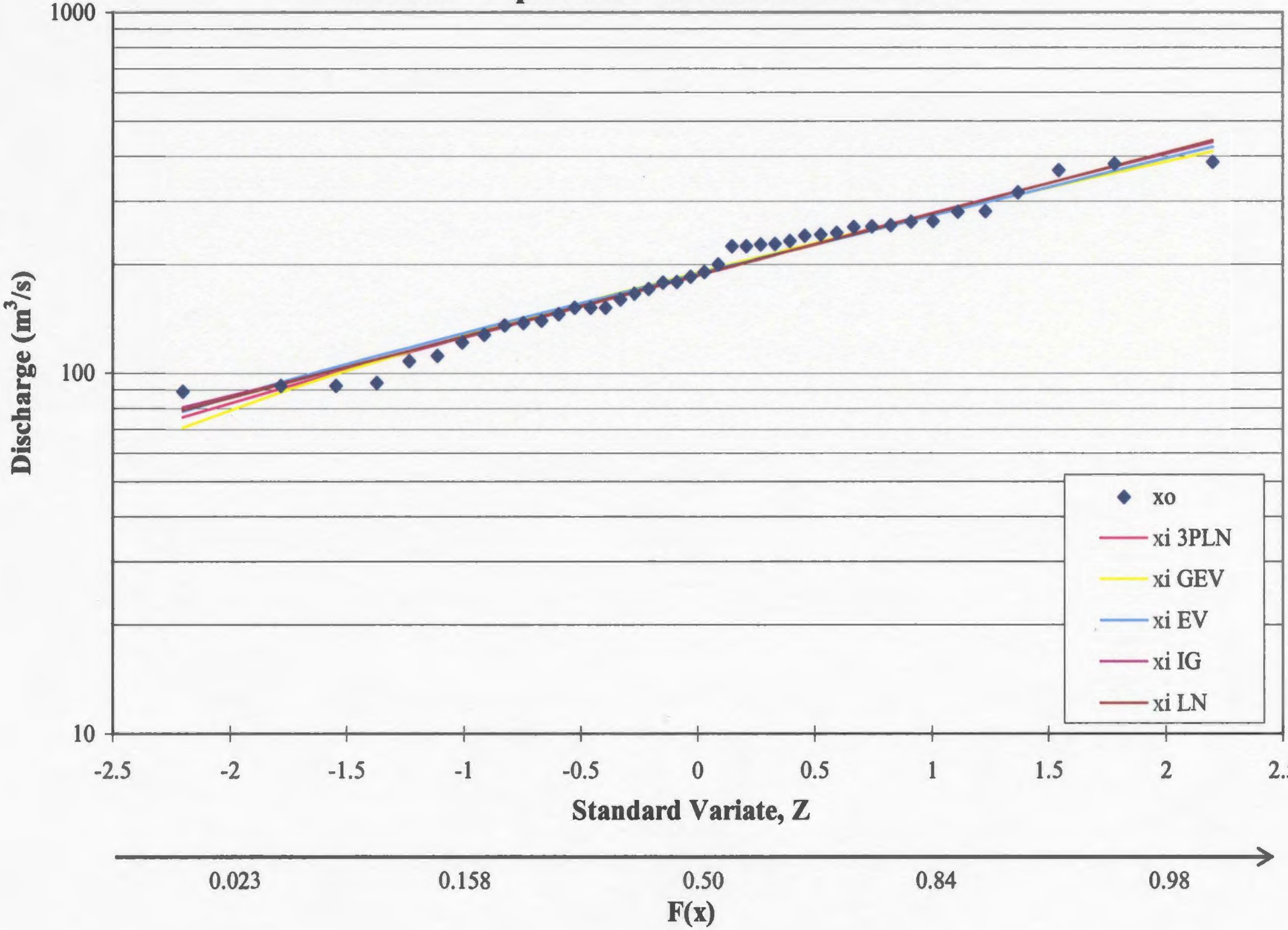

Comparison of Distribution Estimates to Observed Discharge Data
02ZG002 - Tides Brook below Freshwater Pond
Discharge (m3/s)
1000
100
10
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN

Comparison of Distribution Estimates to Observed Discharge Data
02ZG003 - Salmonier River near Lamaline

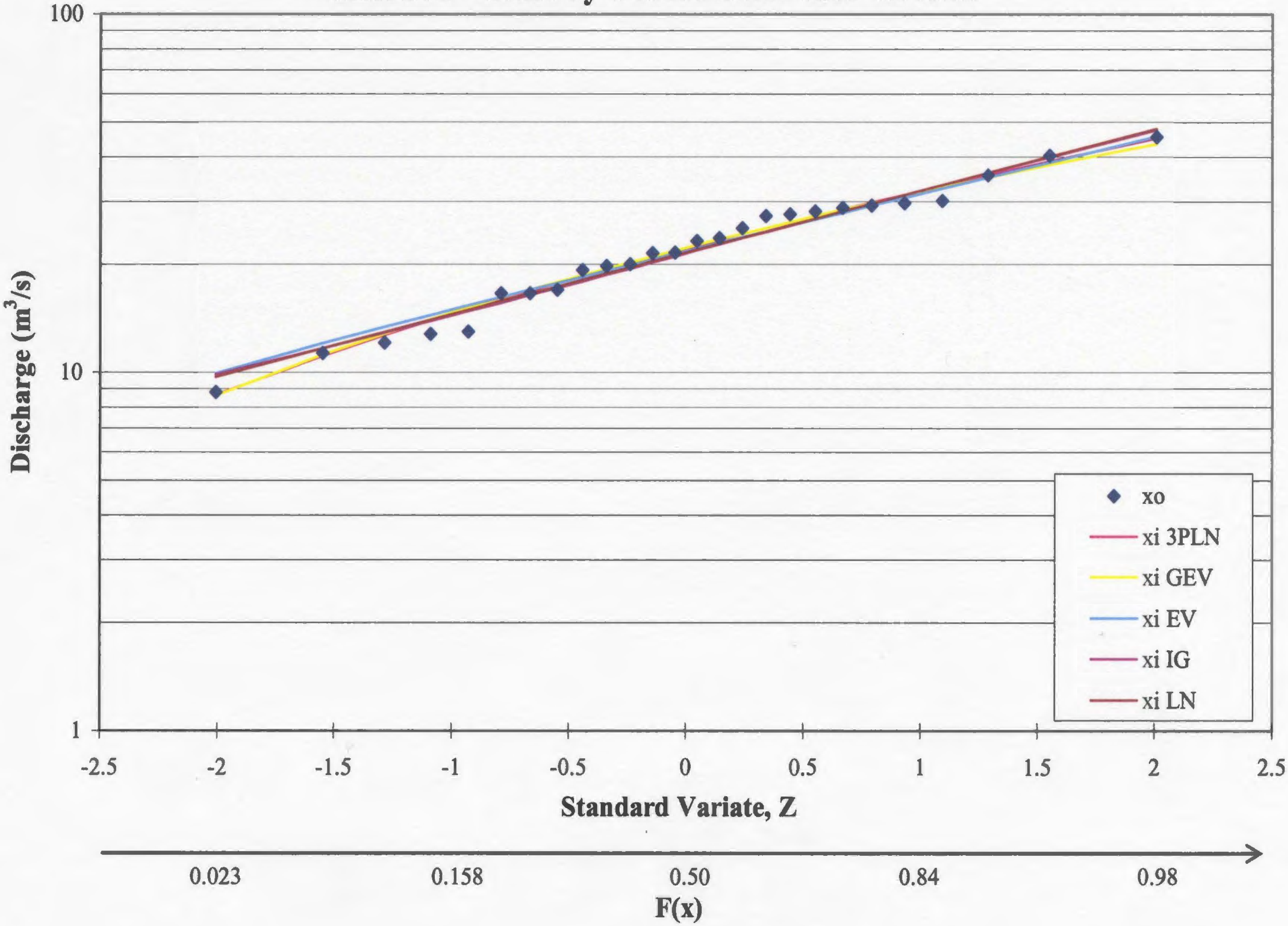



Comparison of Distribution Estimates to Observed Discharge Data
02ZG004 - Rattke Brook near Boat Harbour
Discharge (m3/s)
100
10
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN

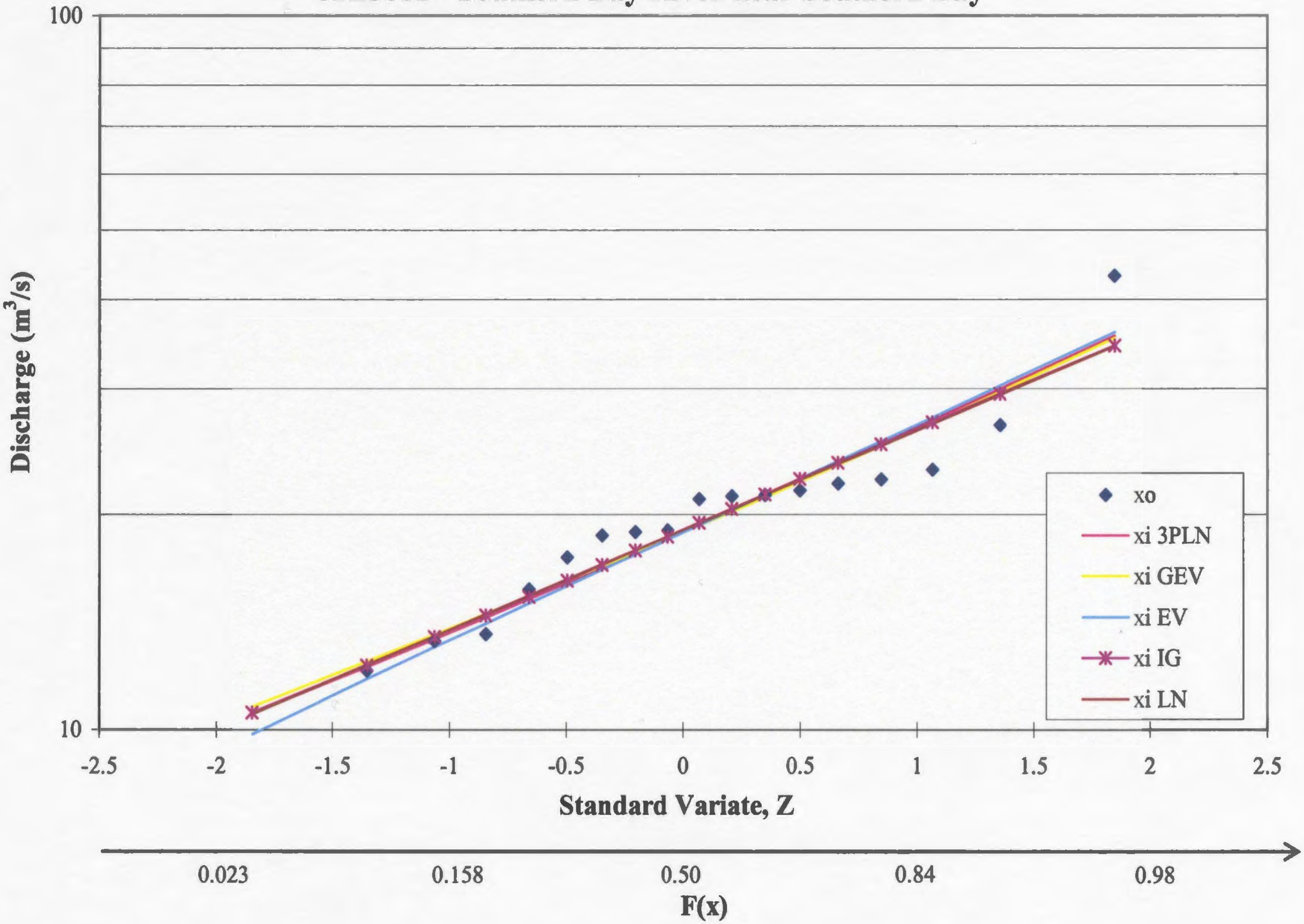
Comparison of Distribution Estimates to Observed Discharge Data
02ZH001 - Pipers Hole River at Mothers Brook

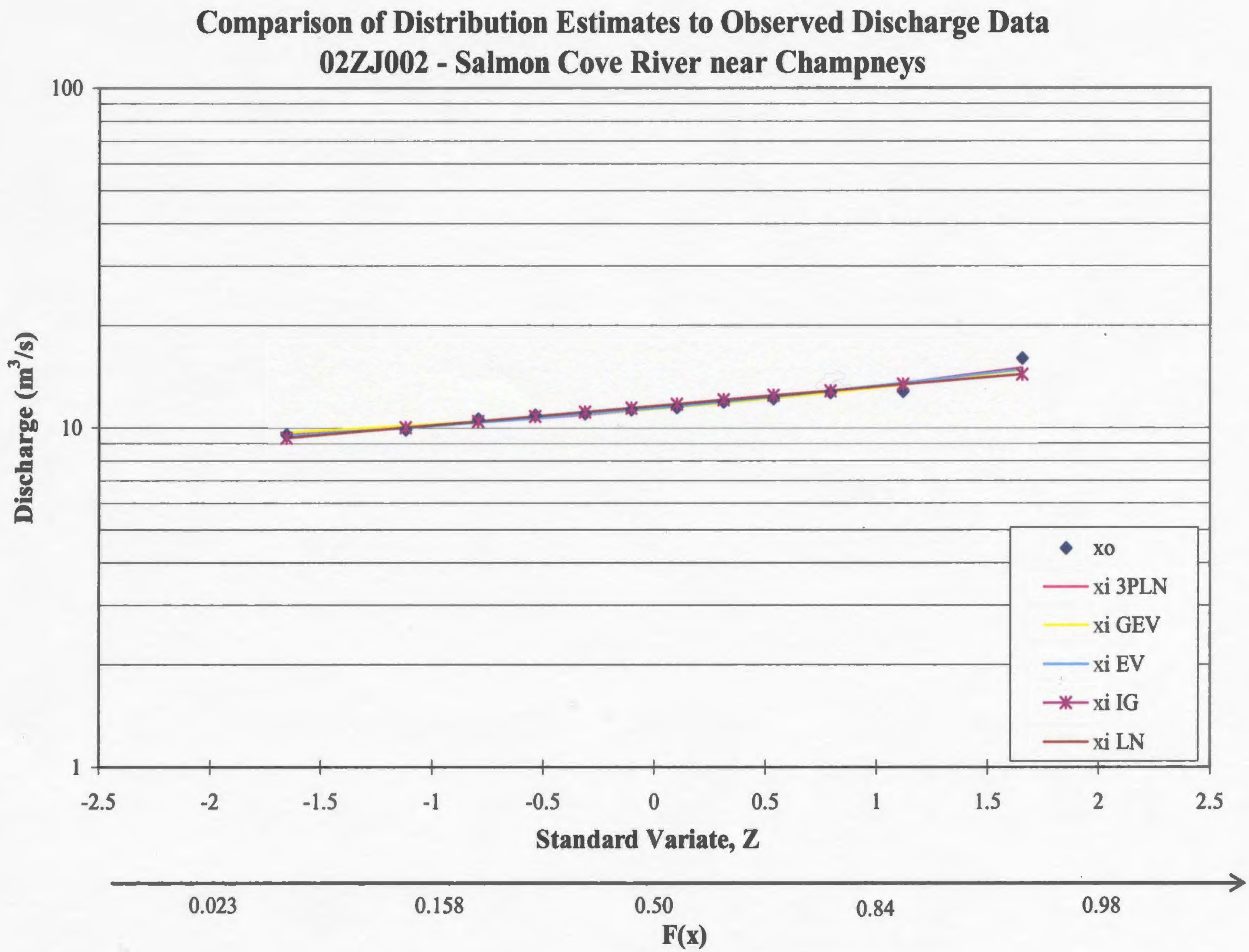


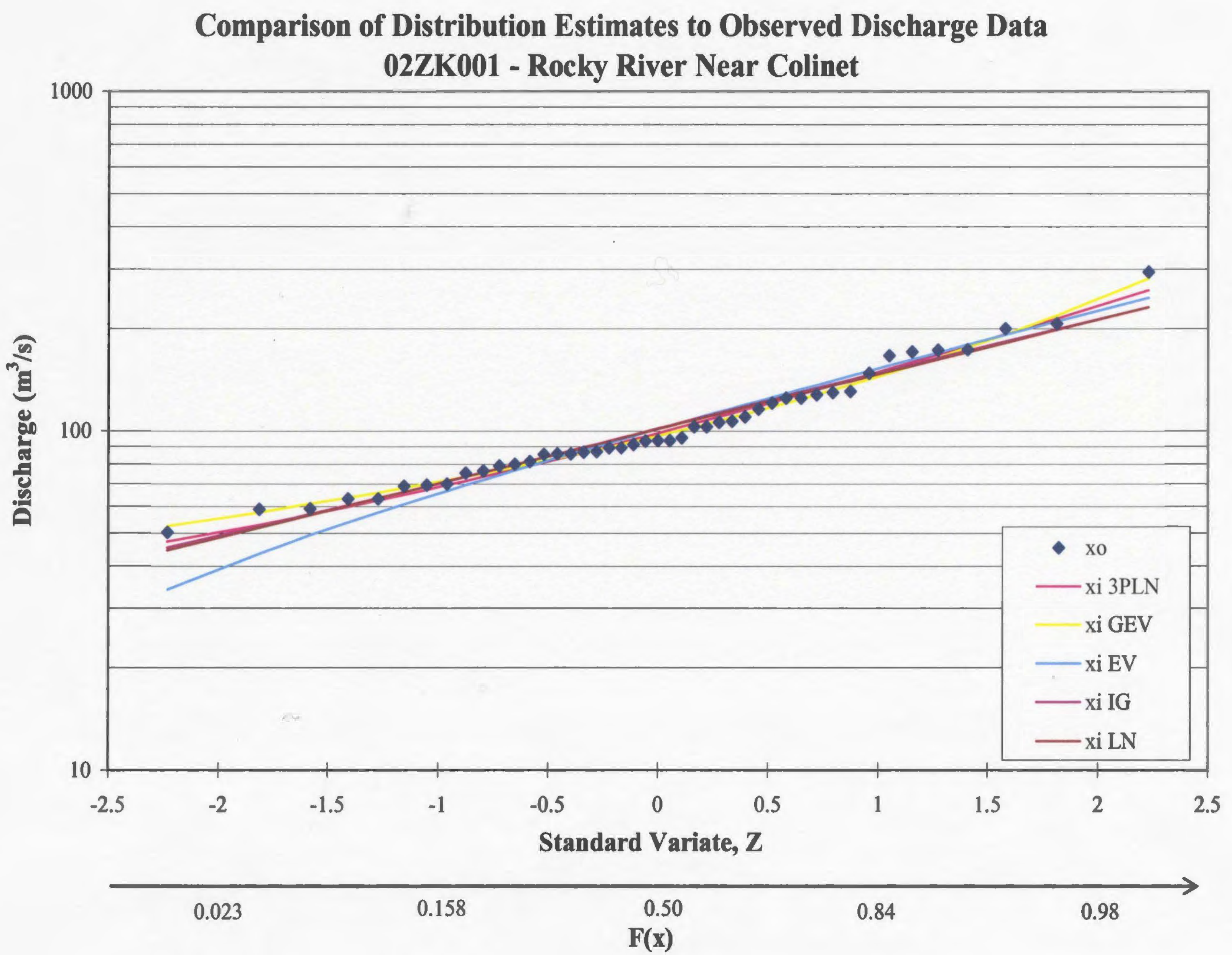
Comparison of Distribution Estimates to Observed Discharge Data
02ZH002 - Come by Chance River near Goobies



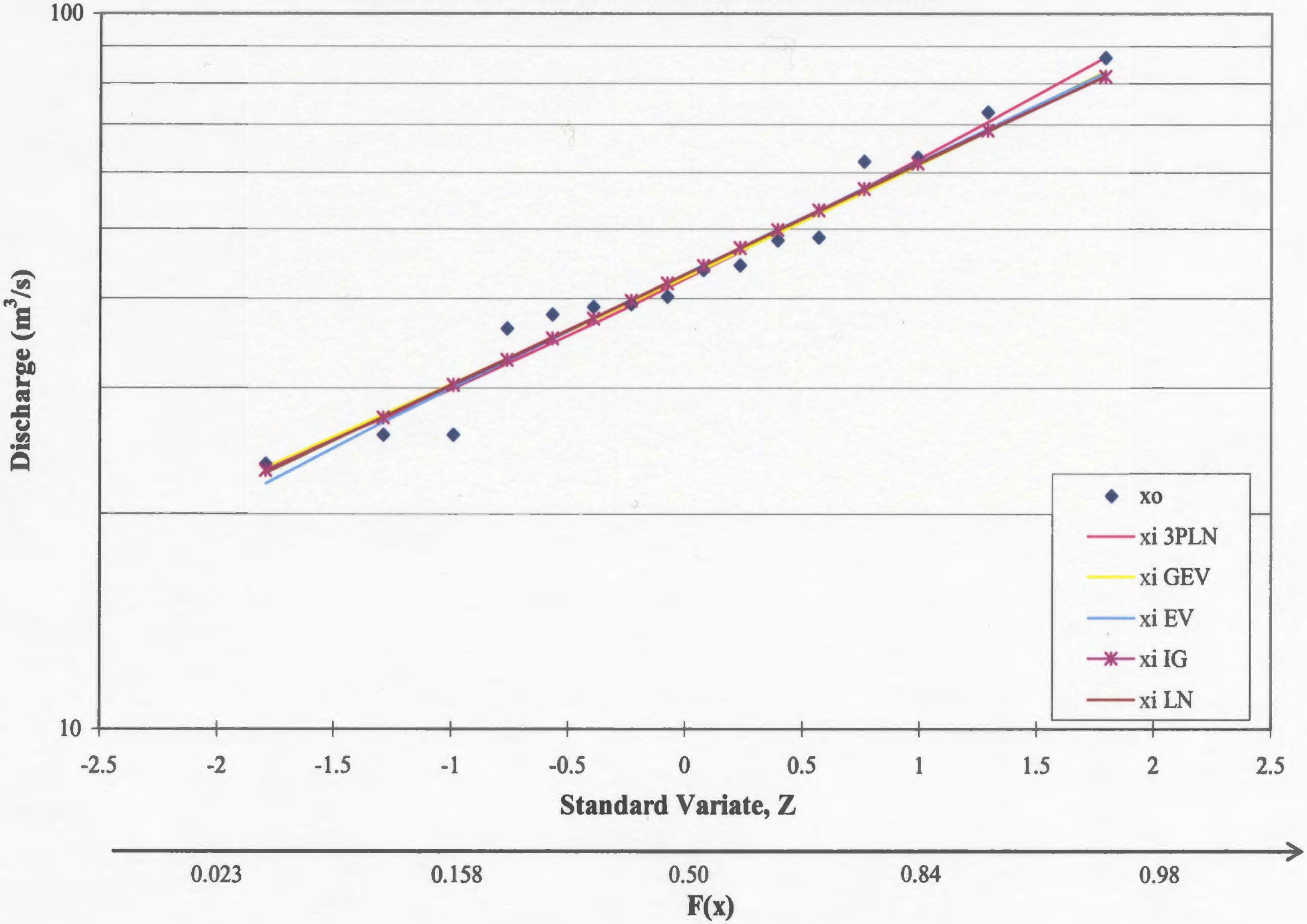
Comparison of Distribution Estimates to Observed Discharge Data
02ZJ001 - Southern Bay River near Southern Bay



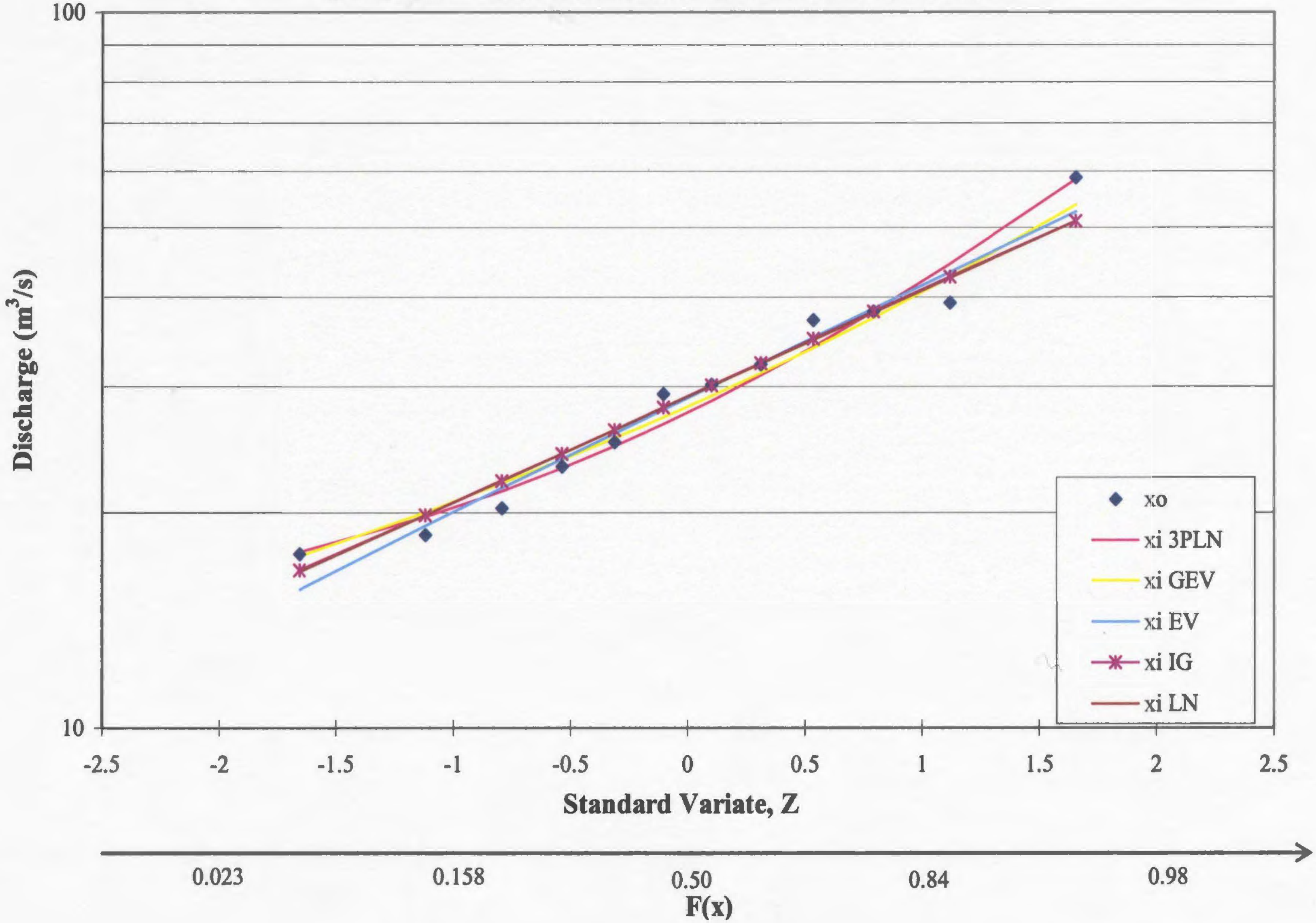

Comparison of Distribution Estimates to Observed Discharge Data
02ZJ002 - Salmon Cove River near Champneys
Discharge (m^3/s)
100
10
1
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN


Comparison of Distribution Estimates to Observed Discharge Data
02ZK001 - Rocky River Near Colinet
1000
100
10
Discharge (m3/s)
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN

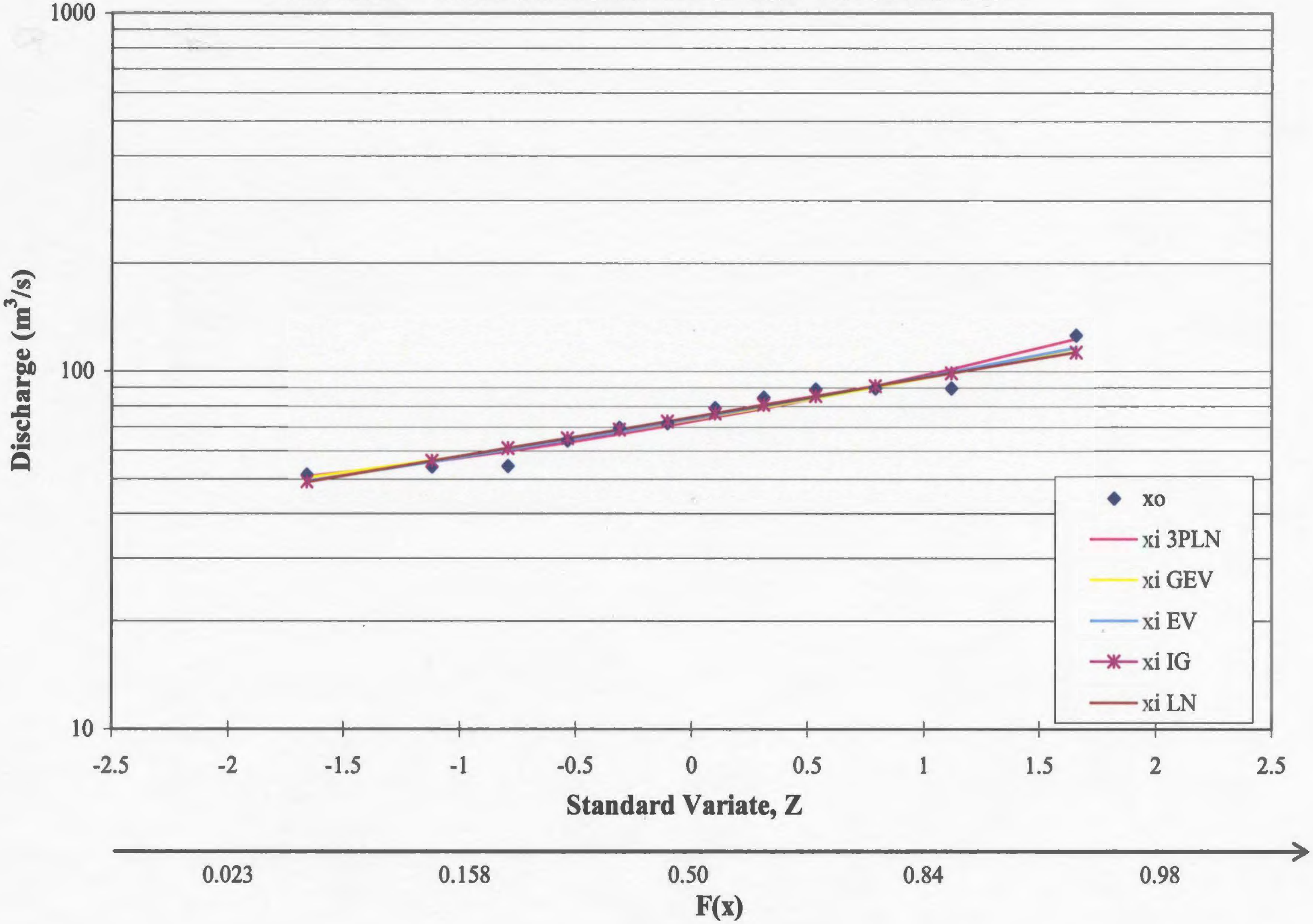
Comparison of Distribution Estimates to Observed Discharge Data
02ZK002 - Northeast River near Placentia



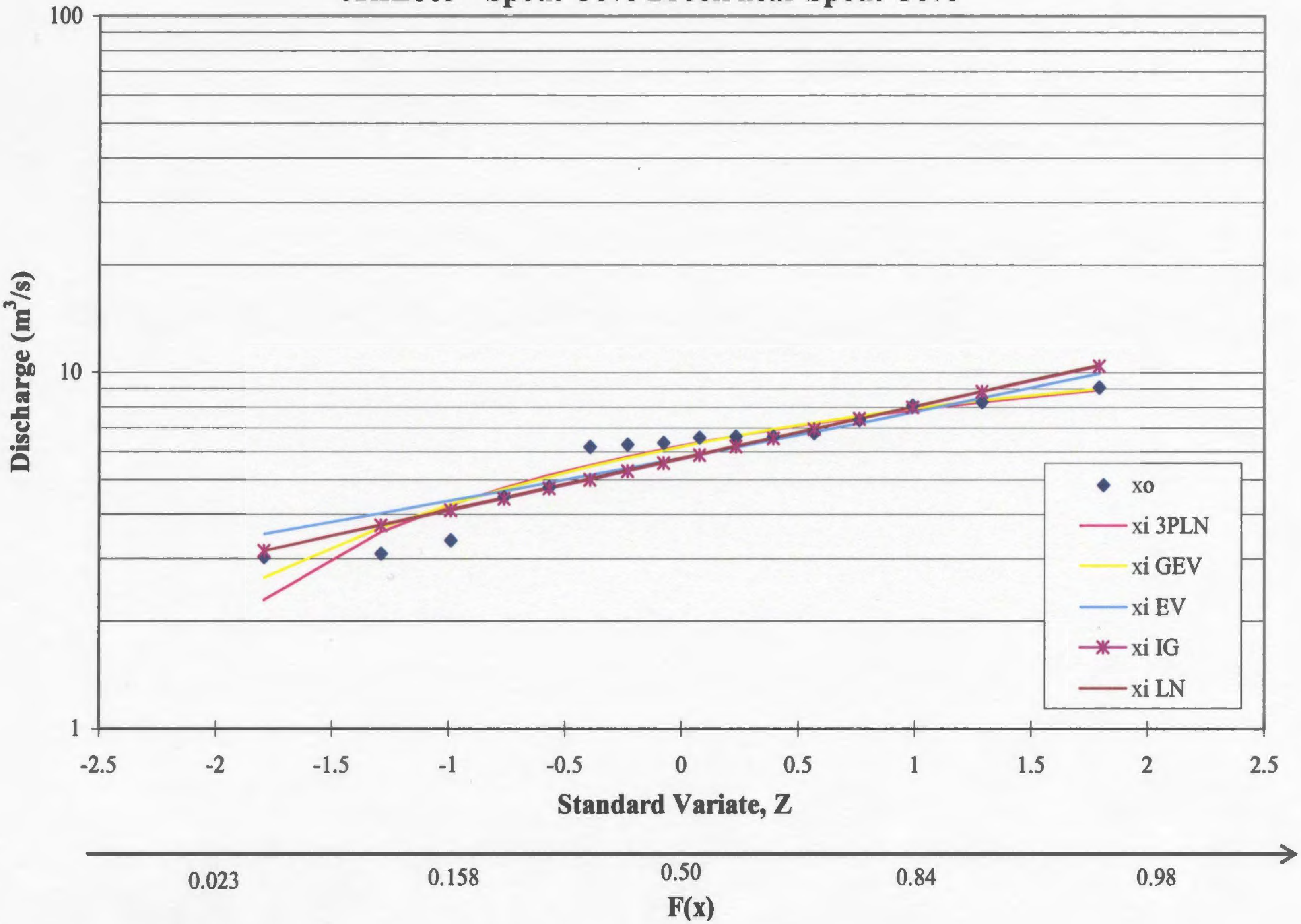
Comparison of Distribution Estimates to Observed Discharge Data
02ZK003 - Little Barachois River near Placentia



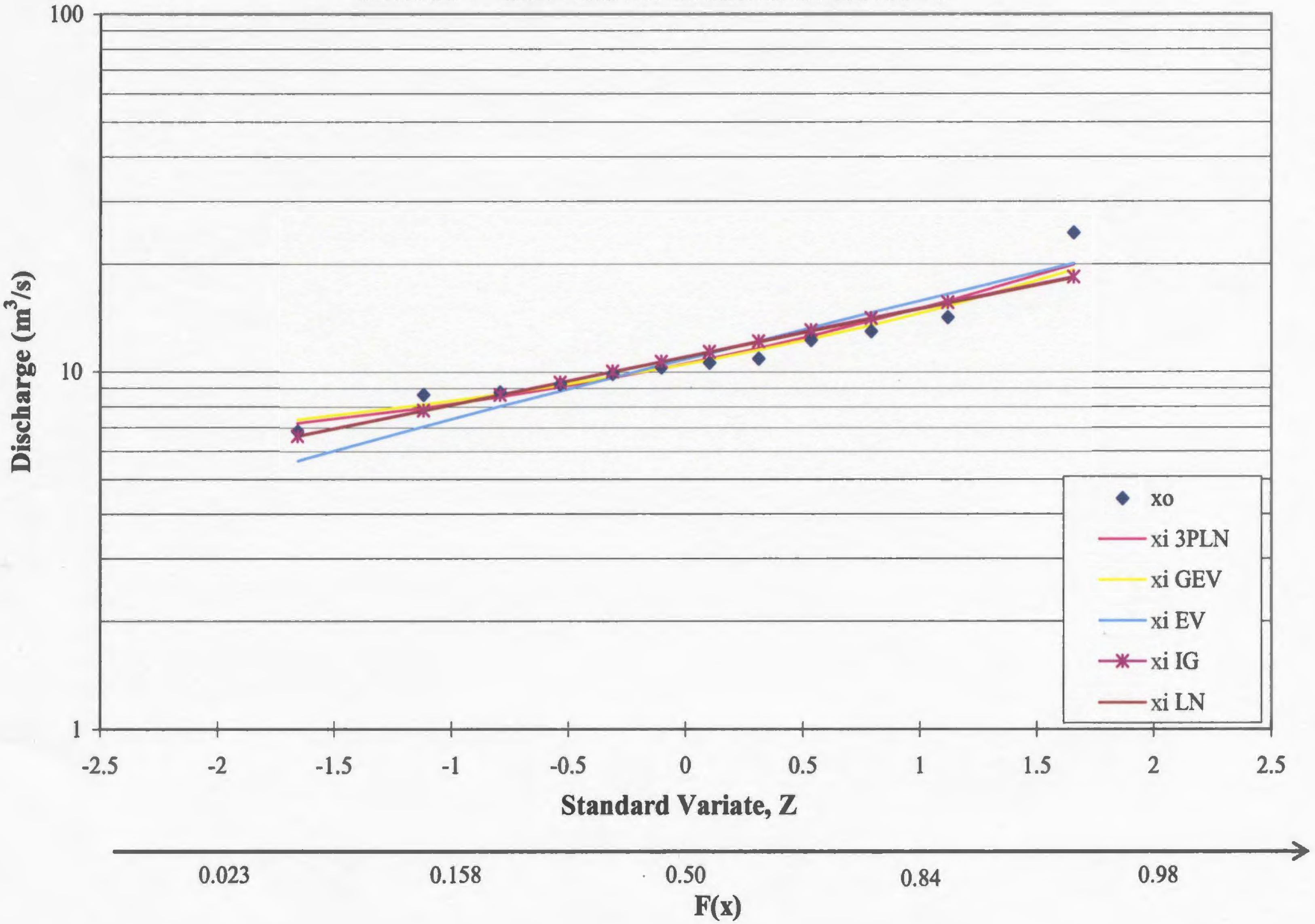
Comparison of Distribution Estimates to Observed Discharge Data
02ZK004 - Little Salmonier River near North Harbour

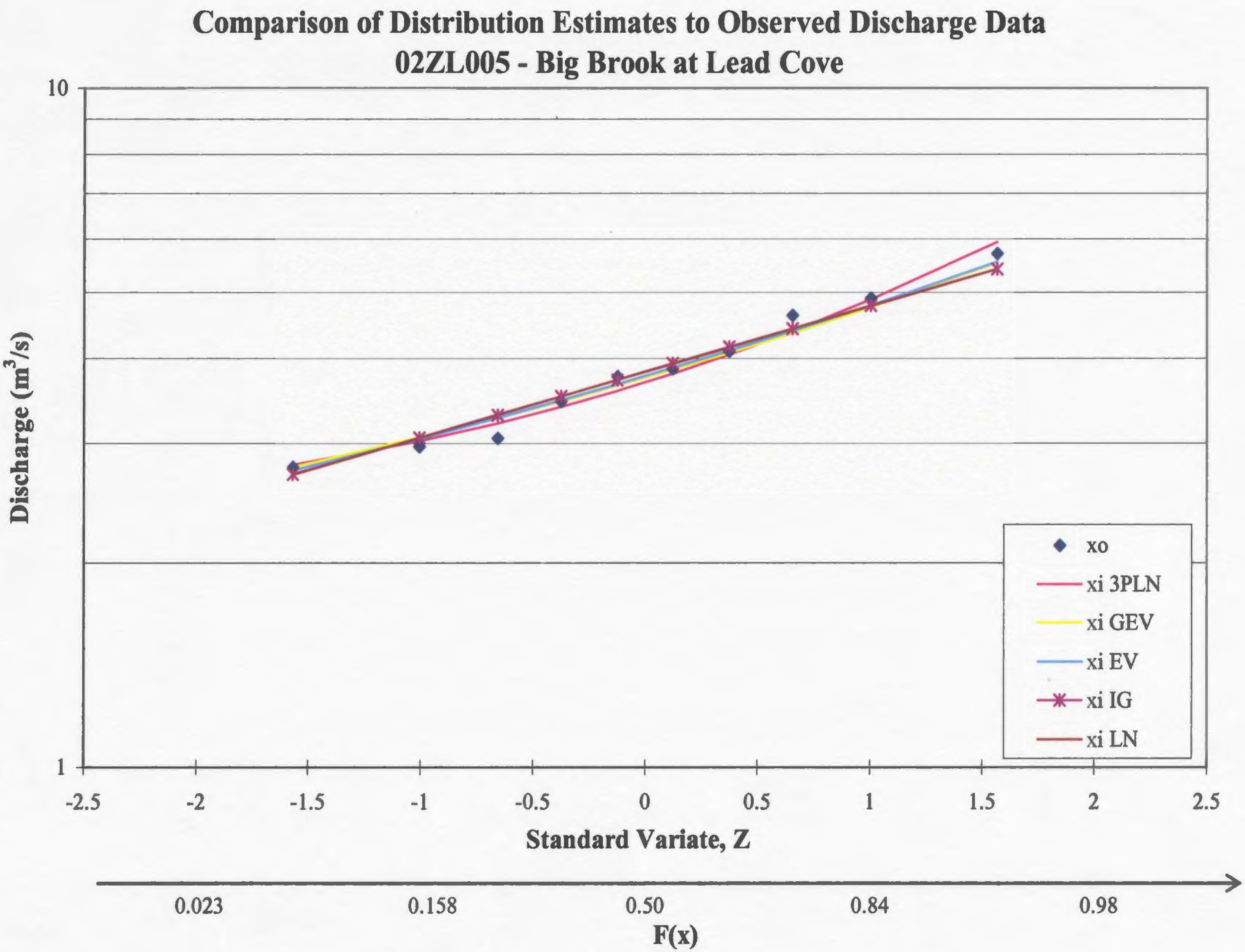


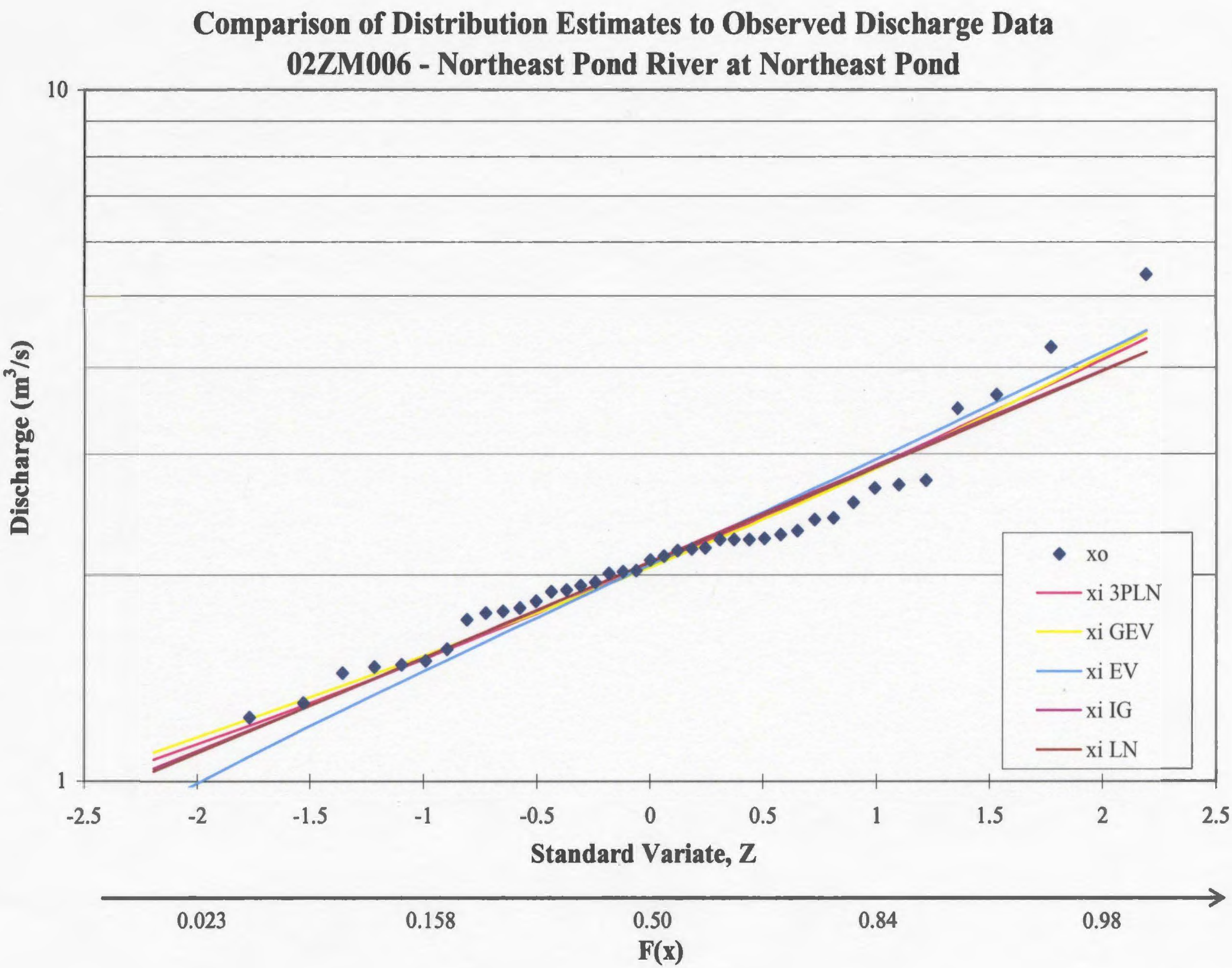
Comparison of Distribution Estimates to Observed Discharge Data
02ZL003 - Spout Cove Brook near Spout Cove



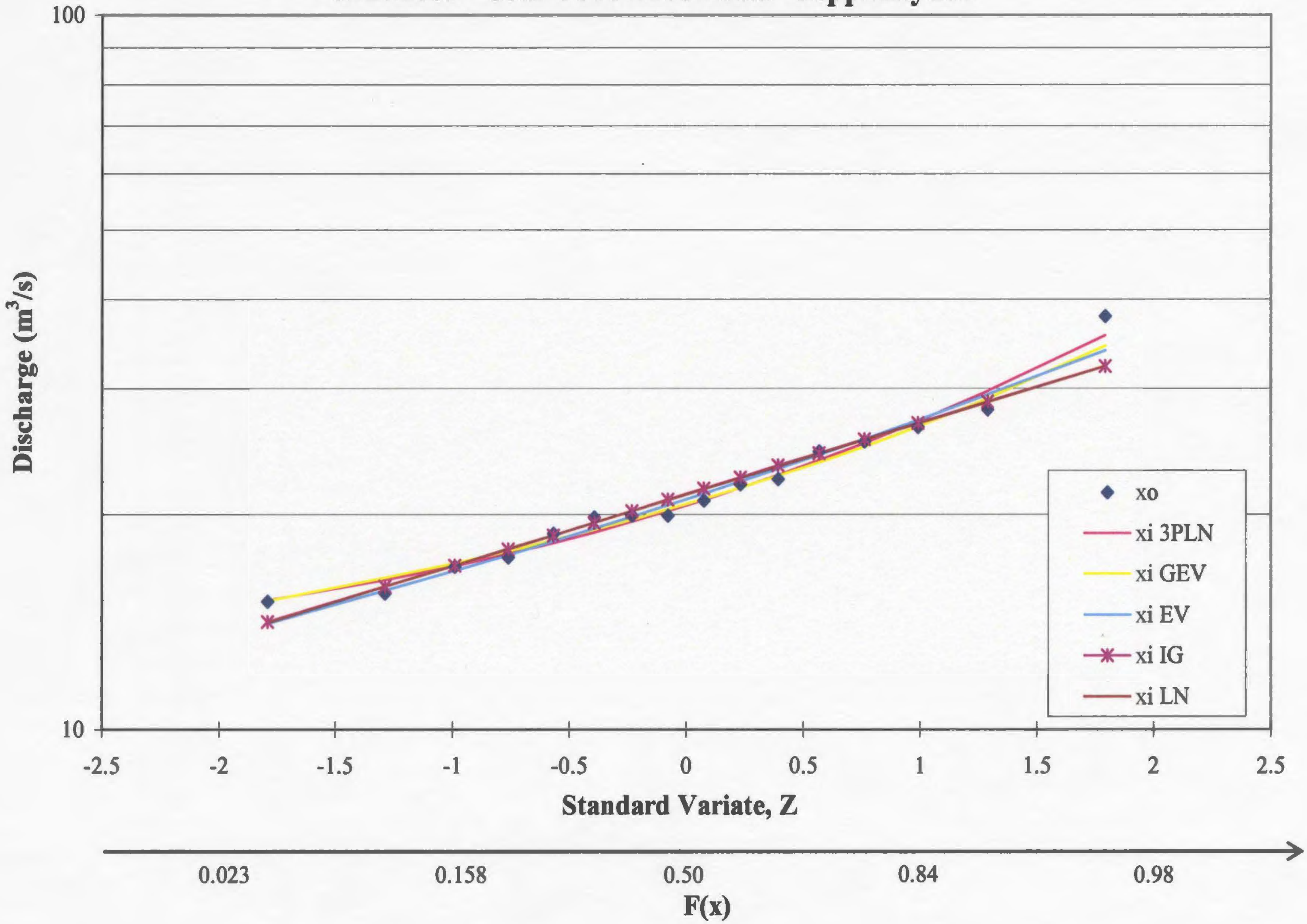
Comparison of Distribution Estimates to Observed Discharge Data
02ZL004 - Shearstown Brook at Shearstown

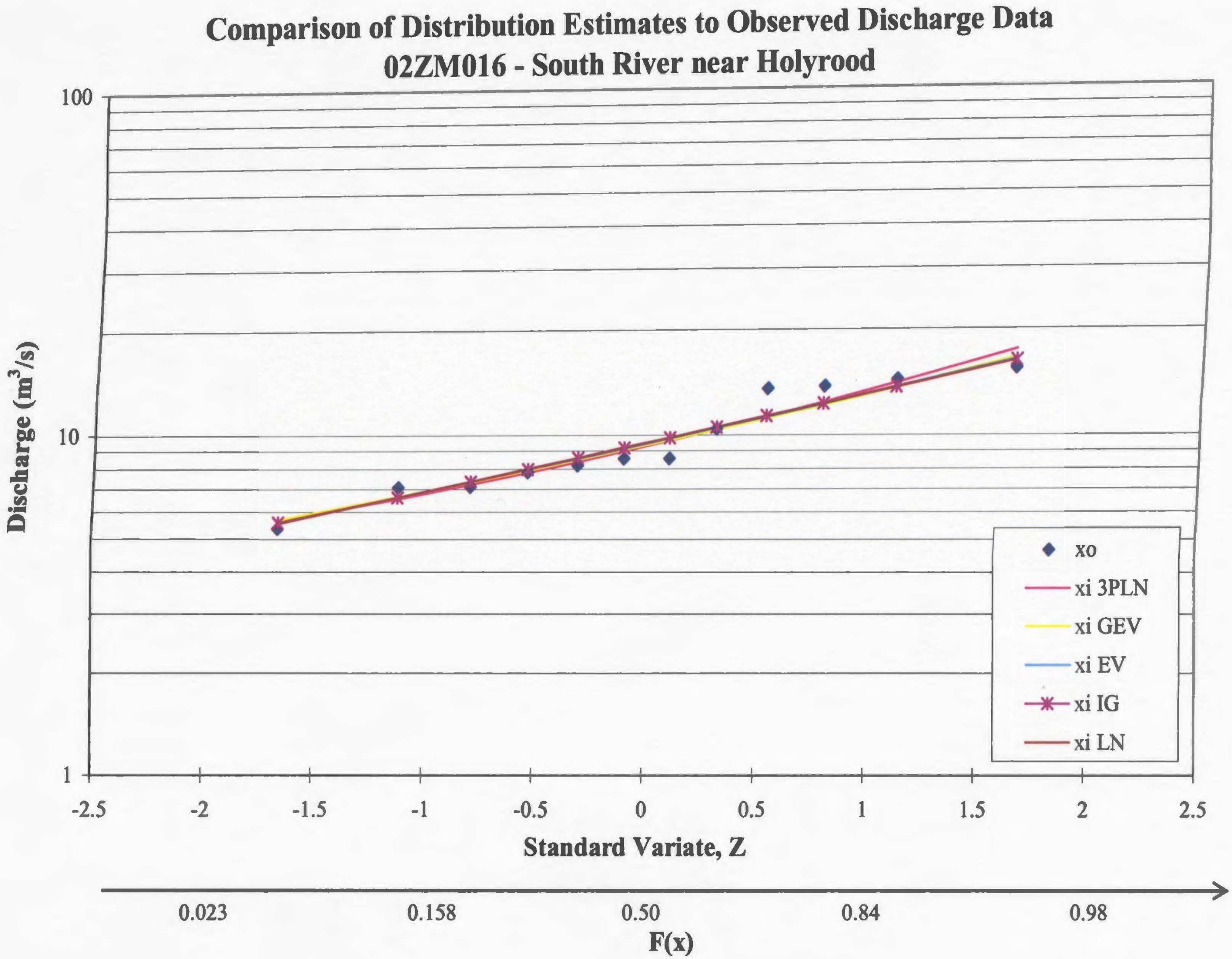



Comparison of Distribution Estimates to Observed Discharge Data
02ZL005 - Big Brook at Lead Cove
Discharge (m3/s)
10
1
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN

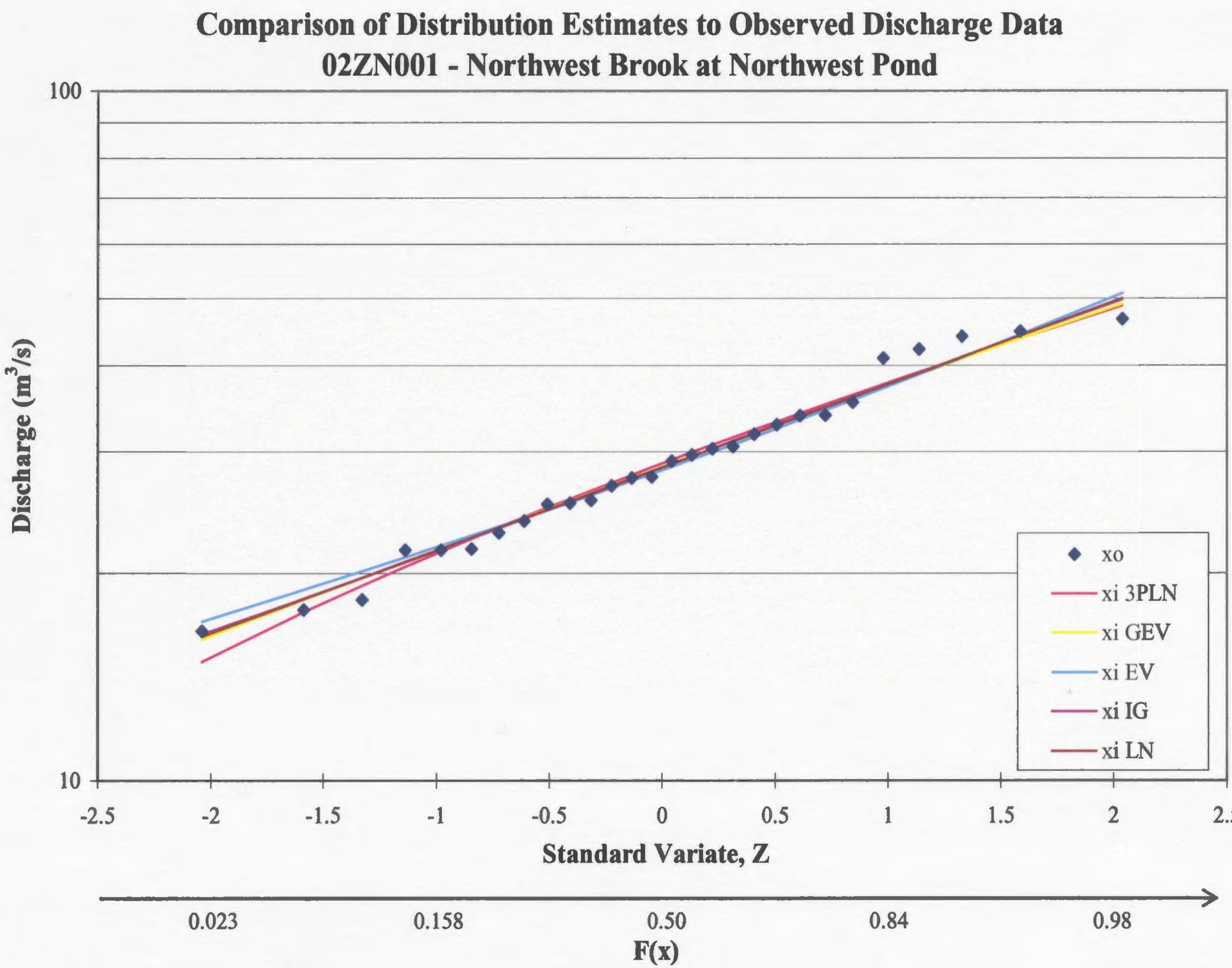

Comparison of Distribution Estimates to Observed Discharge Data
02ZM006 - Northeast Pond River at Northeast Pond
Discharge (m3/s)
10
1
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN

Comparison of Distribution Estimates to Observed Discharge Data
02ZM009 - Seal Cove Brook near Cappahayden

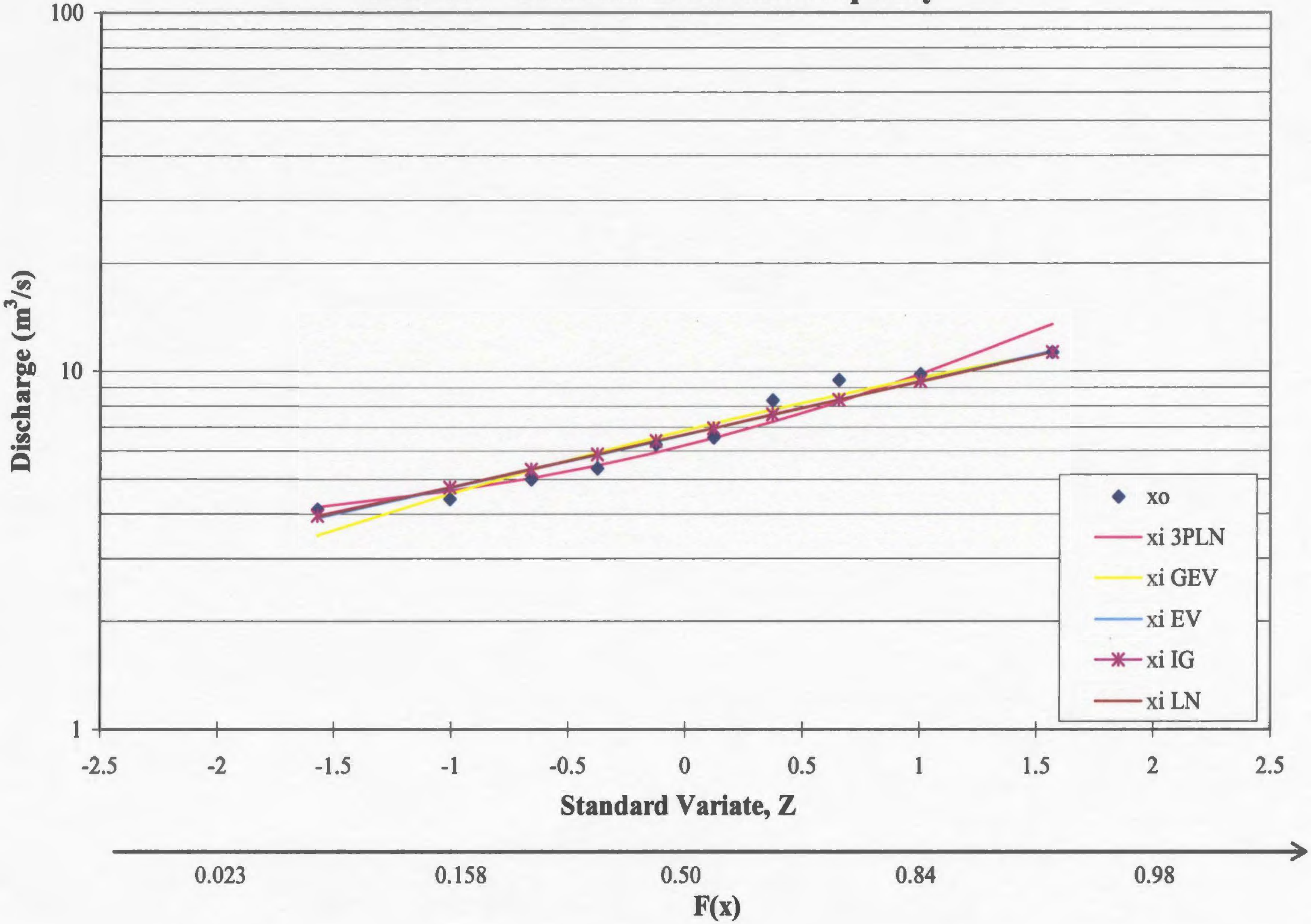




Comparison of Distribution Estimates to Observed Discharge Data
02ZM016 - South River near Holyrood
Discharge (m3/s)
100
10
1
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN

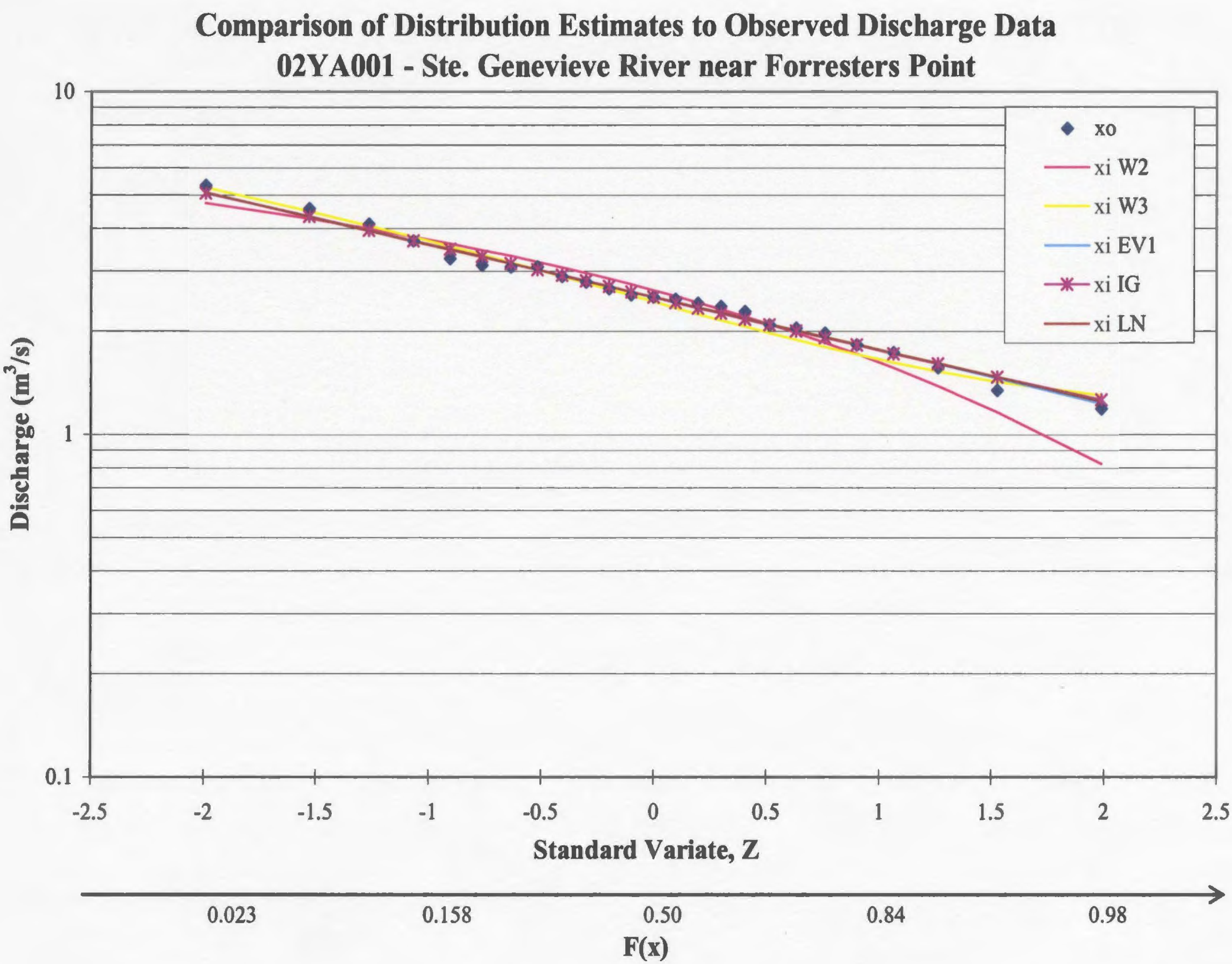

Comparison of Distribution Estimates to Observed Discharge Data
02ZN001 - Northwest Brook at Northwest Pond
100
10
Discharge (m³/s)
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi 3PLN
xi GEV
xi EV
xi IG
xi LN

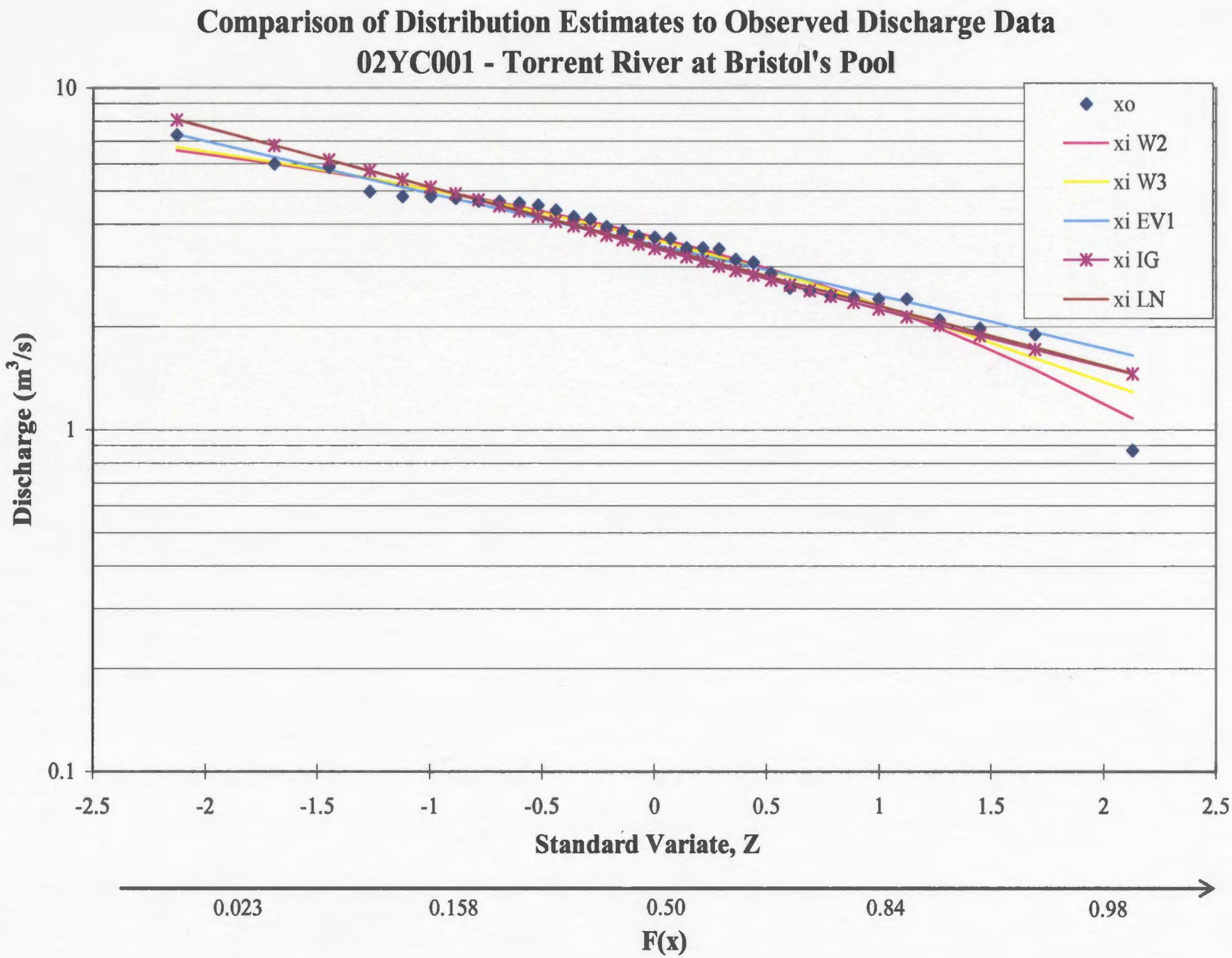
Comparison of Distribution Estimates to Observed Discharge Data
02ZN002 - St. Shotts River neat Trepassey



Appendix D

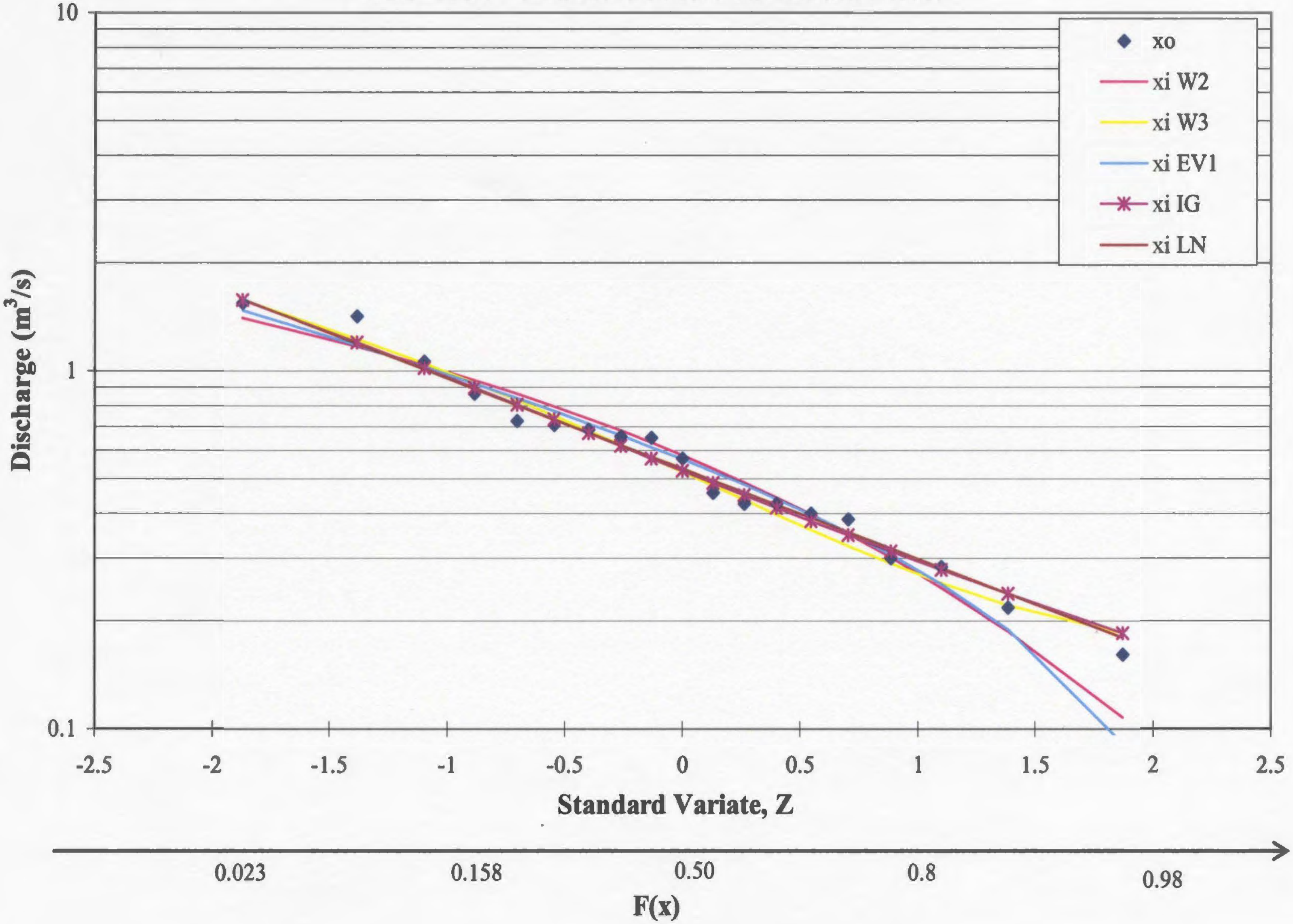
Single Site Analysis Results – Low Flow Analysis Data Plots

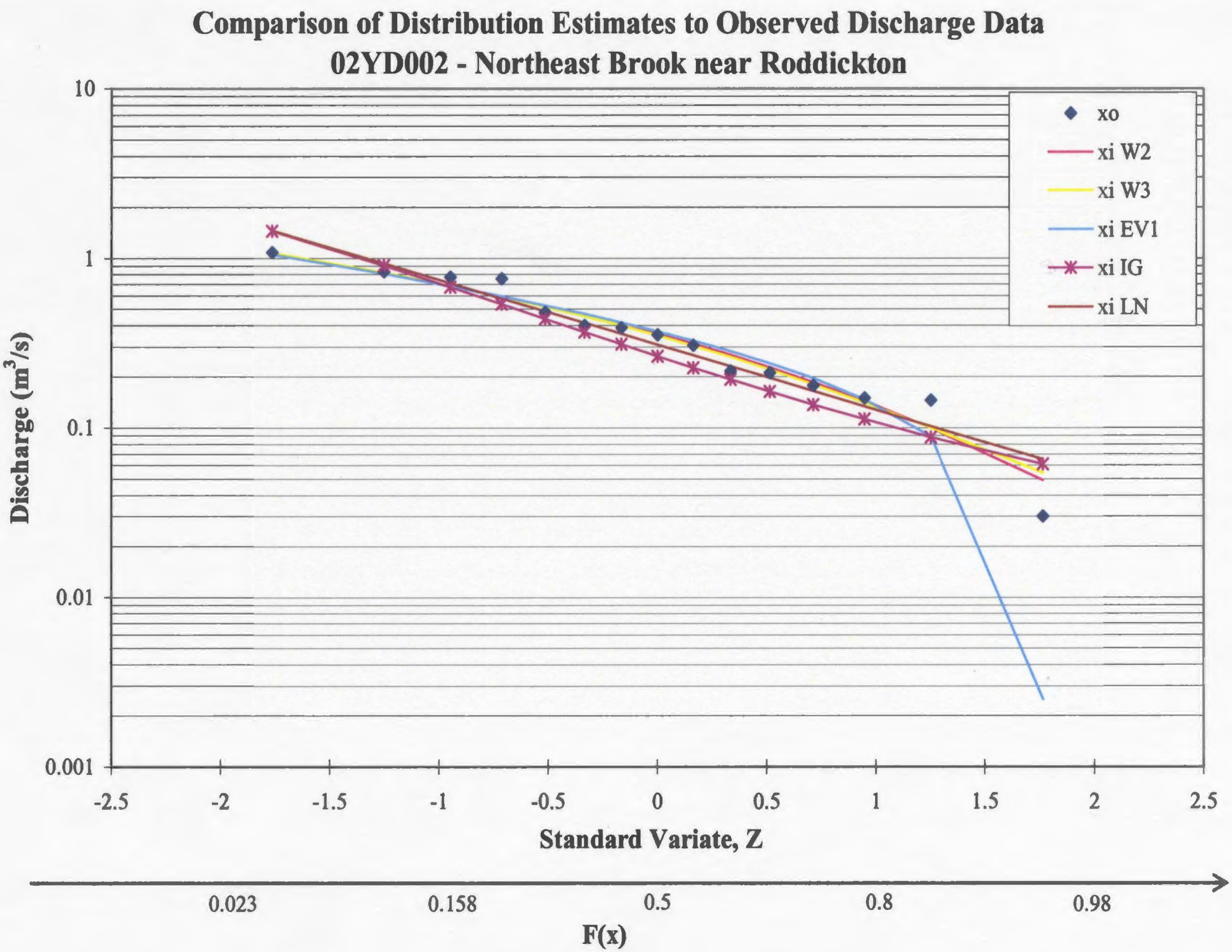

Comparison of Distribution Estimates to Observed Discharge Data
02YA001 - Ste. Genevieve River near Forresters Point
Discharge (m³/s)
10
1
0.1
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi W2
xi W3
xi EV1
xi IG
xi LN



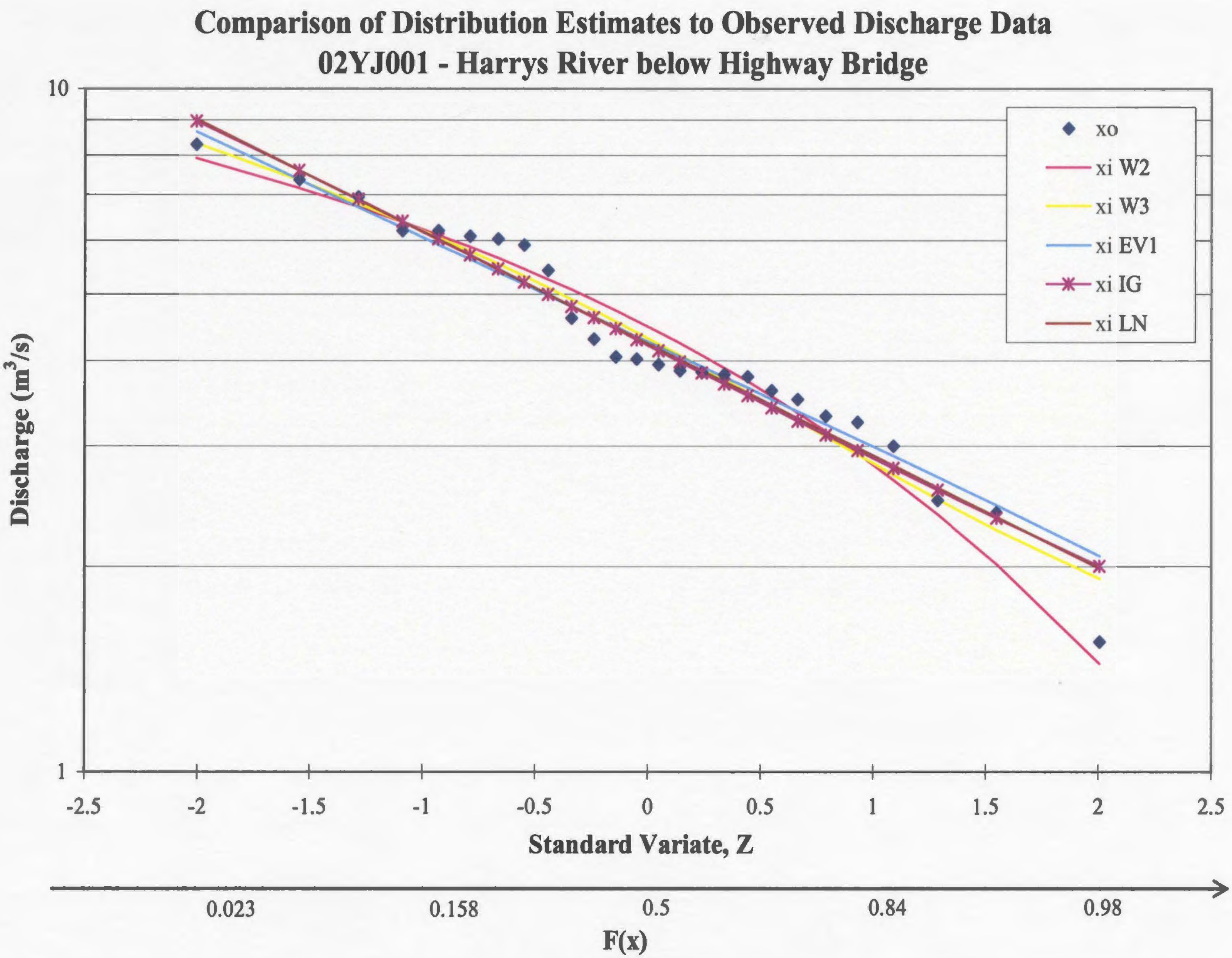
Comparison of Distribution Estimates to Observed Discharge Data
02YC001 - Torrent River at Bristol's Pool
Discharge (m³/s)
10
1
0.1
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi W2
xi W3
xi EV1
xi IG
xi LN

Comparison of Distribution Estimates to Observed Discharge Data
02YD001 - Beaver Brook near Roddickton

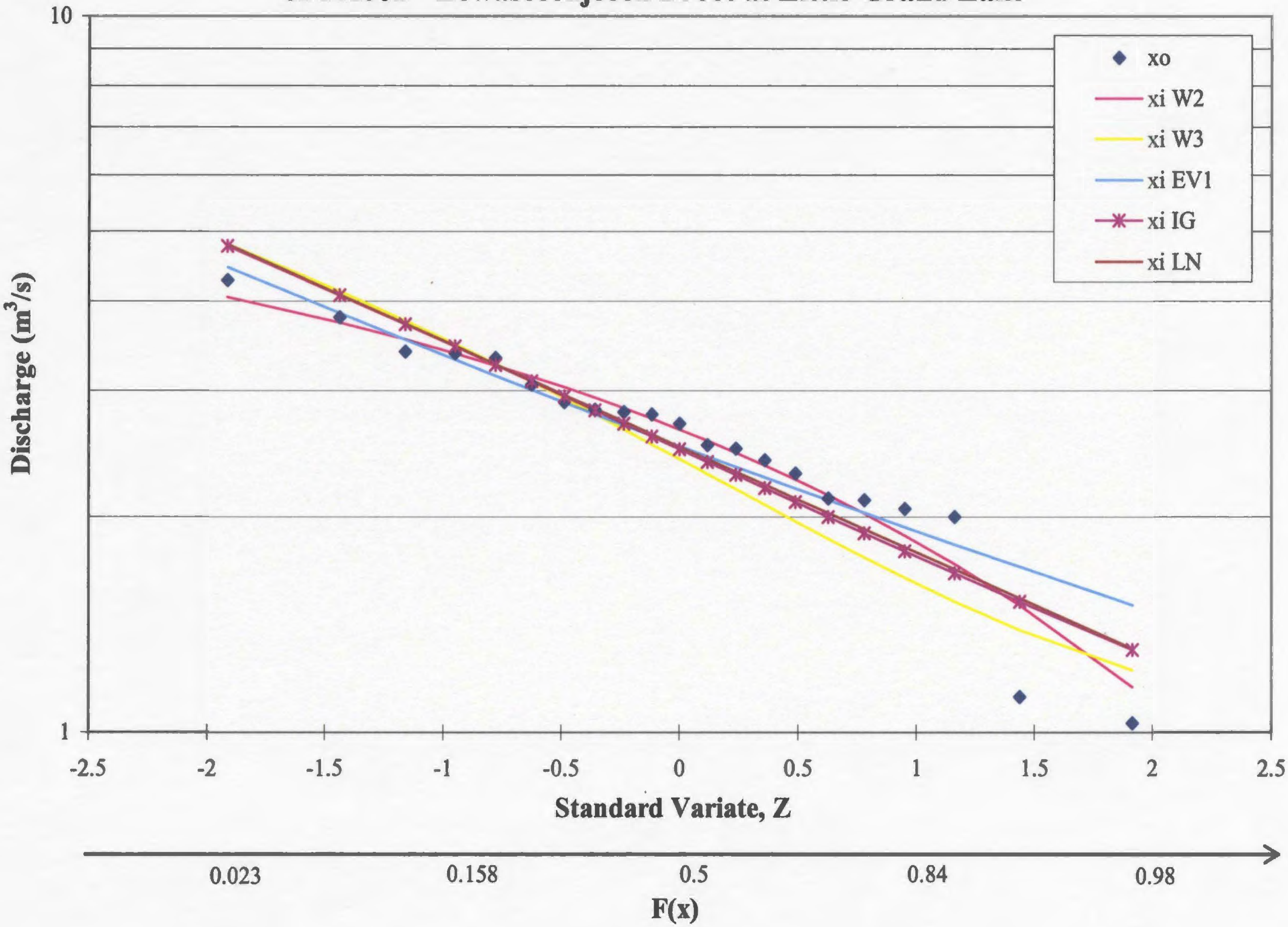


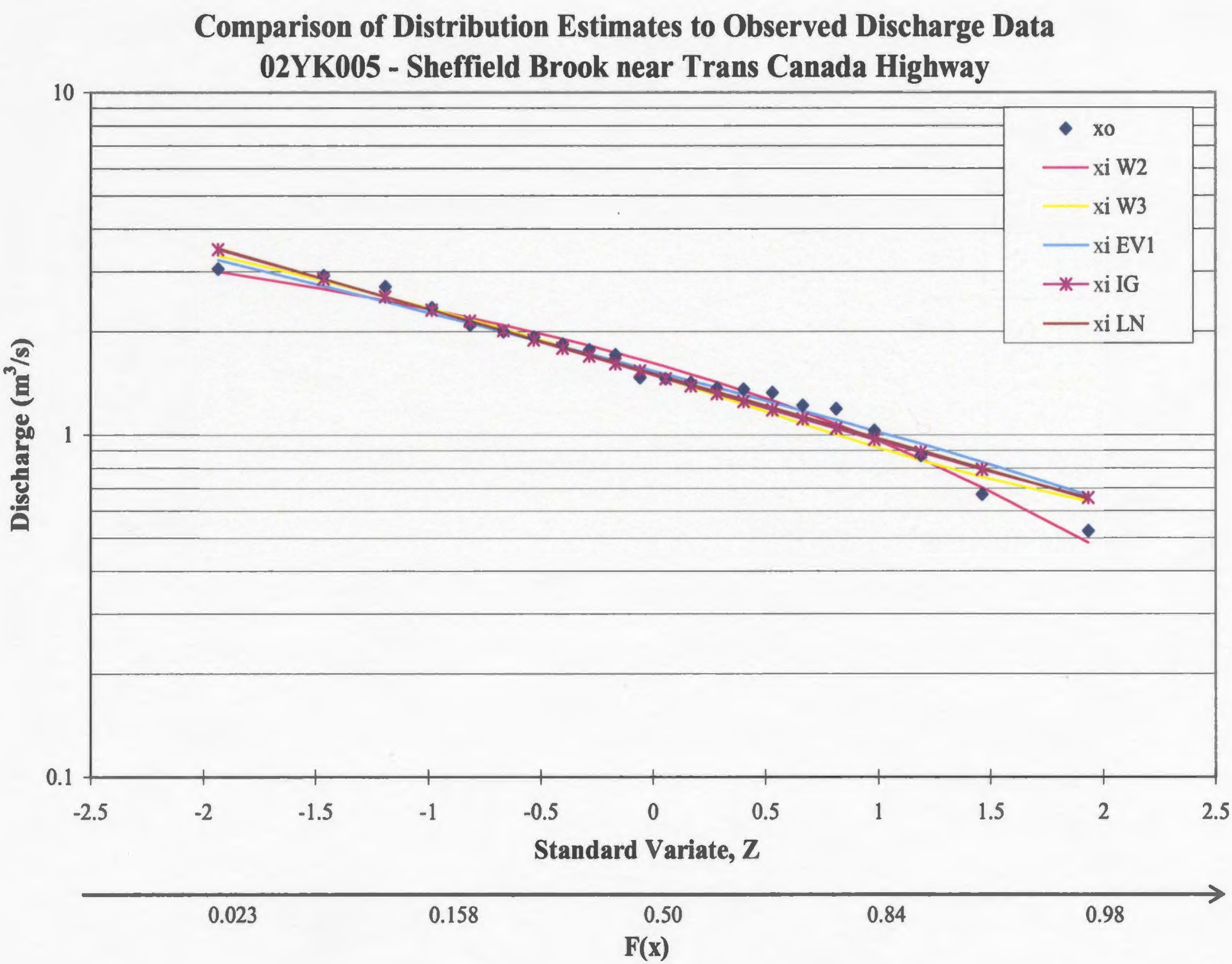


Comparison of Distribution Estimates to Observed Discharge Data
02YD002 - Northeast Brook near Roddickton
Discharge (m3/s)
10
1
0.1
0.01
0.001
xo
xi W2
xi W3
xi EV1
xi IG
xi LN
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.5
0.8
0.98
F(x)

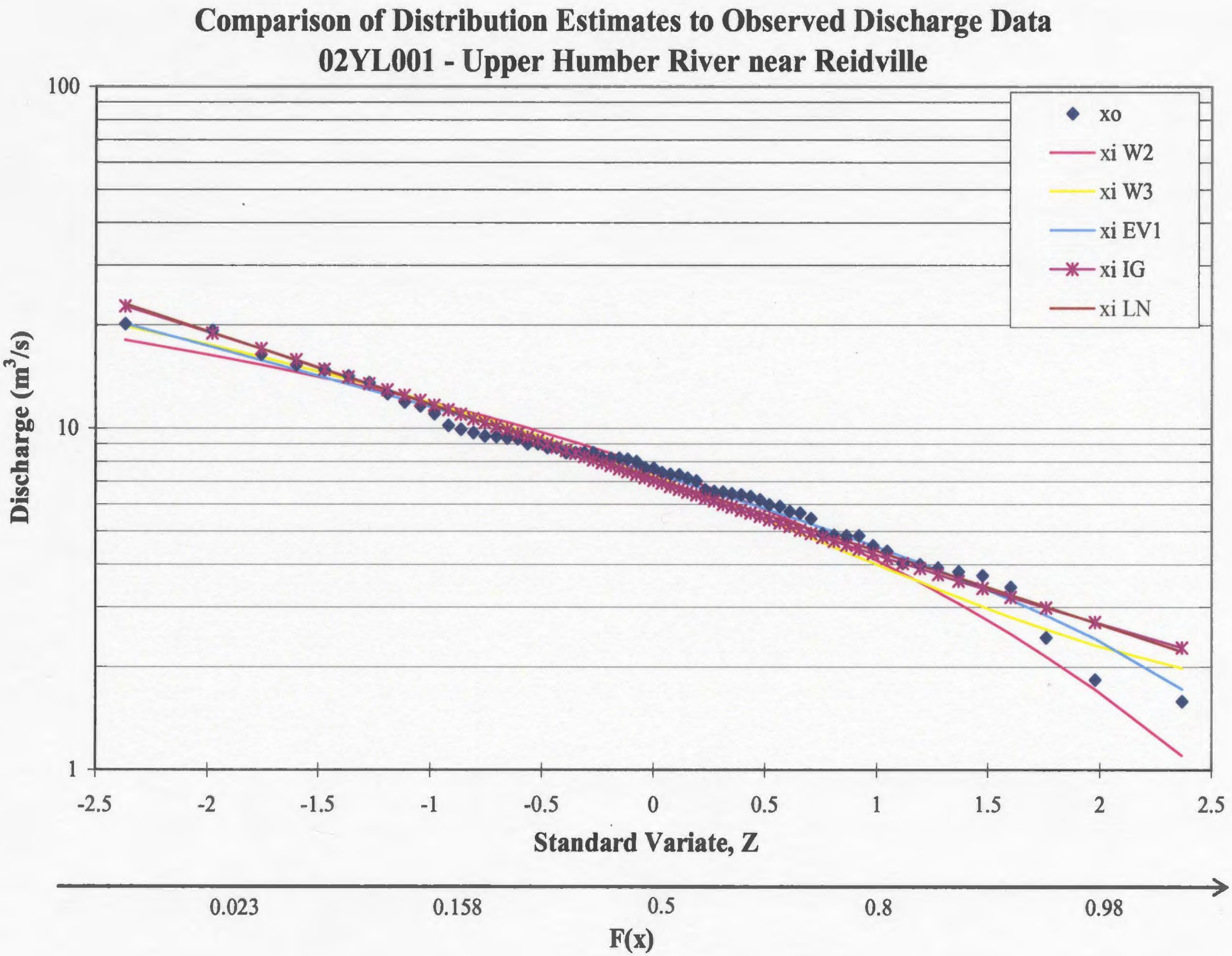

Comparison of Distribution Estimates to Observed Discharge Data
02YJ001 - Harrys River below Highway Bridge
10
1
Discharge (m3/s)
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.5
0.84
0.98
F(x)
xo
xi W2
xi W3
xi EV1
xi IG
xi LN

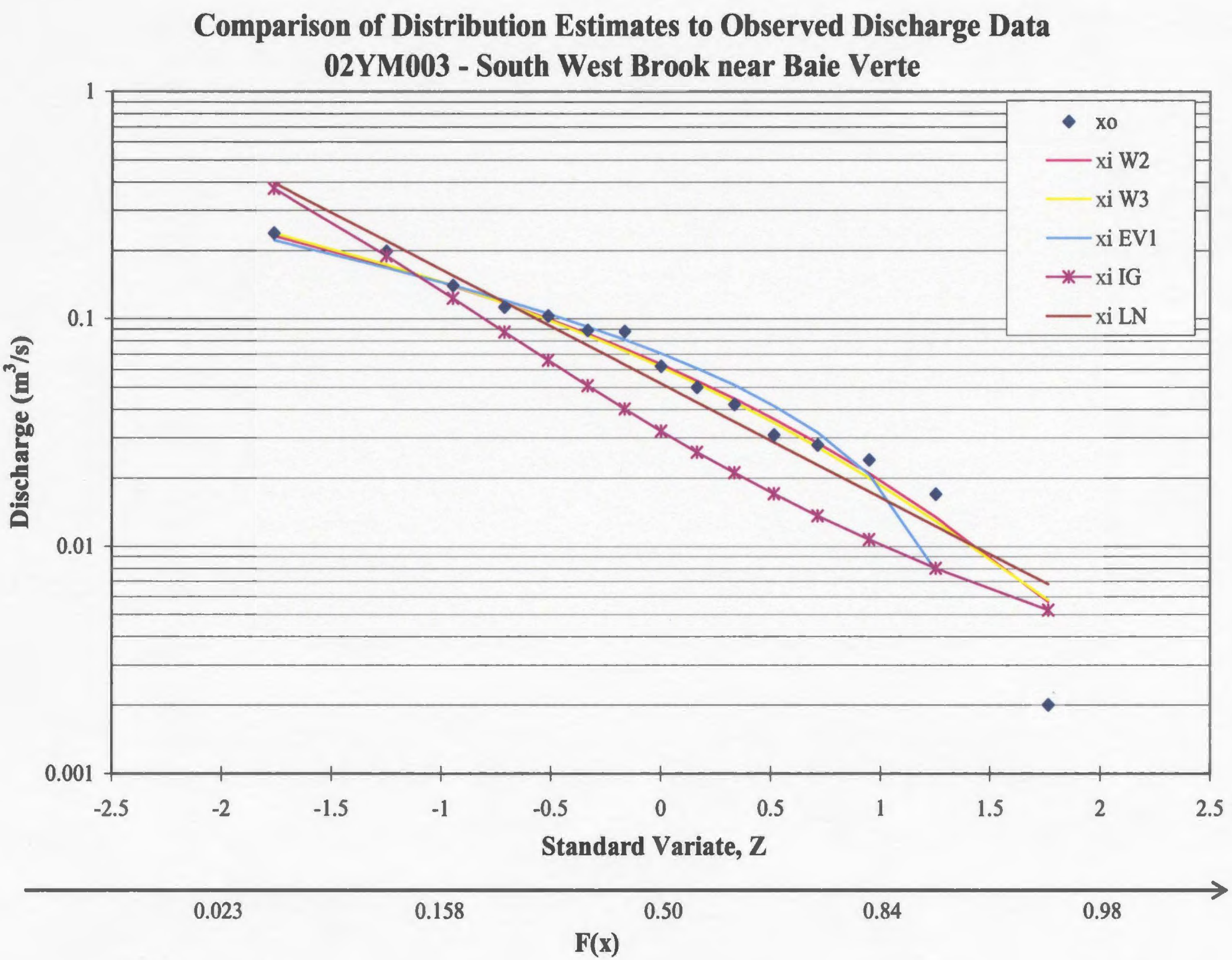
Comparison of Distribution Estimates to Observed Discharge Data
02YK002 - Lewaseechjeech Broot at Little Grand Lake



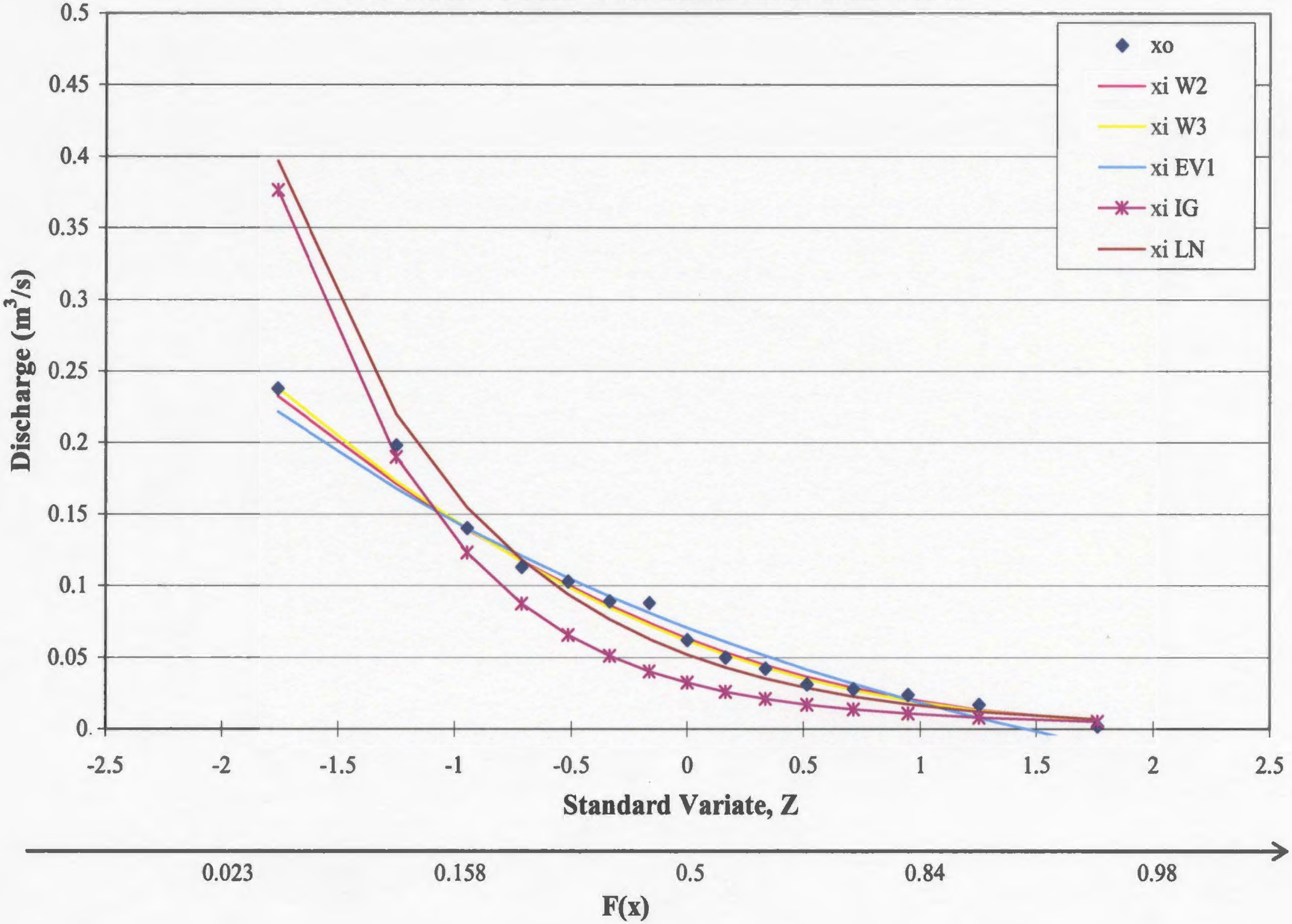


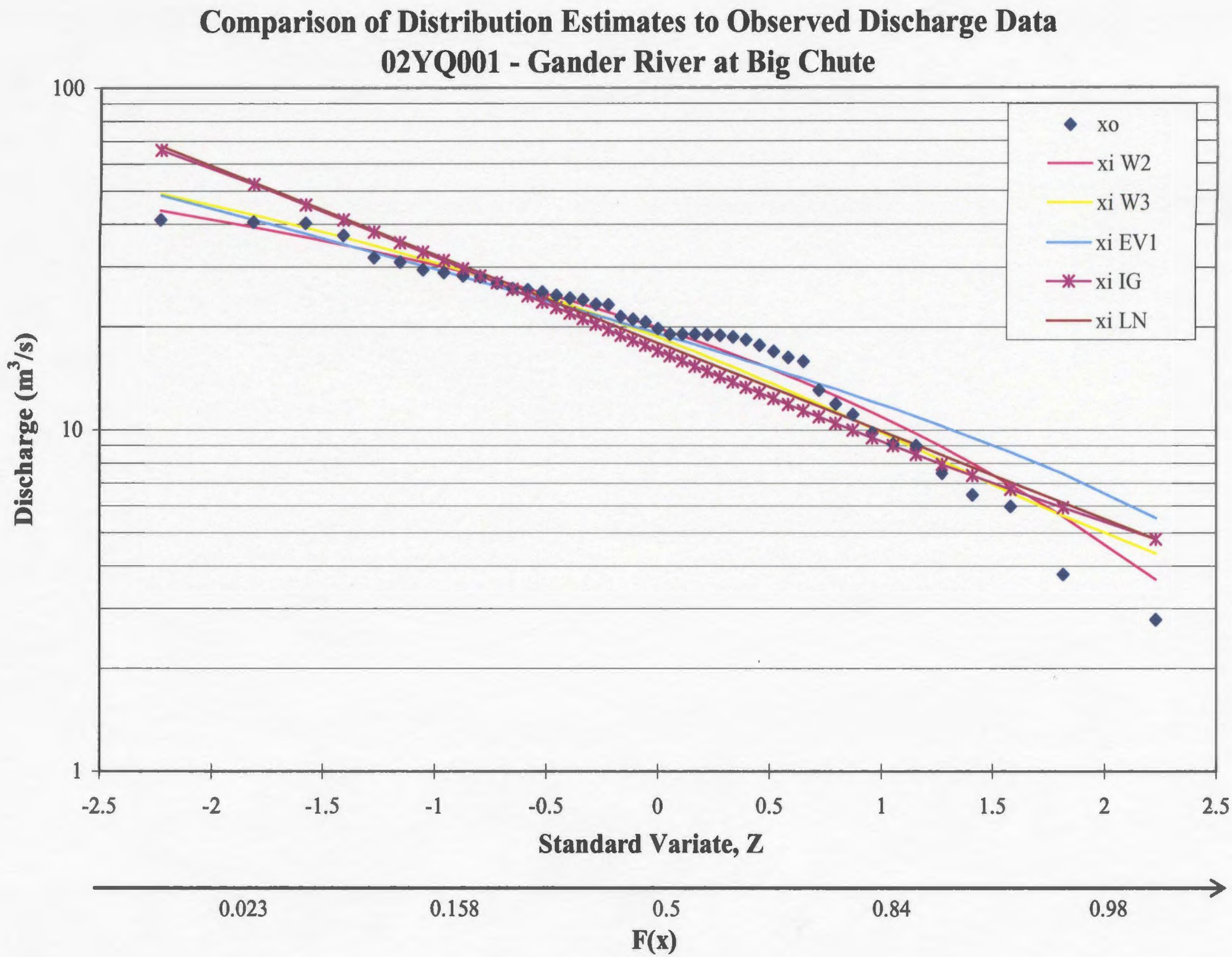
Comparison of Distribution Estimates to Observed Discharge Data
02YK005 - Sheffield Brook near Trans Canada Highway
Discharge (m^3/s)
10
1
0.1
xo
xi W2
xi W3
xi EV1
xi IG
xi LN
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)

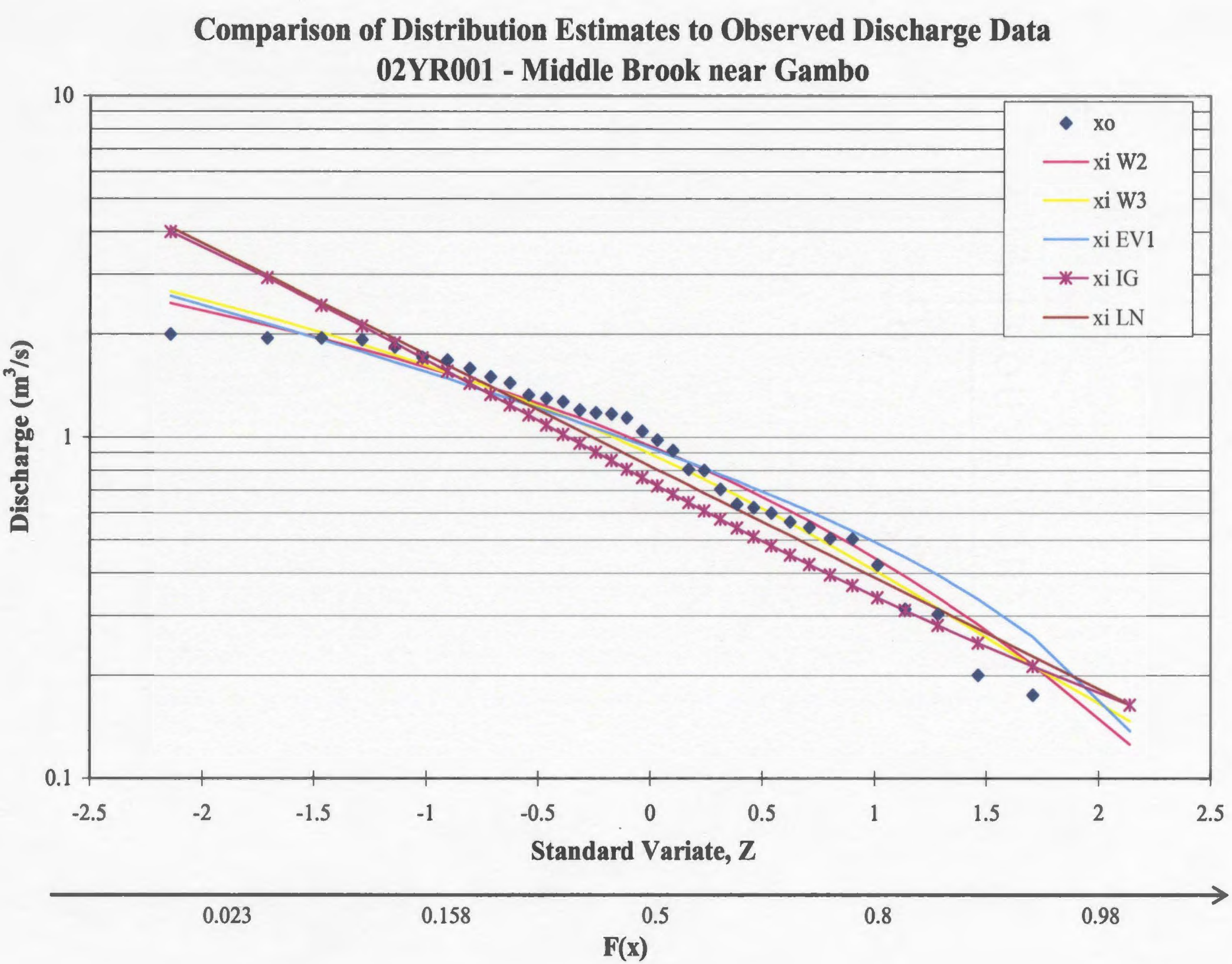

Comparison of Distribution Estimates to Observed Discharge Data
02YL001 - Upper Humber River near Reidville
Discharge (m³/s)
100
10
1
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.5
0.8
0.98
F(x)
xo
xi W2
xi W3
xi EV1
xi IG
xi LN


Comparison of Distribution Estimates to Observed Discharge Data
02YM003 - South West Brook near Baie Verte
Discharge (m³/s)
1
0.1
0.01
0.001
xo
xi W2
xi W3
xi EV1
xi IG
xi LN
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)

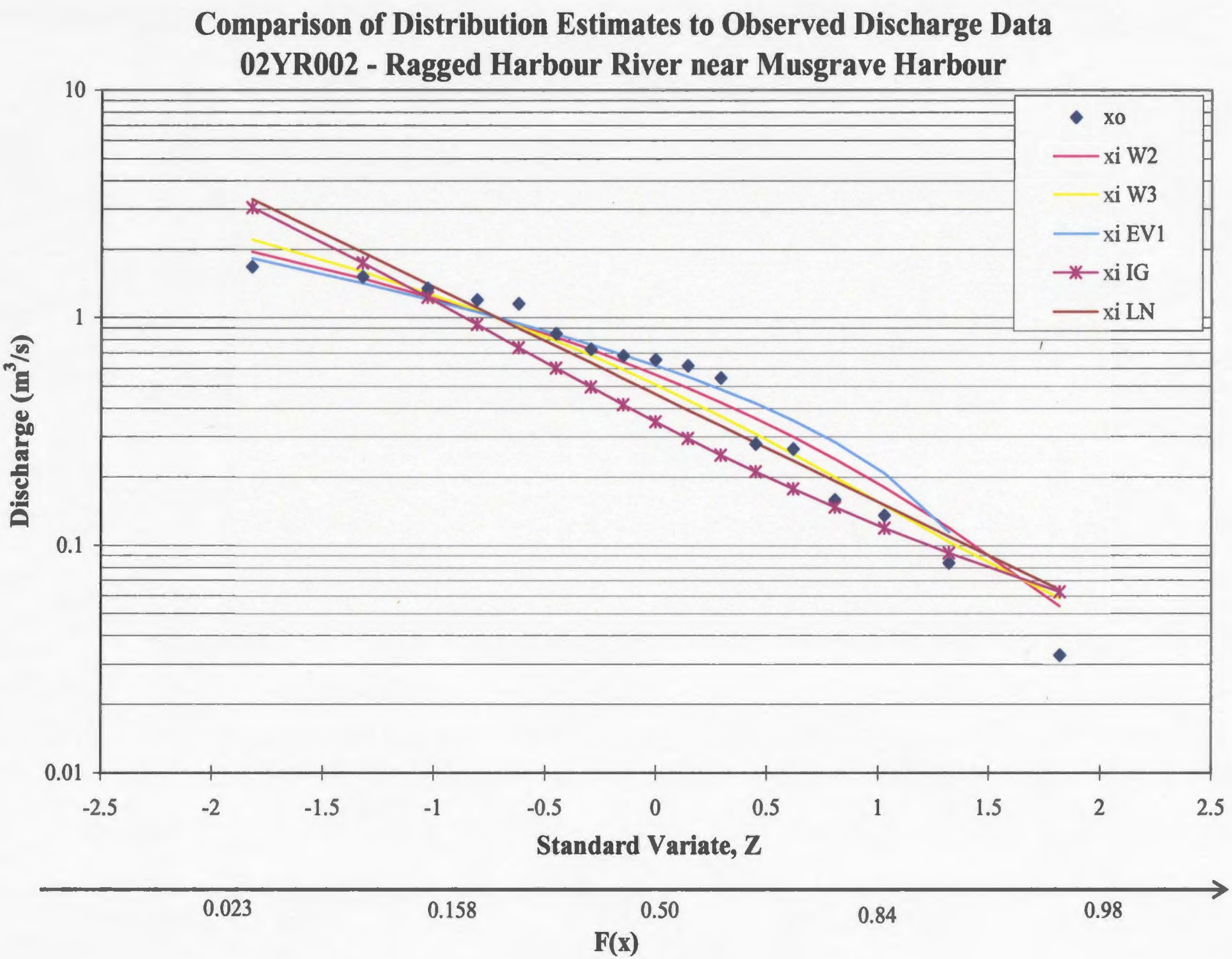
Comparison of Distribution Estimates to Observed Discharge Data
02YM003 - South West Brook near Baie Verte



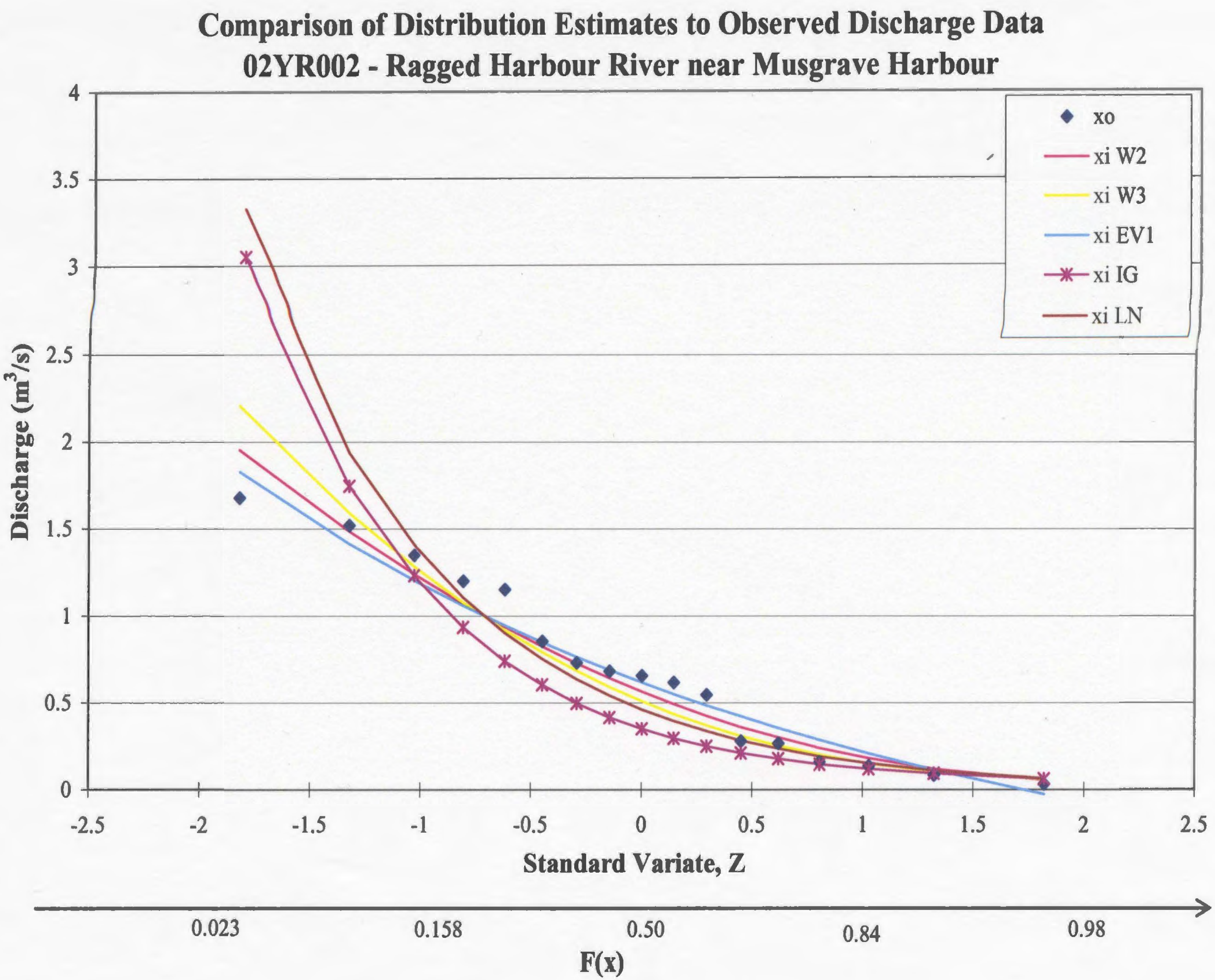

Comparison of Distribution Estimates to Observed Discharge Data
02YQ001 - Gander River at Big Chute
100
10
1
Discharge (m3/s)
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.5
0.84
0.98
F(x)
xo
xi W2
xi W3
xi EV1
xi IG
xi LN

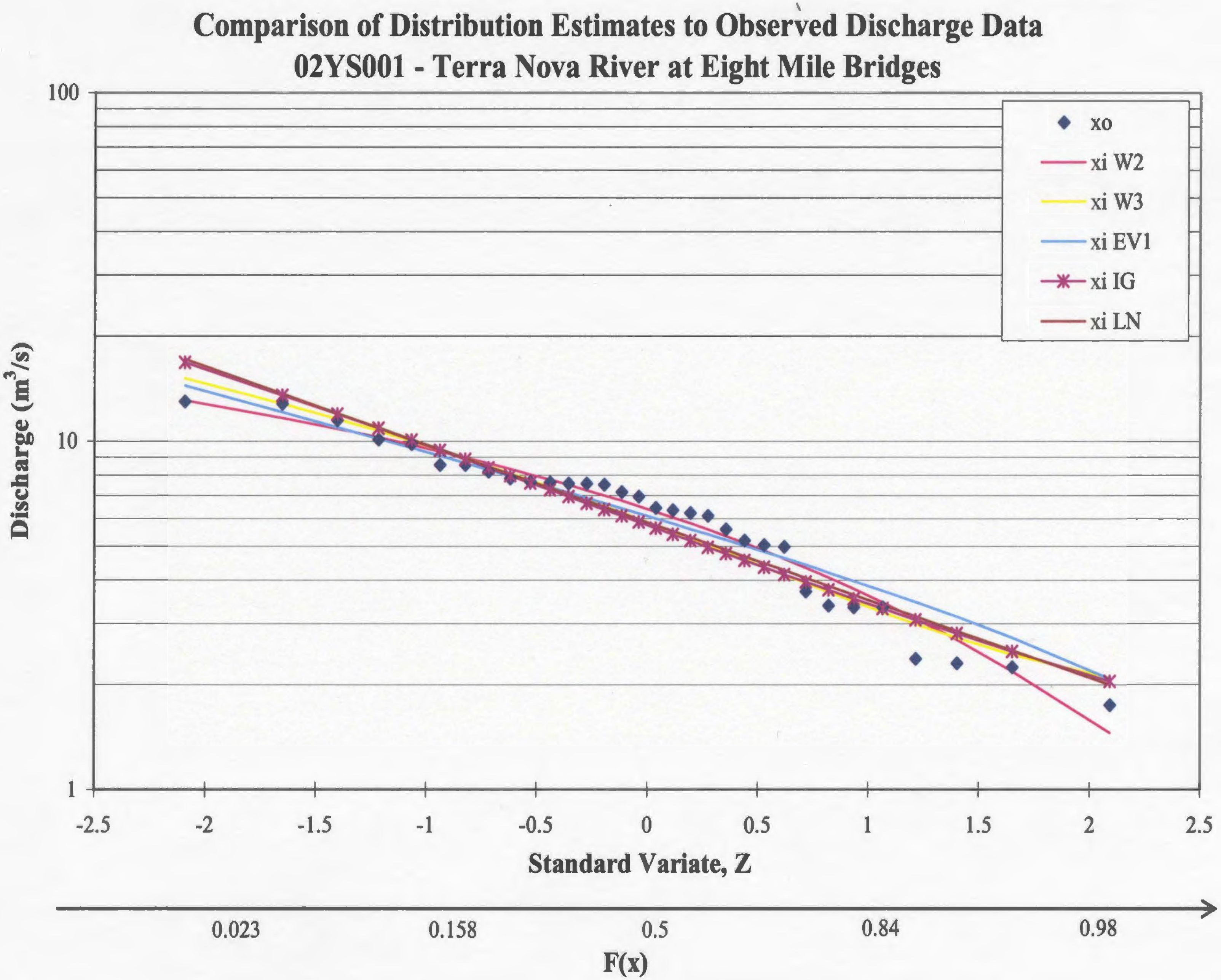


Comparison of Distribution Estimates to Observed Discharge Data
02YR001 - Middle Brook near Gambo
Discharge (m3/s)
10
1
0.1
xo
xi W2
xi W3
xi EV1
xi IG
xi LN
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.5
0.8
0.98
F(x)

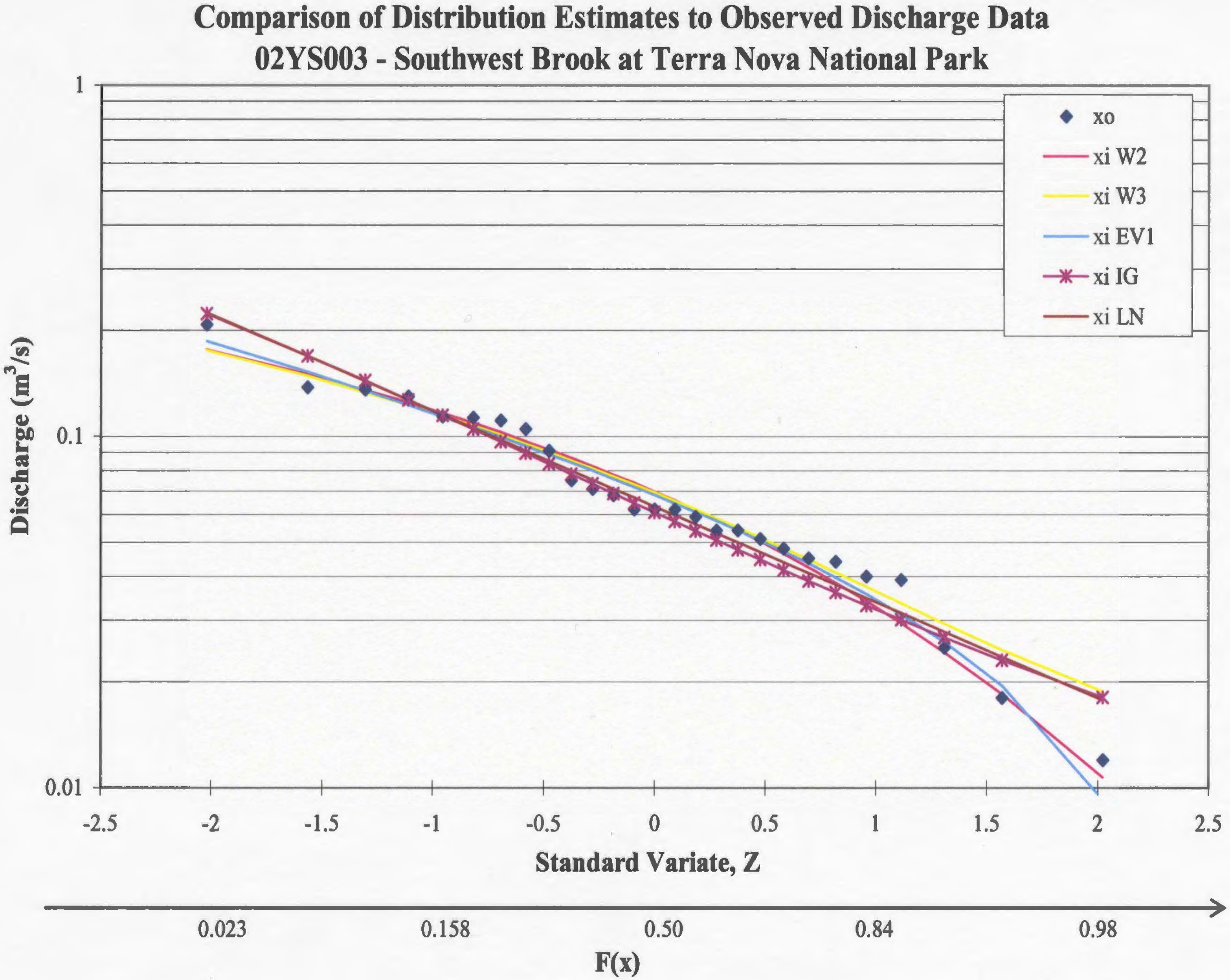


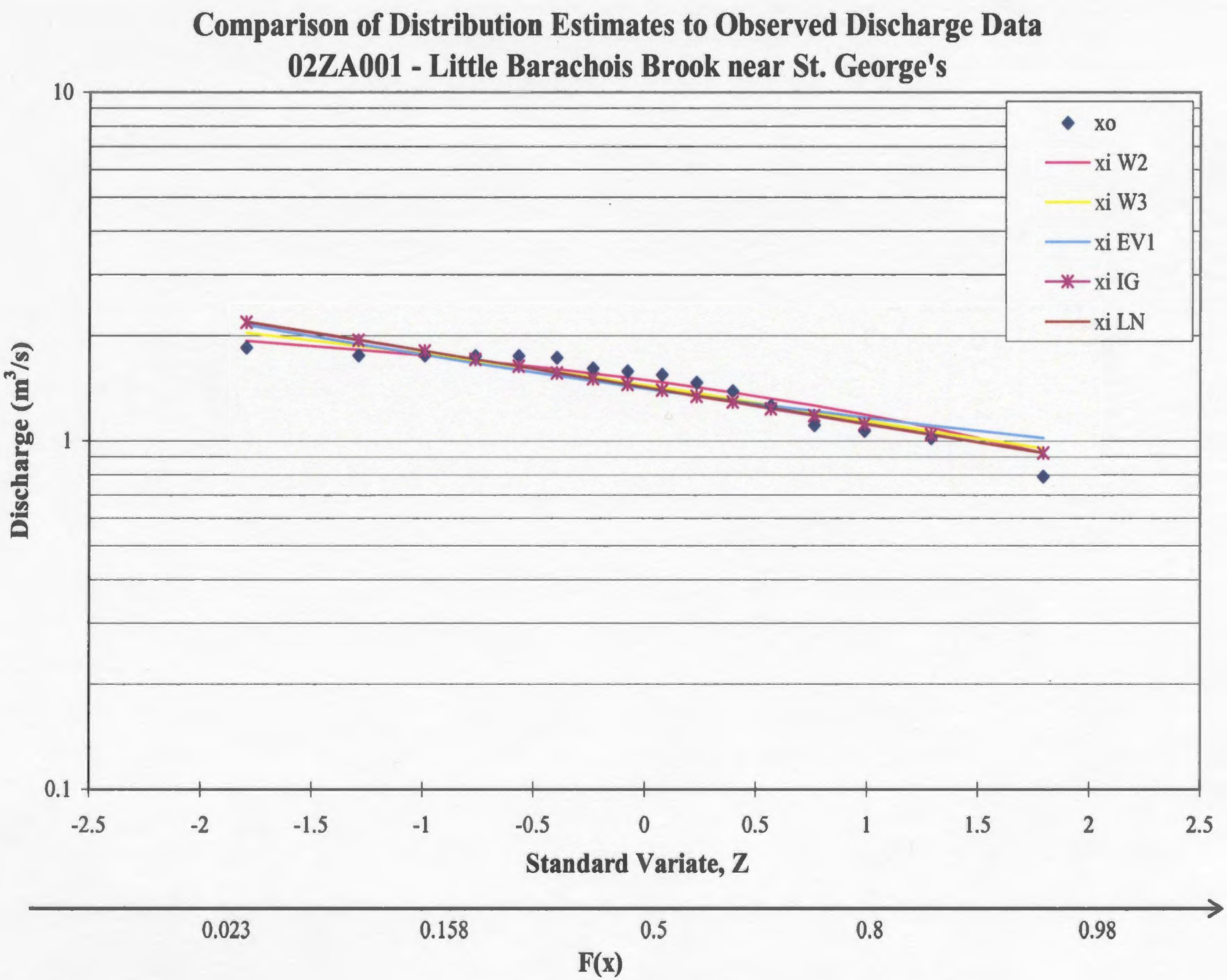
Comparison of Distribution Estimates to Observed Discharge Data
02YR002 - Ragged Harbour River near Musgrave Harbour
Discharge (m3/s)
10
1
0.1
0.01
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi W2
xi W3
xi EV1
xi IG
xi LN

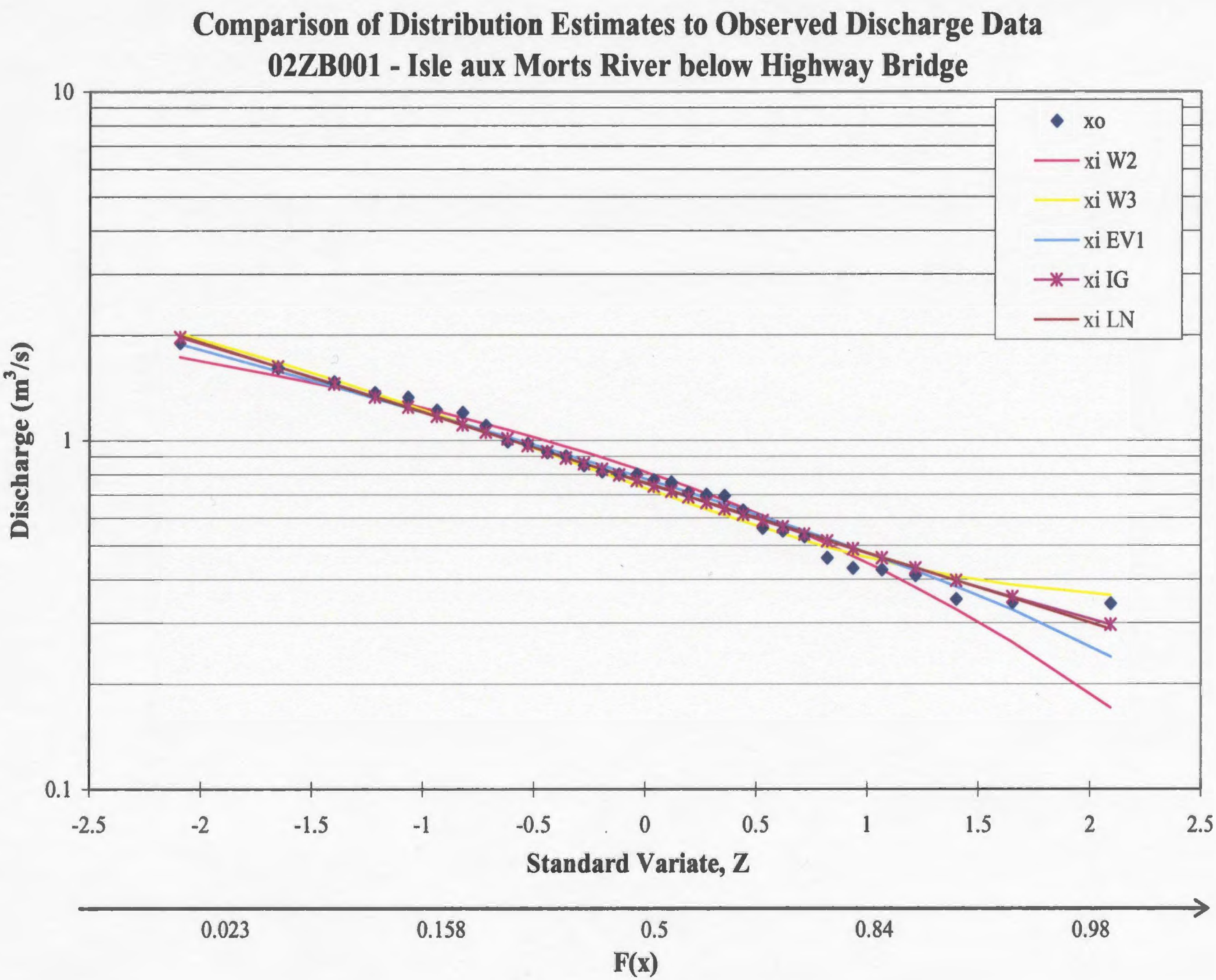

Comparison of Distribution Estimates to Observed Discharge Data
02YR002 - Ragged Harbour River near Musgrave Harbour
xo
xi W2
xi W3
xi EV1
xi IG
xi LN
Discharge (m³/s)
0
0.5
1
1.5
2
2.5
3
3.5
4
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)

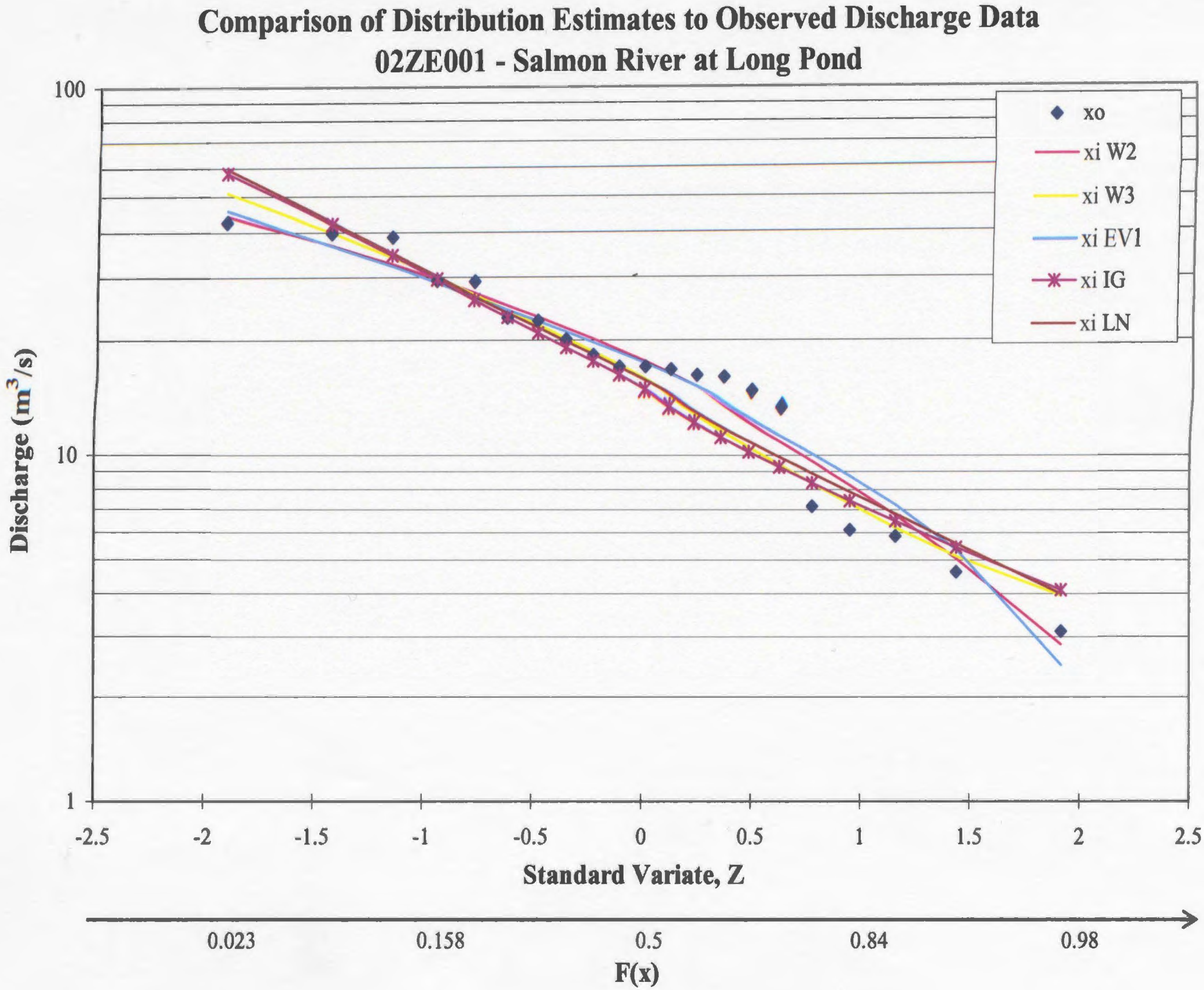


Comparison of Distribution Estimates to Observed Discharge Data
02YS001 - Terra Nova River at Eight Mile Bridges
Discharge (m3/s)
100
10
1
xo
xi W2
xi W3
xi EV1
xi IG
xi LN
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.5
0.84
0.98
F(x)

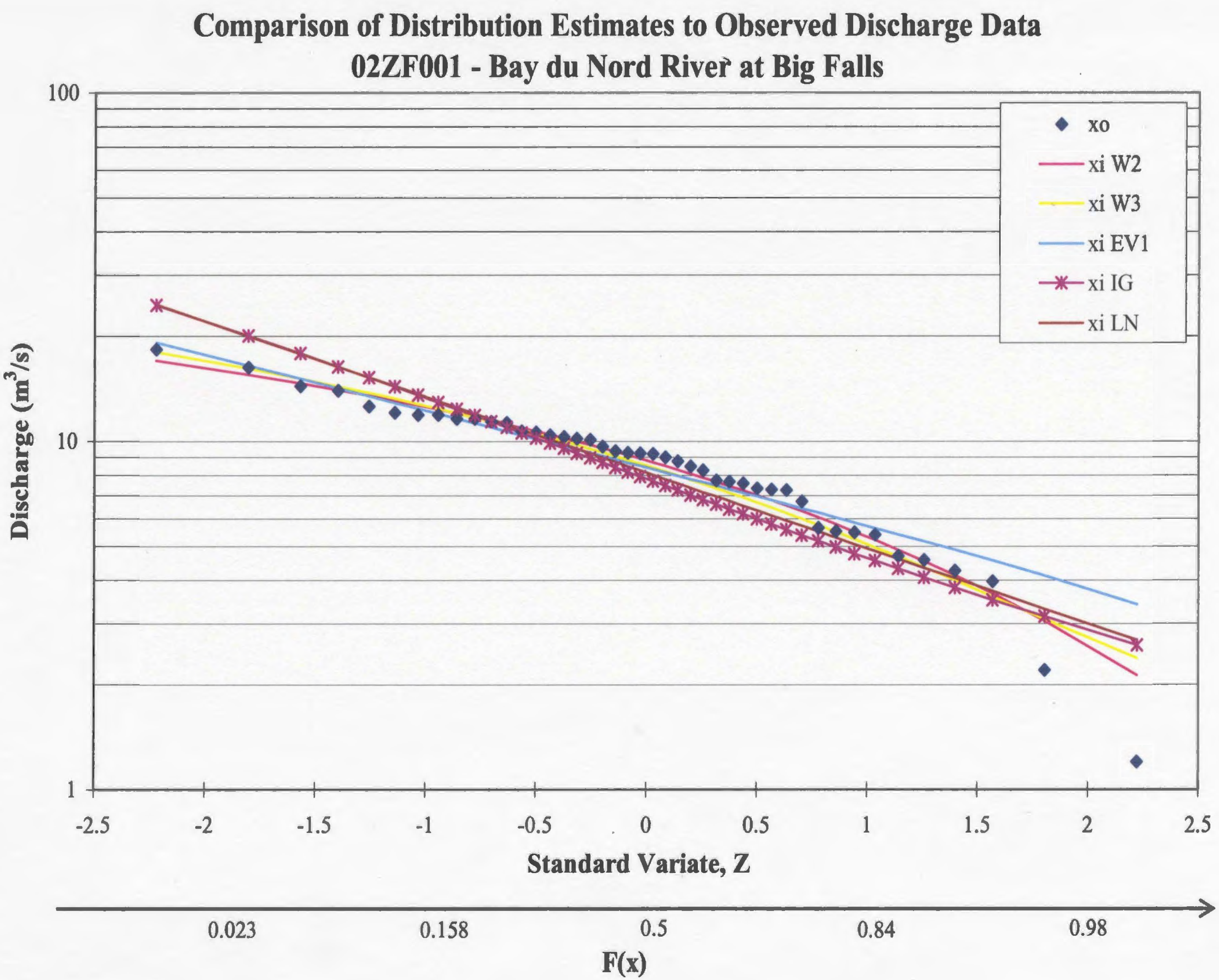

Comparison of Distribution Estimates to Observed Discharge Data
02YS003 - Southwest Brook at Terra Nova National Park
Discharge (m^3/s)
1
0.1
0.01
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi W2
xi W3
xi EV1
xi IG
xi LN


Comparison of Distribution Estimates to Observed Discharge Data
02ZA001 - Little Barachois Brook near St. George's
10
1
0.1
Discharge (m³/s)
xo
xi W2
xi W3
xi EV1
xi IG
xi LN
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.5
0.8
0.98
F(x)


Comparison of Distribution Estimates to Observed Discharge Data
02ZB001 - Isle aux Morts River below Highway Bridge
Discharge (m3/s)
10
1
0.1
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.5
0.84
0.98
F(x)
xo
xi W2
xi W3
xi EV1
xi IG
xi LN

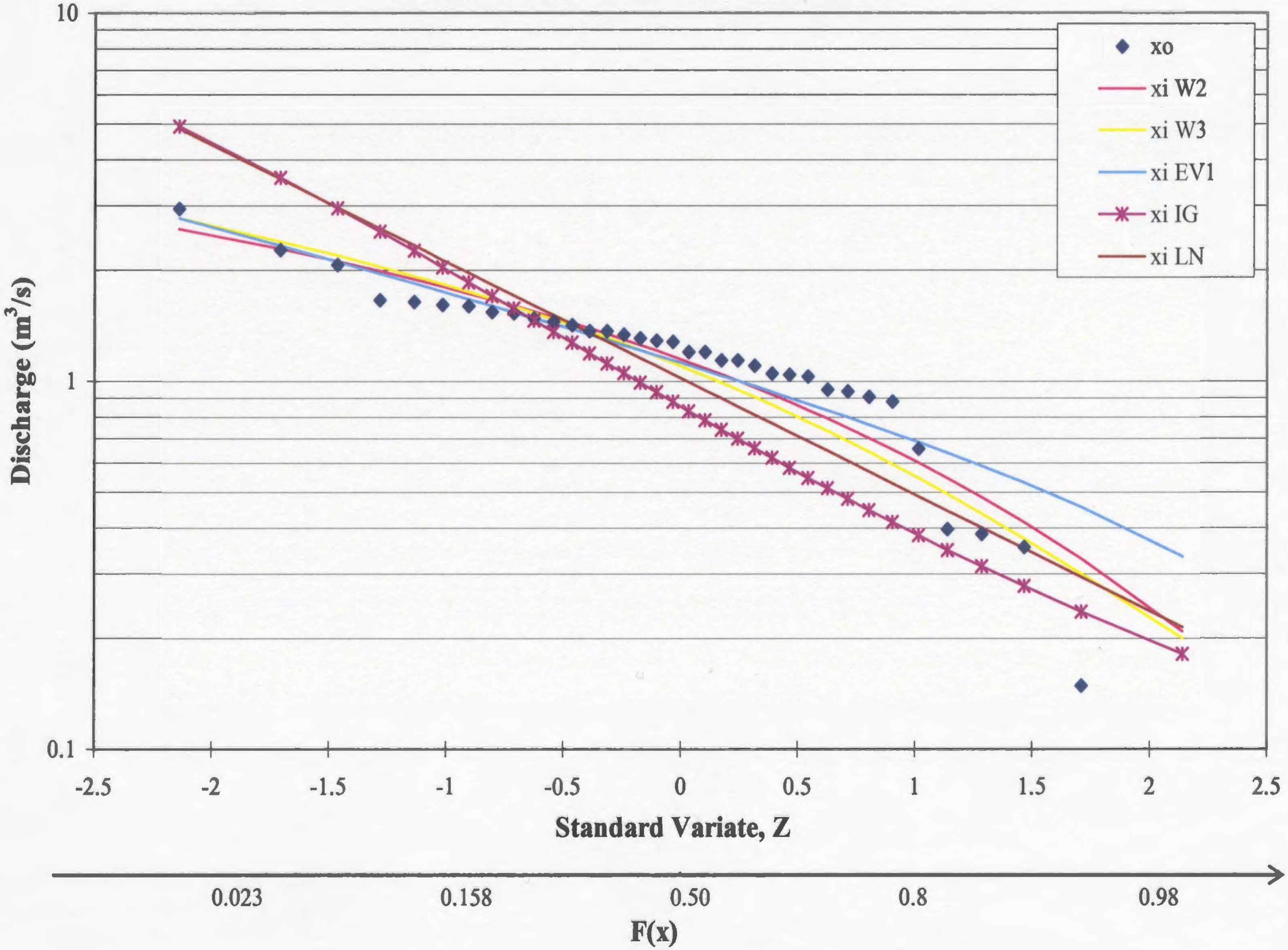


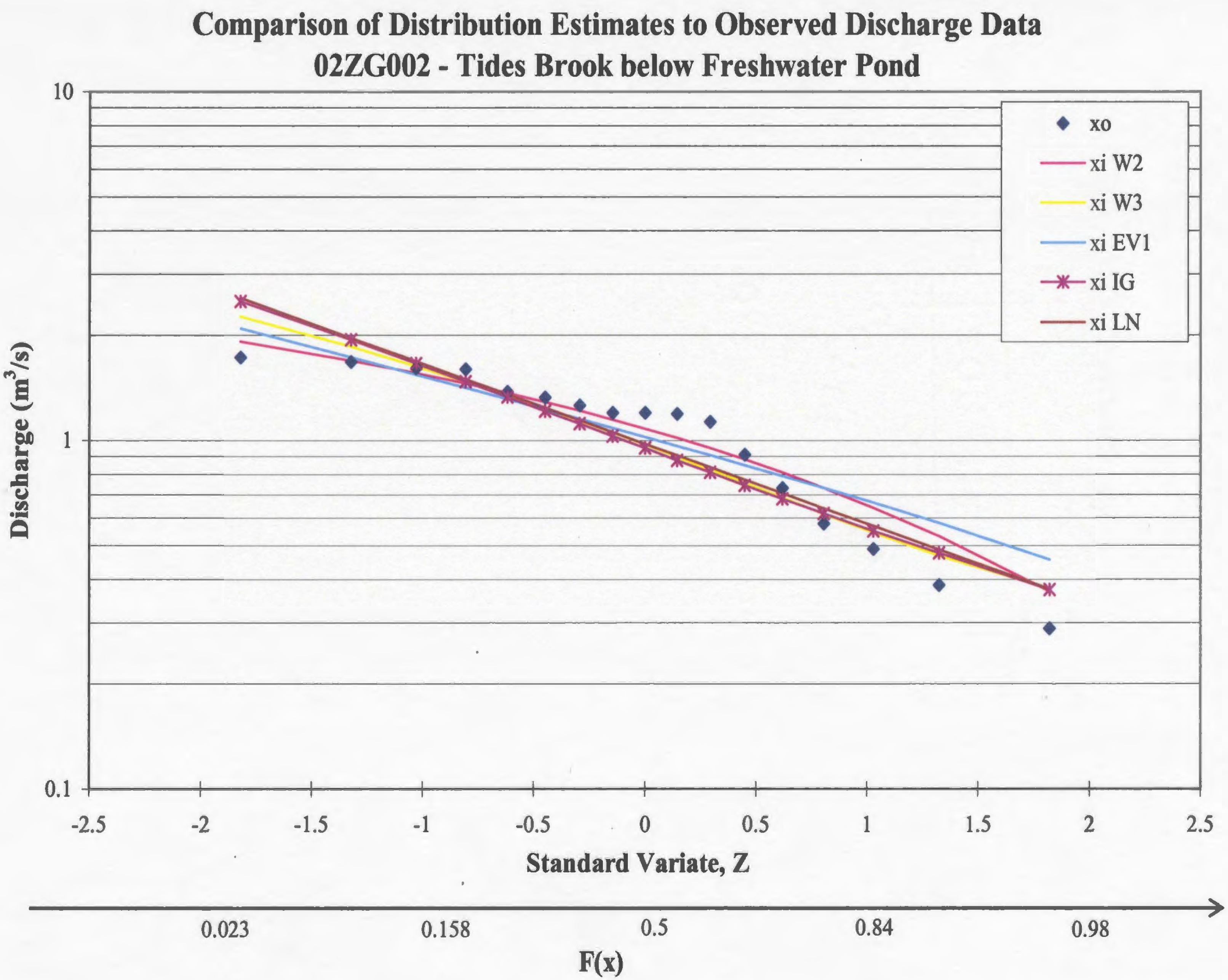
Comparison of Distribution Estimates to Observed Discharge Data
02ZE001 - Salmon River at Long Pond
Discharge (m³/s)
100
10
1
xo
xi W2
xi W3
xi EV1
xi IG
xi LN
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.5
0.84
0.98
F(x)



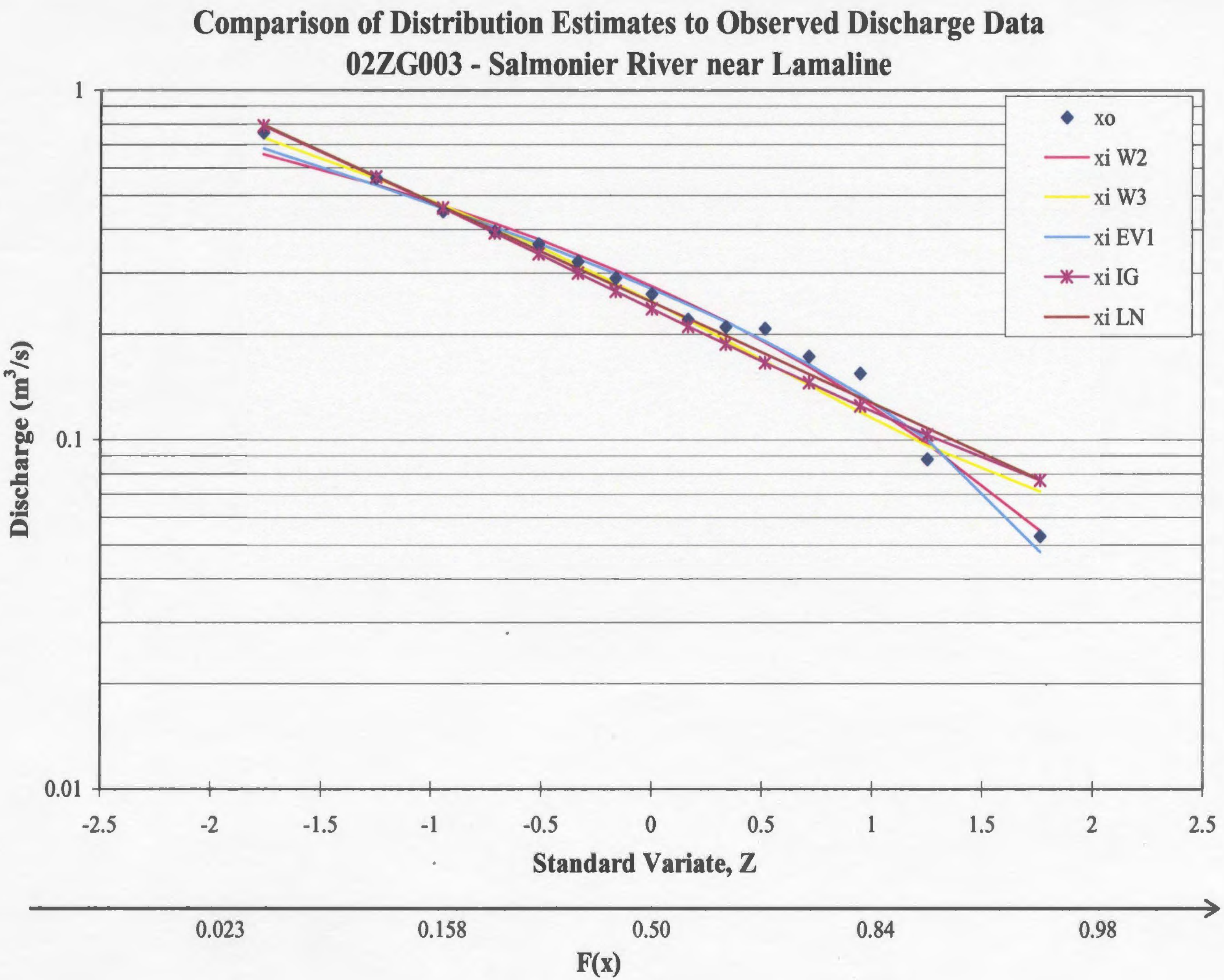
Comparison of Distribution Estimates to Observed Discharge Data
02ZF001 - Bay du Nord River at Big Falls
Discharge (m3/s)
100
10
1
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.5
0.84
0.98
F(x)
xo
xi W2
xi W3
xi EV1
xi IG
xi LN

Comparison of Distribution Estimates to Observed Discharge Data
02ZG001 - Garnish River near Garnish

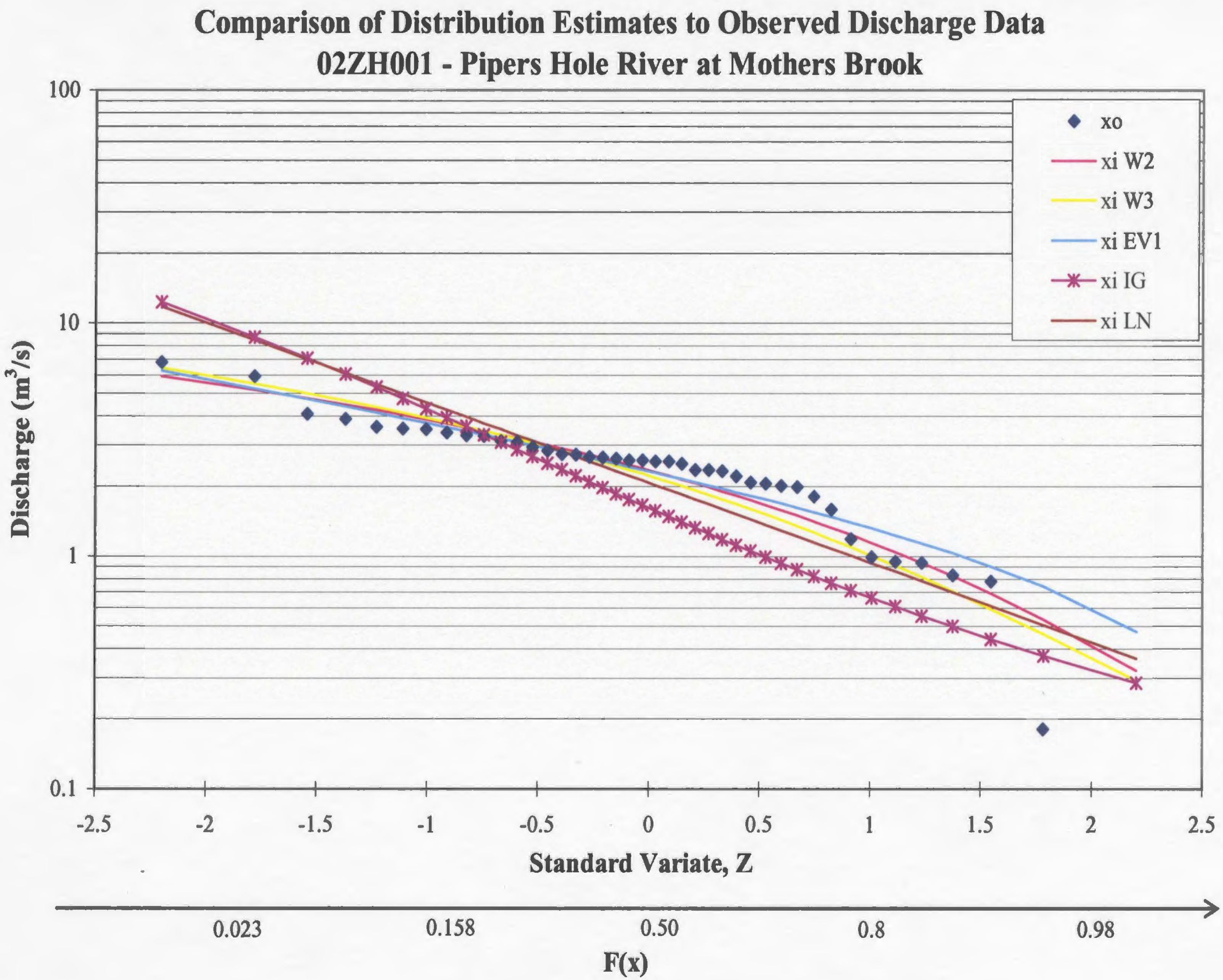


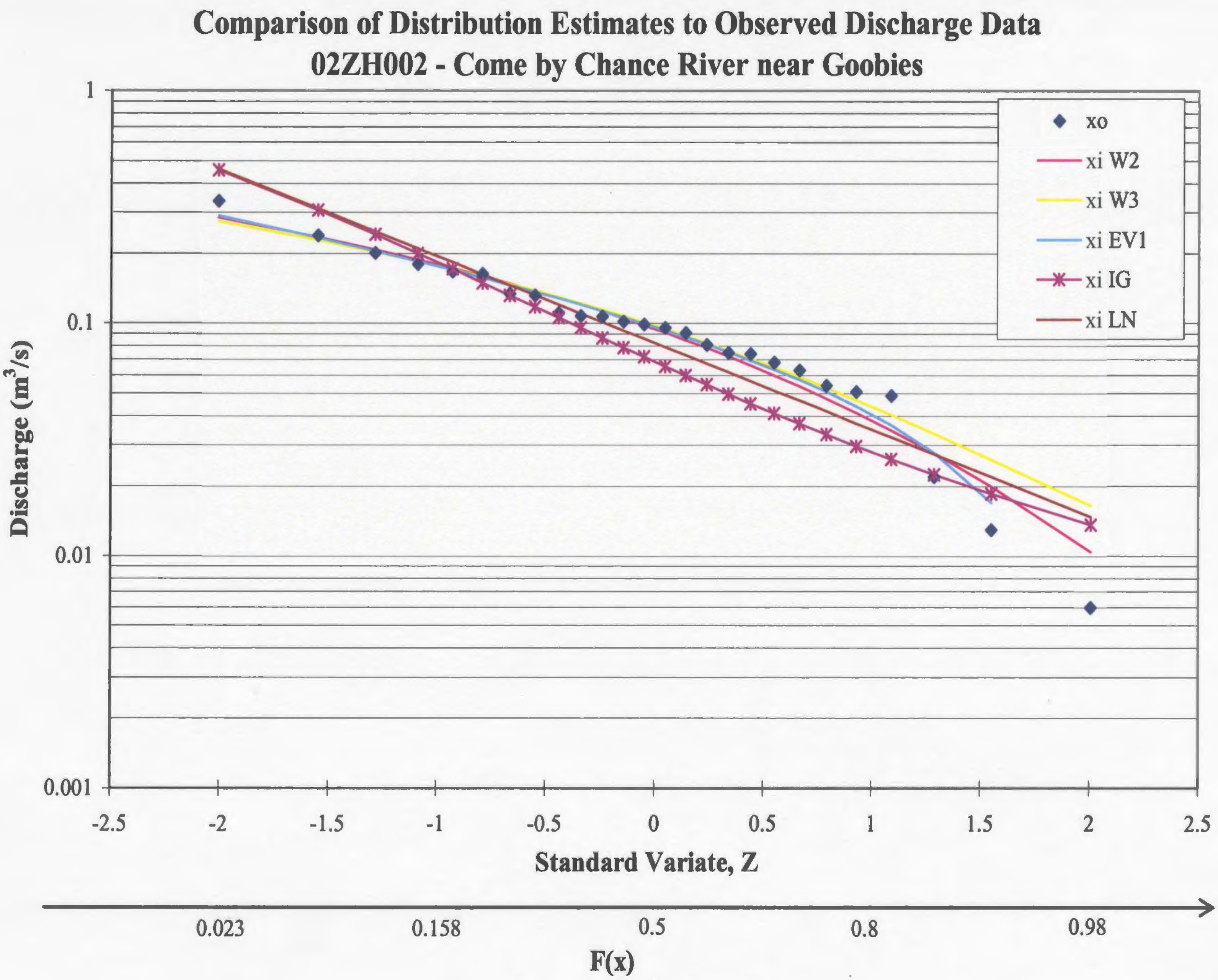


Comparison of Distribution Estimates to Observed Discharge Data
02ZG002 - Tides Brook below Freshwater Pond
Discharge (m³/s)
10
1
0.1
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.5
0.84
0.98
F(x)
xo
xi W2
xi W3
xi EV1
xi IG
xi LN

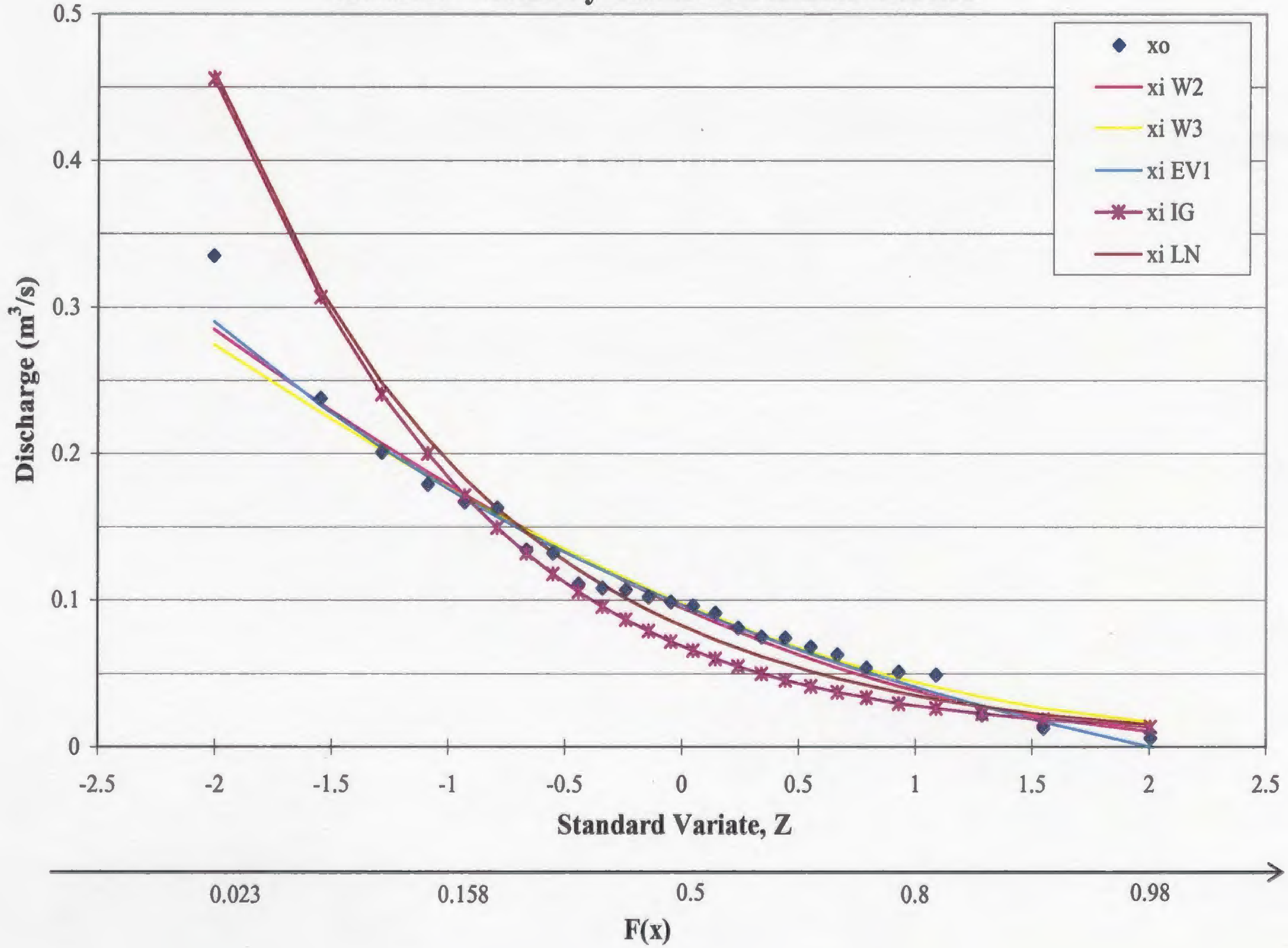


Comparison of Distribution Estimates to Observed Discharge Data
02ZG003 - Salmonier River near Lamaline
Discharge (m³/s)
1
0.1
0.01
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi W2
xi W3
xi EV1
xi IG
xi LN

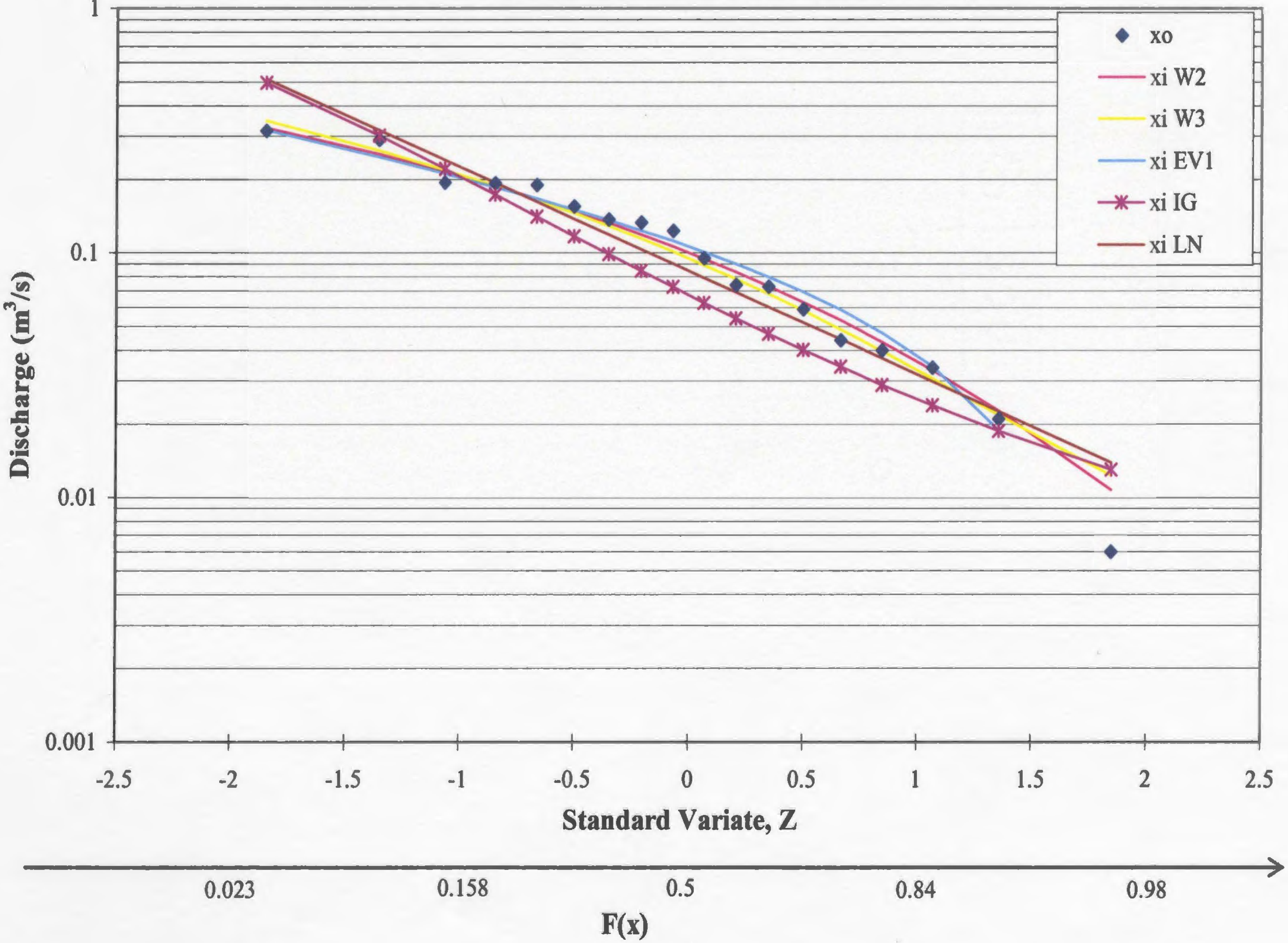

Comparison of Distribution Estimates to Observed Discharge Data
02ZH001 - Pipers Hole River at Mothers Brook
Discharge (m^3/s)
100
10
1
0.1
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.8
0.98
F(x)
xo
xi W2
xi W3
xi EV1
xi IG
xi LN

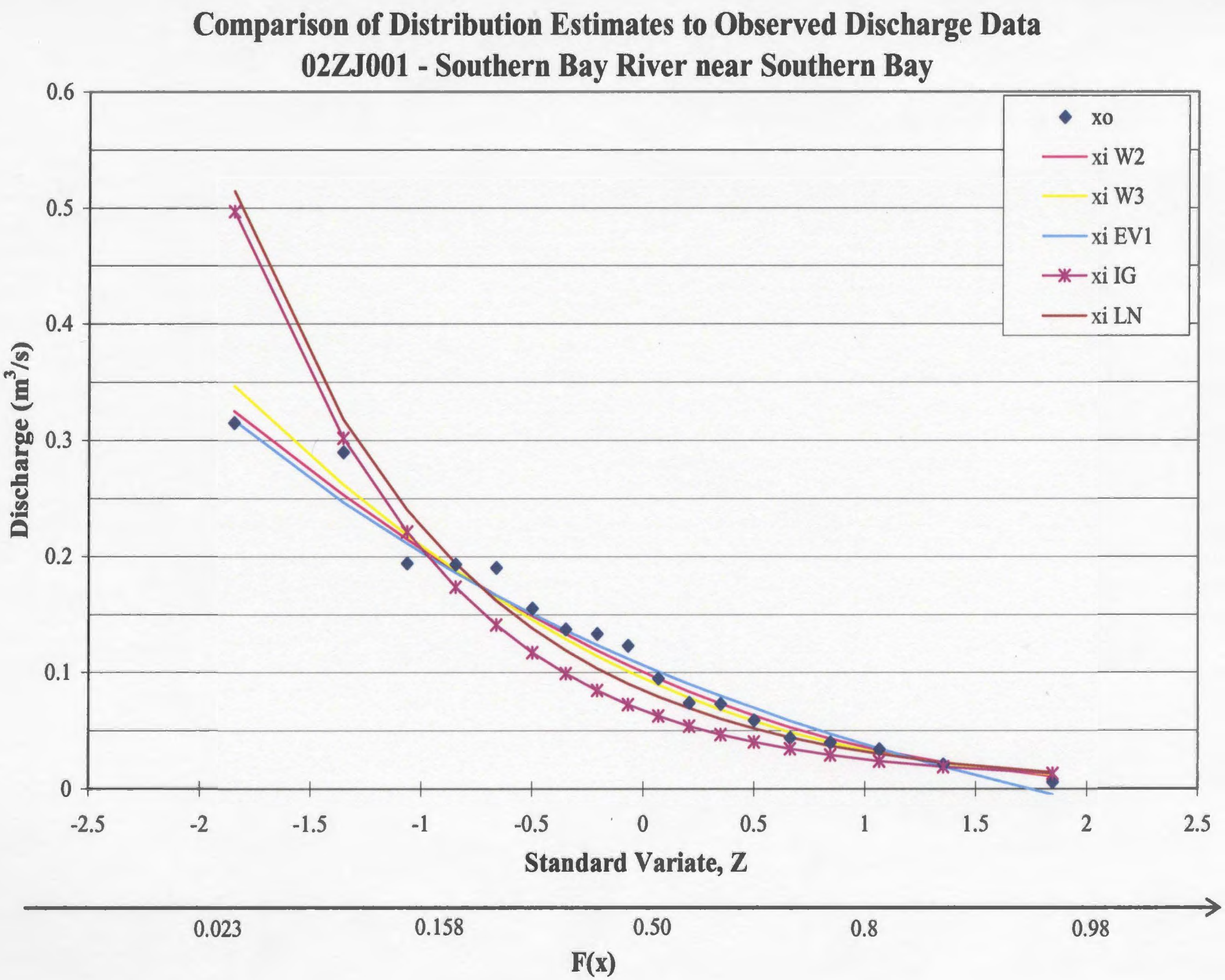

Comparison of Distribution Estimates to Observed Discharge Data
02ZH002 - Come by Chance River near Goobies
xo
xi W2
xi W3
xi EV1
xi IG
xi LN
Discharge (m³/s)
1
0.1
0.01
0.001
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.5
0.8
0.98
F(x)

Comparison of Distribution Estimates to Observed Discharge Data
02ZH002 - Come by Chance River near Goobies

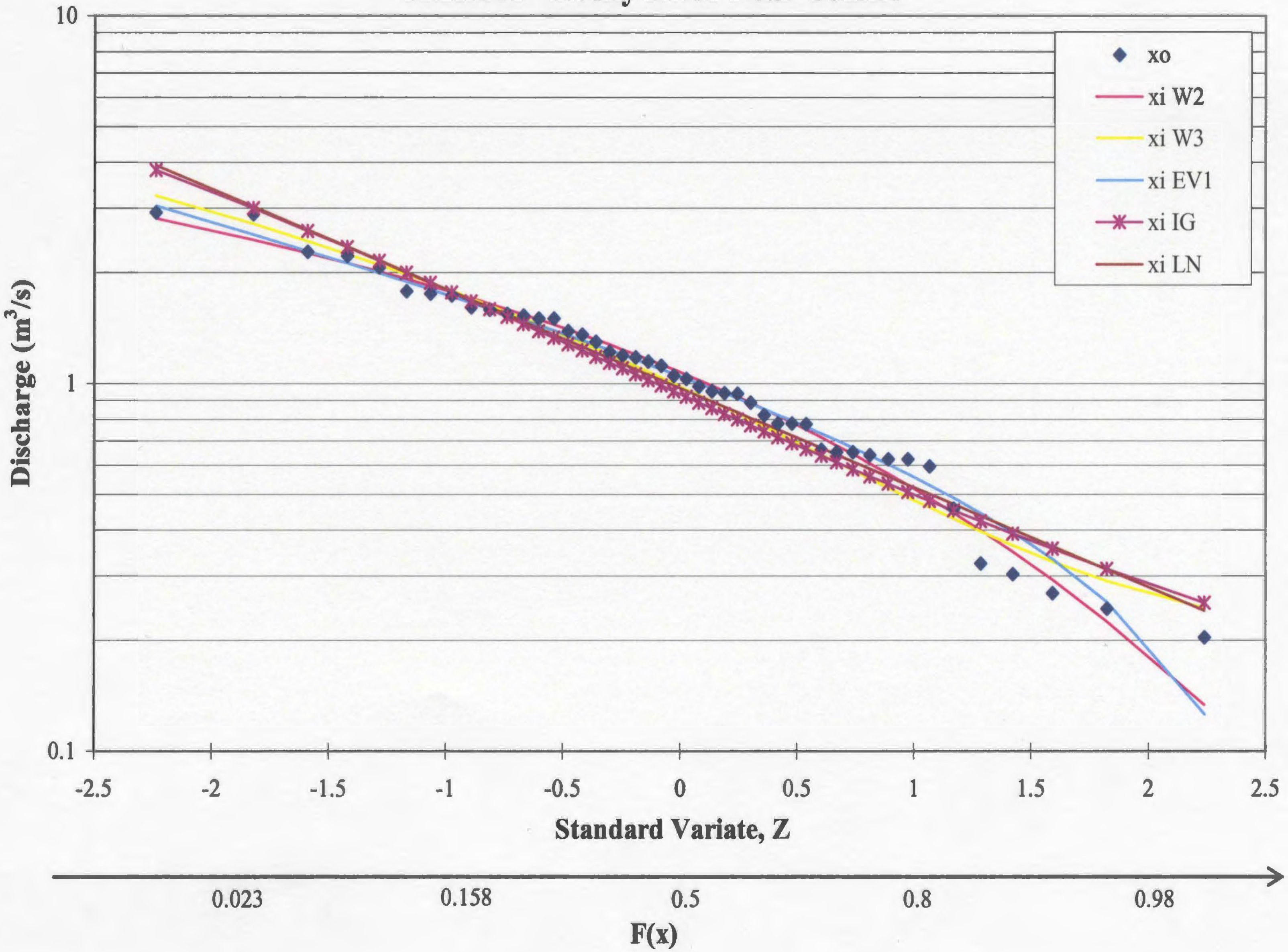


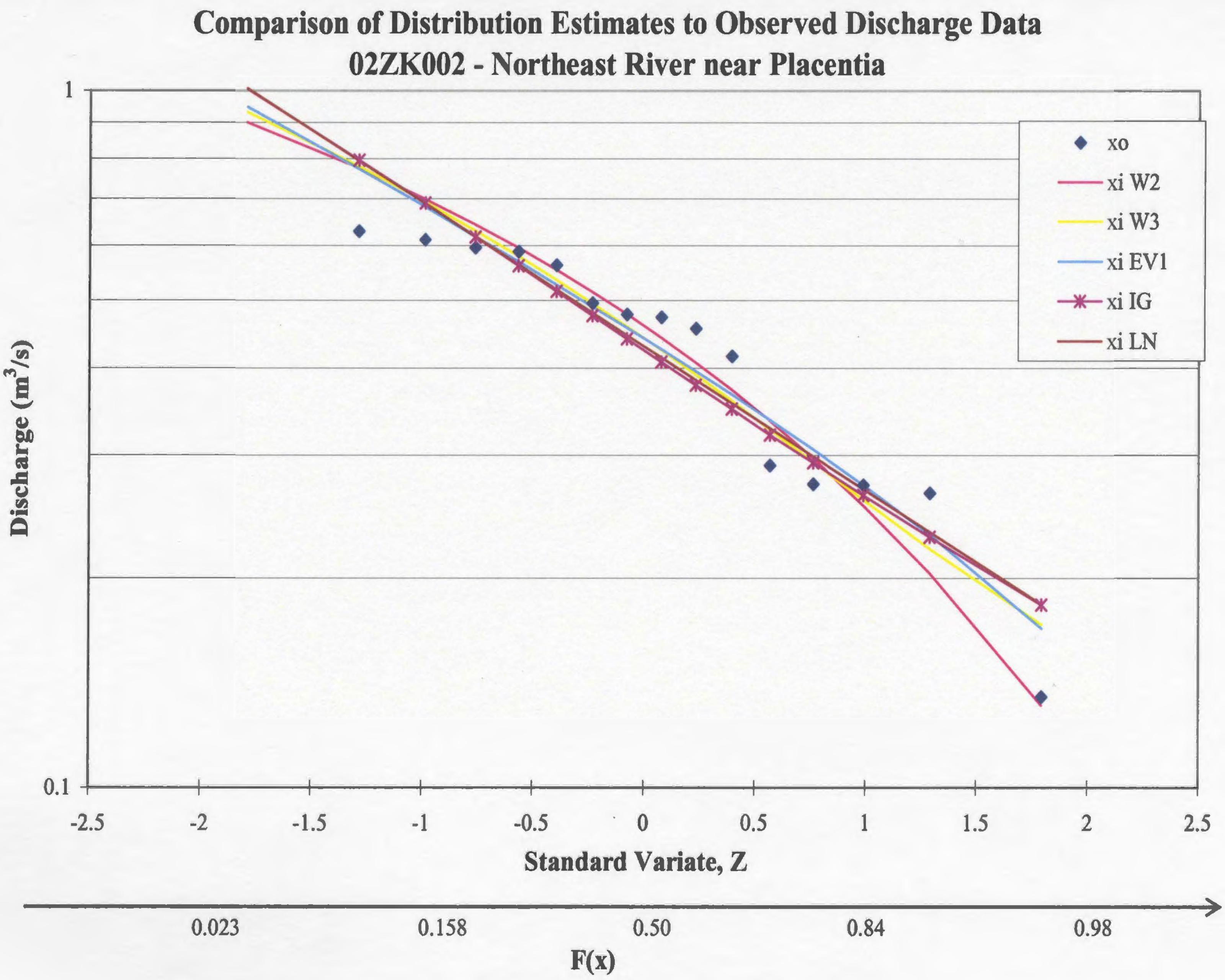
Comparison of Distribution Estimates to Observed Discharge Data
02ZJ001 - Southern Bay River near Southern Bay



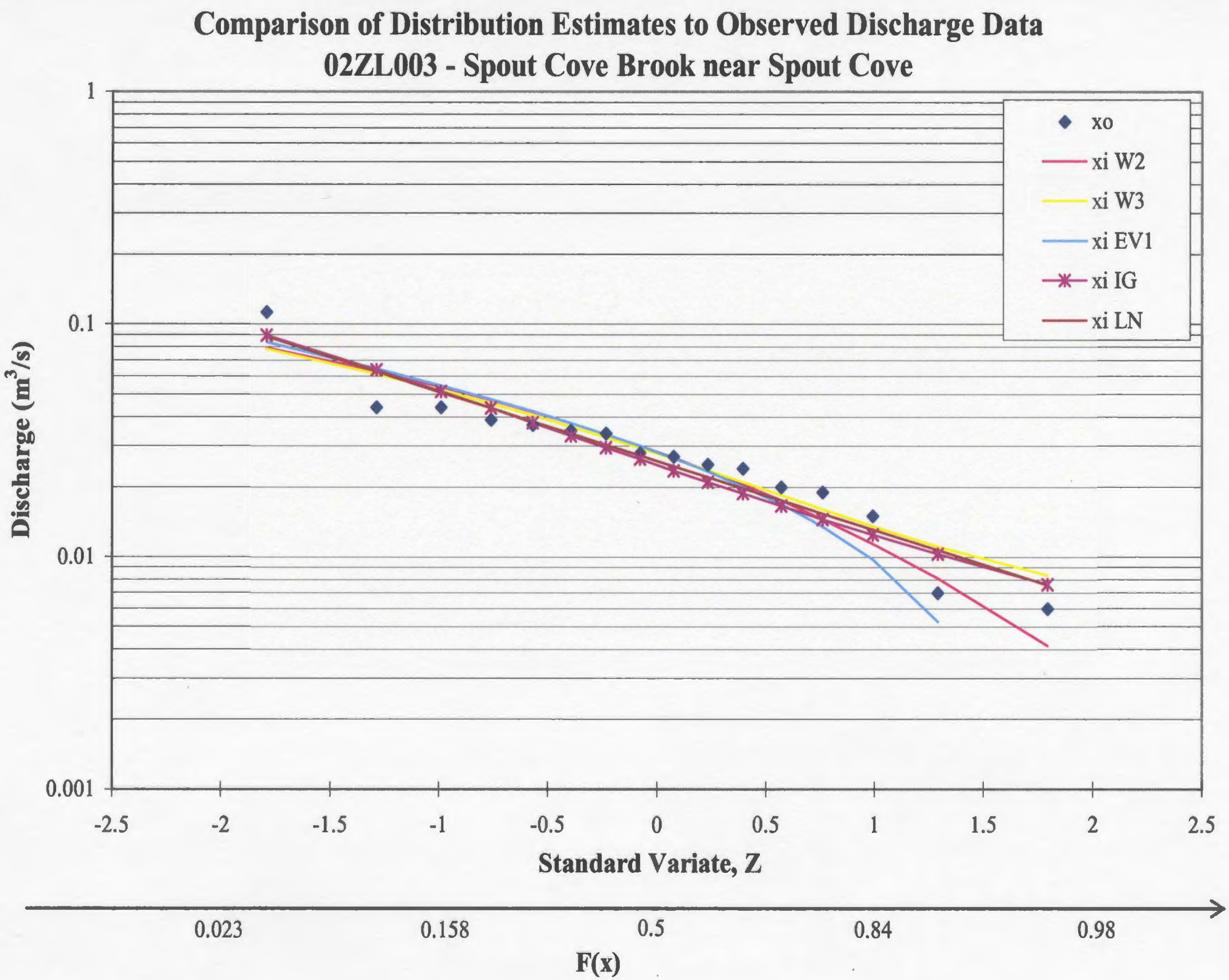

Comparison of Distribution Estimates to Observed Discharge Data
02ZJ001 - Southern Bay River near Southern Bay
Discharge (m³/s)
0.6
0.5
0.4
0.3
0.2
0.1
0
xo
xi W2
xi W3
xi EV1
xi IG
xi LN
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.8
0.98
F(x)

Comparison of Distribution Estimatesn to Observed Discharge Data
02ZK001 - Rocky River Near Colinet

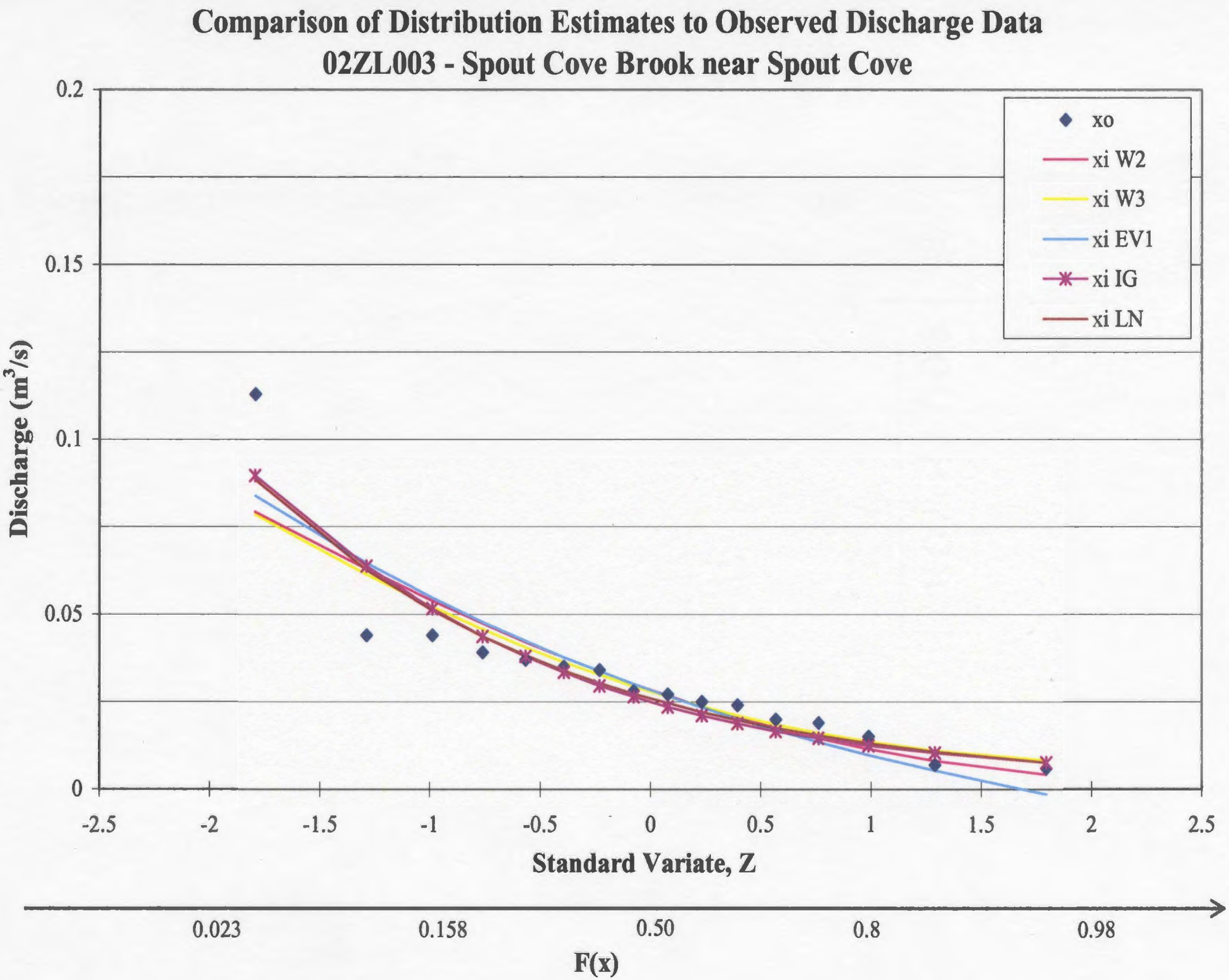




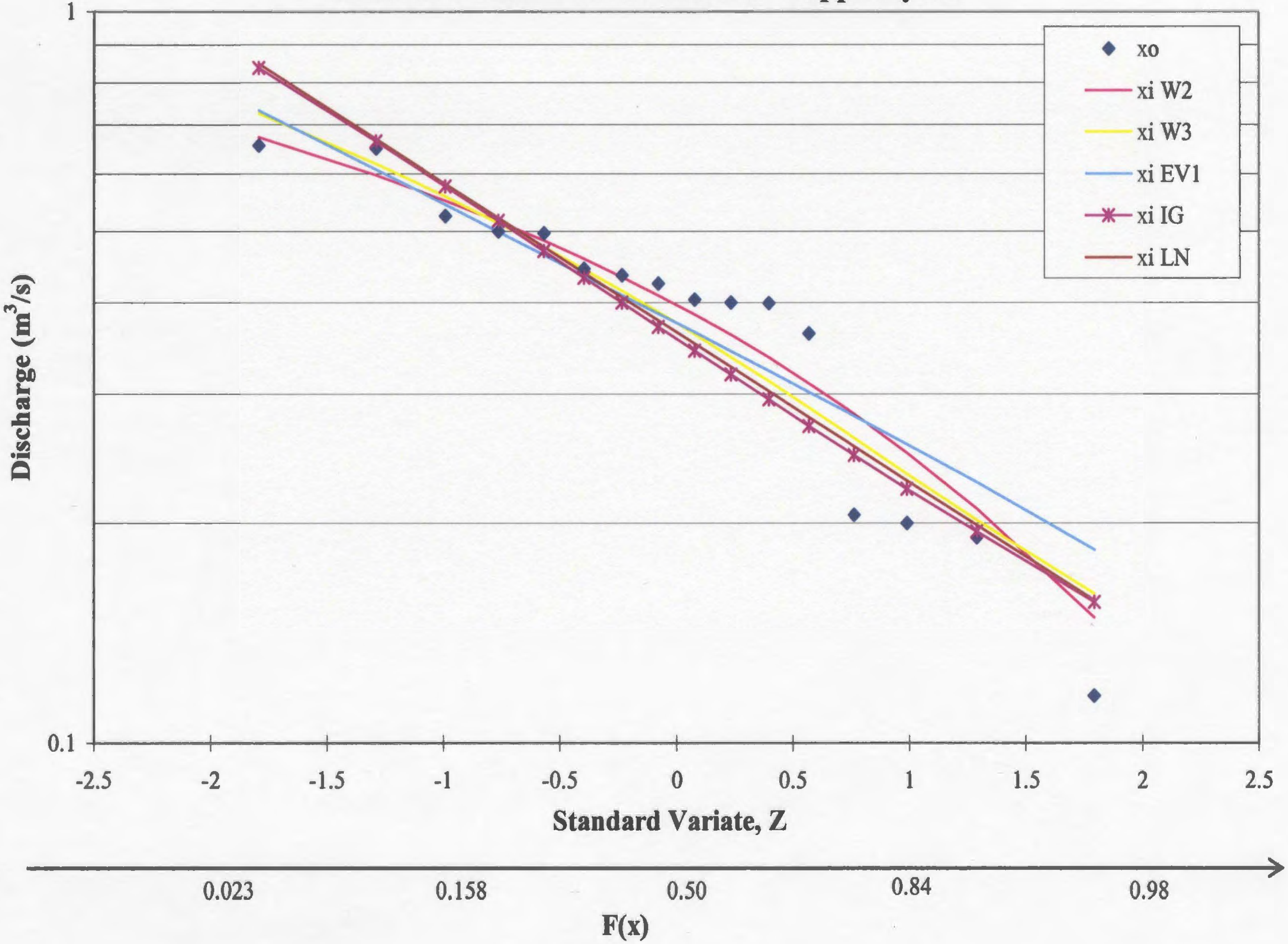
Comparison of Distribution Estimates to Observed Discharge Data
02ZK002 - Northeast River near Placentia
Discharge (m3/s)
1
0.1
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.84
0.98
F(x)
xo
xi W2
xi W3
xi EV1
xi IG
xi LN



Comparison of Distribution Estimates to Observed Discharge Data
02ZL003 - Spout Cove Brook near Spout Cove
Discharge (m3/s)
1
0.1
0.01
0.001
xo
xi W2
xi W3
xi EV1
xi IG
xi LN
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.5
0.84
0.98
F(x)


Comparison of Distribution Estimates to Observed Discharge Data
02ZL003 - Spout Cove Brook near Spout Cove
Discharge (m3/s)
0.2
0.15
0.1
0.05
0
xo
xi W2
xi W3
xi EV1
xi IG
xi LN
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
Standard Variate, Z
0.023
0.158
0.50
0.8
0.98
F(x)

Comparison of Distribution Estimates to Observed Discharge Data
02ZM009 - Seal Cove Brook near Cappahayden



Comparison of Distribution Estimates to Observed Discharge Data
02ZN001 - Northwest Brook at Northwest Pond