2006

Statistical models in environmental and life sciences

Lakshminarayan Rajaram
University of South Florida

Statistical Models in Environmental and Life Sciences

by

Lakshminarayan Rajaram

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Mathematics
College of Arts and Sciences
University of South Florida

Major Professor: Chris P. Tsokos, Ph.D.
Marcus Mcwaters, Ph.D.
Kandethody Ramachandran, Ph.D.
Geoffrey Okogbaa, Ph.D.

Date of Approval:
April 6, 2006

Keywords: trivariate normal, simulation, covariance structure, bioavailability, extreme value distribution

Dedication

I dedicate this dissertation to my wife, Ha, to my son, Ganesh, and to my brother, Swamy. Especially to Swamy, who spent his entire life nurturing all of us, our parents and siblings, with his unflinching devotion and love. His loving words of encouragement from the depths of his magnanimous heart and that ever-present radiant smile have solely contributed to my commitment to complete the doctoral degree.

Acknowledgments

I take this opportunity to acknowledge a number of people who have helped me to succeed in accomplishing my goal. Firstly, I would like to express my profound appreciation to Dr. Chris P. Tsokos for suggesting the research problems, for his continual support and encouragement all through my graduate studies. I also want to thank the other members of my committee-Dr. Marcus McWaters, Dr. K. Ramachandran, and Dr. Okogbaa- for their guidance and support. Finally, I like to express a special note of thanks to Dr. A.N.V. Rao for all his encouragement.

Table of Contents

List of Tables

List of Figures

Statistical Models in Environmental and Life Sciences

Lakshminarayan Rajaram

ABSTRACT

The dissertation focuses on developing statistical models in environmental and life sciences.

The Generalized Extreme Value distribution is used to model annual monthly maximum rainfall data from 44 locations in Florida. Time dependence of the rainfall data is incorporated into the model by assuming the location parameter to be a function of time, both linear and quadratic. Estimates and confidence intervals are obtained for return levels of return periods of 10, 20, 50, and 100 years. Locations are grouped into statistical profiles based on their similarities in return level graphs for all locations, and locations within each climatic zone.

A family of extreme values distributions is applied to model simulated maximum drug concentration (C_{max}) data of an anticoagulant drug. For small samples ($n \leq 100$) data exhibited bimodality. The results of investigating a mixture of two extreme value distributions to model such bimodal data using two-parameter Gumbel, Pareto and Weibull concluded that a mixture of two Weibull distributions is the only suitable model.

For large samples $(n > 100)$, C_{max} data are modeled using the Generalized Extreme Value, Gumbel, Weibull, and Pareto distributions. These results concluded that the Generalized Extreme Value distribution is the only suitable model.

A system of random differential equations is used to investigate the drug concentration behavior in a three-compartment pharmacokinetic model which describes coumermycin's disposition. The rate constants used in the differential equations are assumed to have a trivariate distribution, and hence, simulated from the trivariate truncated normal probability distribution.

Numerical solutions are developed under different combinations of the covariance structure and the nonrandom initial conditions. We study the dependence effect that such a pharmacokinetic system has among the three compartments as well as the effect of variance in identifying the concentration behavior in each compartment. We identify the time delays in each compartment.

We extend these models to incorporate the identified time delays. We provide the graphical display of the time delay effects on the drug concentration behavior as well as the comparison of the deterministic behavior with and without the time delay, and effect of different sets of time delay on deterministic and stochastic behaviors.

Chapter One

Review of Literature and Statistical Modeling Using Extreme Value Distributions

1.0 Introduction and the Focus of Chapter One

In real life, examples such as "*How tall should one design an embankment so that the sea reaches this level only once in 100 years?*"; "*What is the lowest value the Dow Jones Industrial Average can reach in the next three years"?*; "*How high can the drug concentration in bloodstream go before causing toxicity?* require estimation. But since no data or only few have been observed – as by definition extreme events are rare – a branch of statistics that helps us to deal with such rare situations and that gives a scientific alternative to pure guesswork is the Extreme Value Theory (EVT).

The explosion of the space shuttle Challenger in 1983 was the consequence of an extreme event: the exceptionally low temperature (15^{o} F lower than the next coldest previous launch) the night before launching ultimately led to failure of the O-rings (objects that are used to seal mechanical parts against fluid movement, air or liquid) which caused the disaster. Using standard EVT-analysis, one could have predicted that one should not launch at such cold temperature, despite having no measurements at such low temperatures.

If one seeks to estimate about everyday events, it might not matter if extreme data are cut off. But if one asks questions about events that do not happen very often, one should apply Extreme Value Theory; especially as these are the situations where one has the most to lose or win. For the layperson, events such as earthquakes, hurricanes, and

stock market crashes seem to follow no rule, but careful analysis has helped to discover distributions that acceptably model these extreme events [19].

The most important feature of an extreme value analysis is the objective to quantify the stochastic behavior of the maximum and the minimum of *i.i.d.* random variables. The distributional properties of extremes (maximum and minimum), extreme and intermediate order statistics, and exceedances over high (or below low) thresholds are determined by the upper and lower tails of the underlying distribution. The extreme value analysis requires the estimation of the probability of events that are more extreme than any that have already been observed [23]. Extreme value theory is a unique discipline that develops statistical techniques for describing the unusual phenomena such as rainfall, floods, wind gusts, air pollution, earthquakes, risk management, insurance, and financial matters.

Focus of Chapter One

Section 1.1 contains an extensive literature review that describes the evolution of the extreme value theory (EVT) to its present status. Section 1.2 focuses on the probability framework for the EVT.

Sections 1.3 through 1.7 present a survey of statistical modeling using Generalized Extreme Value (GEV), Gumble, Frechet, Weibull, and Generalized Pareto (GP) distributions. For each of these distributions, a thorough discussion of the probability density function, maximum likelihood estimation, quantile expression, and model diagnostics is given along with a numerical example for a few of them.

Section 1.8 provides a brief overview of the aims and objectives of the present study with emphasis being placed on each research problem.

1.1 Literature Review of Extreme Value Theory

Historically, work on extreme value problems can be traced back to as early as 1709 when Nicholas Bernoulli discussed the mean largest distance from the origin given *n* points lying at random on a straight line of a fixed length *t* [45]. Probably the first paper that described an application of extreme values in flood flows was by Fuller [35] in 1914, and Griffith [44], in 1920, brought out an application while discussing the phenomena of rupture and flow in solids. A paper written in 1922 by von Bortkiewicz [12] may have contributed to a systematic development of extreme value theory. His paper dealt with the distribution of range in random samples from a normal distribution, and the concept of *distribution of largest value* was introduced for the first time. In 1923, von Mises [69] evaluated the expected value of this distribution, and Dodd [27] calculated its median and discussed some non-normal parent distributions. A paper that had more direct relevance to the extreme value theory was written in 1927 by Frechet [33] in which he discussed the asymptotic distributions of largest values. In 1928, Fisher and Tippett [31] published the results of their research into the same problem. In addition, they showed that extreme limit distributions can only be one of three types. In 1936, von Mises [70] presented *sufficient* conditions for the weak convergence of the largest order statistic to each of the three types of limit distributions given by Fisher and Tippett. In 1943, Gnedenko [43], in this *breakthrough paper*, presented a solid foundation for the extreme value theory and provided *necessary and sufficient* conditions for the weak convergence of the extreme order statistics. Gnedenko's work was refined later on by many others that include Mejzler [68] in 1949 and de Haan [25] in 1970.

Following the theoretical developments of the extreme value theory during 1920s and mid 1930s, many scholarly papers dealing with the variety of practical applications of the theory were published in late 1930s and 1940s. E.J. Gumbel played a pioneering role during 1940s and 1950s, and from the application point of view, he made many significant contributions to the extreme value theory. He presented all of these in his *statistics of extremes* [45, 46] in 1958 and this work was pivotal in promoting extreme value theory as a tool for modeling the extremal behavior of observed physical processes. The Generalized Extreme Value (GEV), Gumbel, Frechet, Weibull, and the Generalized Pareto (GP) distributions are just the tip of the iceberg of an entirely new and quickly growing branch of statistics. The Gumbel distribution has light or medium tails, Frechet distribution has heavy tails, and Weibull distribution has bounded or short tails. The Pareto distribution is used to model how income is distributed, and to estimate finite limit of human lifespan [1].

Since the publication of Gnedenko's limit theorem [42] for maxima in 1941, and Gumbel's *statistics of extremes*, extreme value theory has found applications in engineering, environmental modeling, and finance [84]. Some recent applications of Gumbel distribution have been for fire protection and insurance problems [76], the prediction of earthquake magnitudes, modeling of extremely high temperatures [16], and the prediction of high return levels of wind speeds relevant for the design of civil engineering structures [71]. Recent applications of Frechet include estimation of the probabilities of extreme occurrences in Germany's stock index and prediction of the behavior of solar proton peak fluxes [15, 100]. Weibull applications include modeling of failure strengths of load-sharing and window glasses [49, 8], analysis of corrosion

failures of lead-sheathed cables at the Kennedy Space center [62], and estimating the occurrence probability of giant freak waves in the sea area around Japan [101]. Internet traffic, structural reliability and biotech analyses are few of the other prime targets for EVT applications, as their data distributions display heavy tails, too. Another more recent application of EVT is in the finance industry. A quantity known as Value-at-Risk (VaR) has become the standard risk measurement to protect portfolio holders against adverse market conditions and prevent them from taking extraordinary risks [5, 6]. The VaR is defined as the α-quantile of the Profit-and-Loss (P&L) distribution of value V_t at time $\boldsymbol{t}$ over the holding period, or horizon, $\boldsymbol{h}$. Existing standard methods to calculate VaR assume normality of the data which is inappropriate since the unconditional distribution of financial time series is known to be heavy-tailed. This gave birth to the use of EVT methods to model the tail and to estimate VaR more reliably. In 1975, Canfield [1**7**], and Canfield and Borgman [18] have discussed the usefulness of this distribution to model time-to-failure data in reliability studies. In 1986, Rossi *et al*. [78, 79, 10] proposed a two-component extreme value distribution for flood frequency analysis. In 1987, Achcar *et al*. [3] have discussed the advantage of transforming a survival data to Gumbel distribution form before analyzing it.

1.2 Probability Framework for Extreme Value Theory

Extreme Value Theory (EVT) is the study of probabilistic extremes and focuses primarily on the asymptotic behavior as sample size approaches infinity [23]. Let $X_1, X_2, ..., X_n$ be a sequence of independent random variables having a common distribution, **F**. The model focuses on the statistical behavior of

$$M_n = \max\{X_1, X_2, ..., X_n\}$$

where X_i^s usually represent values of a process measured on a regular time scale, for example, hourly measurements of stock prices or plasma drug concentrations over a certain period, so that M_n represents the maximum of the process over ***n*** time units of observation. If ***n*** is the number of observations in a day, then M_n corresponds to the daily maximum.

In theory, the distribution of M_n can be derived exactly for all values of ***n***:

$$\begin{aligned} \Pr\{M_n \le z\} &= \Pr\{X_1 \le z, ..., X_n \le z\} \\ &= \Pr\{X_1 \le z\} \bullet ..... \bullet \Pr\{X_n \le z\} = F^n(z) \end{aligned} \tag{1.1}$$

The difficulty that arises in practice is the fact that the distribution function **F** is unknown. One possibility is to use standard statistical techniques to estimate F from observed data, and then substitute this estimate into (1.1). But very small discrepancies in the estimate of F can lead to substantial discrepancies for F^n. This leads to an approach based on asymptotic argument which requires determining what possible limit distributions are possible for M_n as $n \to \infty$. The question then is "what are the possible limit distributions in the extremal case?"

The same problem arises in classical statistics when we insist that with probability 1, $\overline{X}_n$ converges to the population mean. In Central Limit Theorem (CLT), this problem is resolved by allowing a linear scaling, so that $\frac{\overline{X}_n - \mu_n}{\sigma_n} \to N(0,1)$ where $\mu_n = \mu$ and $\sigma_n = \frac{\sigma}{\sqrt{n}}$ are linear re-scalings which prevent the degenerate limits.

The same approach is adopted in obtaining the limits of the distribution of M_n, looking instead for limiting distributions of $\frac{M_n - b_n}{a_n}$ where a_n and b_n are sequences of normalizing coefficients such that $F^n\left(\frac{M_n - b_n}{a_n}\right)$ leads to a non-degenerate distribution as $n \to \infty$. Specifically, we seek $\{a_n > 0\}$ and $\{b_n\}$ such that $F^n\left(\frac{M_n - b_n}{a_n}\right) \to G(z)$ where $G(z)$ does not depend on $\boldsymbol{n}$.

Extremal Types Theorem **:** If there exist sequences of constants $\{a_n > 0\}$ and $\{b_n\}$ such that, as $n \to \infty$, $\Pr\left\{\frac{M_n - b_n}{a_n} \le z\right\} \to G(z)$ where G is a non-generate distribution function, then, G belongs to one of the following families [23]:

I. $$G(z) = \exp\left\{-\exp\left[-\left(\frac{z-b}{a}\right)\right]\right\}, \quad -\infty < z < \infty;$$

II. $$G(z) = \begin{cases} 0, & z \le b; \\ \exp\left\{-\left(\frac{z-b}{a}\right)^{-\alpha}\right\}, & z > b; \end{cases}$$

III. $$G(z) = \begin{cases} \exp\left\{-\left[-\left(\frac{z-b}{a}\right)^{\alpha}\right]\right\}, & z < b, \\ 1, & z \geq b; \end{cases}$$

for parameters a (scale) $> 0, b$ (location) and, in the case of II and III, α (shape) > 0.

In words, the theorem states that the rescaled sample maxima $\frac{M_n - b_n}{a_n}$ converge in distribution to a variable having a distribution within one of these families labeled I, II, and III. Collectively, the three classes of distribution are referred to as the extreme value distributions. The remarkable feature of this result is that the three types of extreme value distributions are the only possible limits for the distributions of $\frac{M_n - b_n}{a_n}$, regardless of the distribution F for the population. It is in this sense that the above theorem provides an extreme value analog of the central limit theorem.

The three types of limits that arise in Extremal Types Theorm have distinct forms of behavior, corresponding to the different forms of tail behavior for F of X_i. These distinct forms can be made precise by looking at the behavior of G at its upper-end point z_+. For the Weibull distribution z_+ is finite, while for both the Frechet and Gumbel distributions $z_+ = \infty$.

The three families of distributions - Gumbel, Frechet, and Weibull - can be combined into a single family of models having distribution functions of the form

$$G(z) = \exp\left\{-\left[1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} \qquad (1.2)$$

defined on $\{z : 1+\xi(z-\mu)/\sigma > 0\}$, where the parameters satisfy $-\infty < \mu < \infty$, $\sigma > 0$ and $-\infty < \xi < \infty$ [23].

The equation (1.2) above is the generalized extreme value family of distributions. This was obtained independently by Von Mises [70] and Jenkinson [54]. Equation (1.2) can be written as:

$$G(z) = \begin{cases} \exp\left\{-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} & -\infty < z \le \mu - \frac{\sigma}{\xi} \text{ for } \xi < 0; \\ & \mu - \frac{\sigma}{\xi} \le z < \infty \text{ for } \xi > 0; \\ \exp\left\{-\exp\left\{-\frac{z-\mu}{\sigma}\right\}\right\} & -\infty < z < \infty \text{ for } \xi = 0; \end{cases} \quad (1.3)$$

The parameters of this distribution are ξ (shape), μ (location), and σ (scale). The distribution in (1.3) is also referred to as the von Mises type extreme value distribution or the von Mises-Jenkinson type distribution.

Shape Parameter (ξ)

The importance of the shape parameter ξ is apparent from the above equations. If ξ is *negative*, the quantities are bounded above, and any extrapolation will lead to a finite limit. On the other hand if it is *0* or *positive*, then the limiting quantity is unbounded, and extrapolation leads to an infinite limit. The shape parameter ξ determines the shape, and depending on its value, the function can change drastically. From the sensitivity to the shape parameter, it is obvious extreme values will change drastically depending on the value of the shape parameter.

- The Type I (Gumbel) class of extreme value distribution is obtained in the limit as $\xi \to 0$. This is the case of an exponentially decreasing tail.
- The Type II (Frechet) class of extreme value distribution corresponds to case $\xi > 0$. When $\xi > 0$, the GEV distribution has a finite lower end point, given by $\mu - \sigma / \xi$. This is the case of a polynomially decreasing tail function and therefore, corresponds to a long-tailed parent distribution.
- The Type III (Weibull) class of extreme value distribution corresponds to case $\xi < 0$. When $\xi < 0$, the GEV has a finite upper end point, also given by $\mu - \sigma / \xi$. The case $\xi < 0$ is that of a finite upper endpoint, and is therefore short-tailed.

1.3 Generalized Extreme Value Distribution

1.3.1 Probability Density Function and Maximum Likelihood

The probability density [60] function corresponding to the equation (1.3) is

$$g(z)=\begin{cases}\exp-\left\{1+\xi\left(\dfrac{z-\mu}{\sigma}\right)\right\}^{\frac{-1}{\xi}}\dfrac{1}{\sigma}\left\{1+\xi\left(\dfrac{z-\mu}{\sigma}\right)\right\}^{\frac{-1}{\xi}-1} & -\infty<z\le\mu-\dfrac{\sigma}{\xi}\ \text{ for } \xi<0;\\ & \mu-\dfrac{\sigma}{\xi}\le z<\infty \quad \text{for } \xi>0;\\ \exp\left\{-e^{-\frac{z-\mu}{\sigma}}\right\}\dfrac{1}{\sigma}e^{-\frac{z-\mu}{\sigma}} & -\infty\le z<\infty \qquad \text{for } \xi=0\end{cases} \tag{1.4}$$

The standard forms of the probability density function and cumulative distribution function for the GEV distributions are obtained from the equations (1.4) and (1.3), respectively, by taking $\mu=0$ and $\sigma=1$.

Maximum Likelihood Estimation

Suppose we have *n* observations $Z_1, Z_2, Z_3, \ldots, Z_n$ for which the GEV distribution (1.3) is appropriate [23]. The log-likelihood for GEV parameters when $\xi\neq 0$ is

$$l(\mu,\sigma,\xi)=-n\log\sigma-\left(1+\frac{1}{\xi}\right)\sum_{i=1}^{n}\log\left[1+\xi\left(\frac{z_i-\mu}{\sigma}\right)\right]-\sum_{i=1}^{n}\left[1+\xi\left(\frac{z_i-\mu}{\sigma}\right)\right]^{\frac{-1}{\xi}}, \tag{1.5}$$

$$\text{provided that } 1+\xi\left(\frac{z_i-\mu}{\sigma}\right)>0,\ \text{for } i=1,2,\ldots..,n \tag{1.6}$$

At parameter combinations for which (1.6) is violated, corresponding to a configuration for which at least one of the observed data falls beyond an end-point of the distribution, the likelihood is zero and the log-likelihood is -∞.

The case $\xi = 0$ requires separate treatment using the Gumbel limit of the GEV distribution. This leads to the log-likelihood

$$l(\mu,\sigma) = -n\log\sigma - \sum_{i=1}^{n}\left(\frac{z_i - \mu}{\sigma}\right) - \sum_{i=1}^{n}\exp\left\{-\left(\frac{z_i - \mu}{\sigma}\right)\right\} \tag{1.7}$$

Maximization of the pair of equations (1.5) and (1.7) with respect to the parameter vector (μ,σ,ξ) leads to the maximum likelihood estimate with respect to the entire GEV family. There is no analytical solution, but for a given dataset the maximization is straightforward using standard numerical optimization algorithms [23].

The maximum likelihood estimators are the values of the unknown parameters that maximize the log-likelihood. In practice these are local maxima found by nonlinear optimization. The regularity conditions that are required for the usual asymptotic properties associated with the maximum likelihood estimator are *not* satisfied by the GEV model because the end-points of the GEV distribution are functions of the parameters; $\mu - \frac{\sigma}{\xi}$ is an upper end-point of the distribution when $\xi < 0$, and a lower end-point when $\xi > 0$. The violation of the regularity conditions means that the standard asymptotic likelihood results are not automatically applicable. Smith [80] studied this problem and obtained the following results:

- When $\xi < -1$, maximum likelihood estimators are unlikely to be obtainable.
- When $-1 < \xi < -1/2$ maximum likelihood estimators are generally obtainable, but do not have the standard asymptotic properties.
- When $\xi > 1/2$, the second and higher moments do not exist.

The standard asymptotic results of consistency, asymptotic efficiency and asymptotic normality hold for these distributions when $\xi > -1/2$.

The elements of the matrix of second-order partial derivatives, evaluated at the maximum likelihood estimators, are known as the *observed information matrix*, and the inverse of this matrix is the *variance-covariance matrix* of the maximum likelihood estimators. The square roots of the diagonal entries of this inverse matrix are estimates of the standard deviations of the three parameter estimates, widely known as *standard errors* of those estimates. All these results are asymptotic approximations valid for large sample sizes, but in practice they are widely used even when the sample sizes are fairly small.

1.3.2 Quantile

The interest is in the quantile z_p for which $G(z_p) = 1 - p$ **,** where ***G*** denotes the cumulative distribution function (*cdf*). Estimates of extreme quantiles are obtained by inverting $G(z) = \exp\left\{-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}$ as follows:

$$z_p = \mu - \frac{\sigma}{\xi}\left[1-\{-\log(1-p)\}^{-\xi}\right], \text{ where } G(z_p) = 1-p$$

$$z_p = \begin{cases} \mu - \frac{\sigma}{\xi}\left[1-\{-\log(1-p)\}^{-\xi}\right] & \text{for } \xi \neq 0 \\ \mu - \sigma\log\{-\log(1-p)\} & \text{for } \xi = 0 \end{cases} \qquad (1.8)$$

z_p is the return level associated with the return period (1/p), since to a reasonable degree of accuracy, the level z_p is expected to be exceeded on average once every (1/p) years,

and more precisely, z_p is exceeded by the maximum in any particular year with probability p .

1.3.3 Model Diagnostics

For checking the validity of a GEV model, graphical goodness-of-fit checks such as probability, quantile, return level and density plots are discussed next [23].

*Probability Plot***:** The probability plot is a comparison of the empirical and fitted distribution functions. With ordered data $z_{(1)} \le z_{(2)} \le ..... \le z_{(n)}$, the empirical distribution function evaluated at $z_{(i)}$ is given by $\widetilde{G}(z_{(i)}) = i/(n+1)$

By substituting the GEV parameter estimates into equation (1.4), the corresponding model based estimates are $\hat{G}(z_{(i)}) = \exp\left\{-\left[1+\hat{\xi}\left(\frac{z_{(i)}-\hat{\mu}}{\hat{\sigma}}\right)\right]^{-1/\hat{\xi}}\right\}$.

If the GEV model fits the data well, $\hat{G}(z_{(i)}) \approx \widetilde{G}(z_{(i)})$ for each *i*, so a probability plot, consisting of the points $\left\{\left(\widetilde{G}(z_{(i)}), \hat{G}(z_{(i)})\right), i = 1,2,.....n\right\}$ should lie close to the unit diagonal. The drawback of the probability plot for extreme value models is that $\hat{G}(z_{(i)})$ and $\widetilde{G}(z_{(i)})$ are bound to approach 1 as $z_{(i)}$ increases. Since the accuracy of the model for large values of z is of interest in extreme value problems, the probability plot seems to provide inadequate information in the region of most interest.

Quantile Plot: The deficiency of the probability plot can be eliminated by the quantile plot, consisting of the points $\left\{\left(\hat{G}^{-1}\left(\frac{i}{n+1}\right), z_{(i)}\right), i = 1,2,...n\right\}$ where

$$\hat{G}^{-1}\left(\frac{i}{m+1}\right) = \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}}\left[1 - \left\{-\log\left(\frac{i}{n+1}\right)\right\}^{-\xi}\right]$$

If the GEV model is an adequate fit, then the points on the quantile plot should lie close to a unit diagonal.

*Return Level Plot***:** Since quantiles enable probability models to be expressed on the scale of data, the relationship of the GEV model to its parameters is most easily interpreted in terms of the quantile expression (1.8). Suppose we define $y_p = -\log(1-p)$.

$$\text{Then, } z_p = \begin{Bmatrix} \mu - \frac{\sigma}{\xi}\left[1 - y_p^{-\xi}\right] & \text{for } \xi \neq 0, \\ \mu - \sigma\log(y_p) & \text{for } \xi = 0 \end{Bmatrix}$$

Suppose z_p is plotted against y_p on a logarithmic scale (or equivalently, if z_p is plotted against $\log y_p$). The plot is linear for ξ=0. If $\xi < 0$, the plot is convex with asymptotic limit as $p \to 0$ at $\mu - \sigma/\xi$. If $\xi > 0$, the plot is concave and has no finite bound [23].

This graph is a *return level plot*. The return level plots are convenient for both model presentation and validation. The tail of the distribution is compressed so that return level estimates for long return periods are displayed, while the linearity of the plot in the case ξ=0 provides a baseline against which to judge the effect of the estimated shape parameter. Confidence interval and empirical estimates of the return level function enable the return level plot to be used as a model diagnostic. If the GEV model is an adequate fit

for the data, the model-based curve and empirical estimates should be in reasonable agreement [23].

*Interpretation of a T-year Return Level***:** If *P(z)* is the probability of a level ***z*** being exceeded in a single year, then that level is often said to have a return period, which is in the inverse of *P(z)* years. For example, a sea level having a probability of being exceeded in a year of 0.01 is said to have a return period of 100 years. A sea level that has a probability of being exceeded once in hundred years is called the *100-year return level*.

*Density Plot***:** The probability, quantile and return level plots are based on a comparison of model-based and empirical estimates of the distribution function. Another diagnostic is based on the comparison of the probability density function of a fitted model with the histogram of the data. This comparison is less informative than the previous plots since the form of a histogram depends on the choice of grouping interval. There is no unique estimator of a density function, and hence, the comparison is more difficult and subjective.

1.3.4 GEV Modeling of Non-stationary Processes

A process is referred to as a non-stationary process if its characteristics change systematically with respect to change in time [23]. For example, the basic level of the annual maximum rainfall may change linearly over an observation period. If the distribution to be used is GEV, then it follows that a suitable model for $\boldsymbol{X_t}$, the annual maximum rainfall level in year ***t*** can be written as $X_t \oplus$ GEV $(\mu(t), \sigma, \xi)$, where $\mu(t) = \beta_0 + \beta_1 t$ for parameters β_0 and β_1. This models the variations through time in the observed process as a linear trend in the location parameter of the generalized extreme value model. The parameter β_1 represents the annual rate of change in annual maximum rainfall.

More complex changes in the location parameter μ may be modeled as a quadratic model given by $\mu(t) = \beta_0 + \beta_1 t + \beta_2 t^2$

Non-stationarity can also be expressed in terms of other extreme value parameters such as $\sigma(t) = e^{\beta_0 + \beta_1 t}$ where the exponential function is used to ensure that the positivity of σ is respected for all values of **t**. Modeling the extreme value model shape parameters as a smooth function of time is rather unrealistic since the shape parameters are difficult to estimate with precision [23].

An existence of a situation in which the extremal behavior of one series being related to that of another variable, referred to as a covariate, is a possibility. For example, in the statistical modeling of maximum drug concentrations in a pharmacokinetic study using extreme value theory, the variable, age, can be one of the covariates.

1.4 Gumbel Distribution

The Type I extreme value distribution has two forms. One is based on the smallest extreme and the other is based on the largest extreme. These are referred to as the *minimum* and *maximum* cases, respectively. The Gumbel distribution is also referred to as a Type I extreme value distribution or log-Weibull distribution.

1.4.1 Probability Density Function and Maximum Likelihood

Probability Density Function

Maximum Case: The probability density function of the Gumbel distribution is

$$f(z) = \frac{1}{\sigma}\exp\left[-\left(\frac{z-\mu}{\sigma}\right) - \exp\left(-\left(\frac{z-\mu}{\sigma}\right)\right)\right]$$ where μ is the location parameter and σ is the scale parameter.

Minimum Case: The probability density function of the Gumbel distribution is

$$f(z) = \frac{1}{\sigma}\exp\left[\left(\frac{z-\mu}{\sigma}\right) - \exp\left(\frac{z-\mu}{\sigma}\right)\right]$$ where μ is the location parameter and σ is the scale parameter.

Cumulative Distribution Function

Maximum Case: The cumulative distribution function of the Gumbel distribution is

$$F(z) = \exp\left\{-\exp\left(-\frac{z-\mu}{\sigma}\right)\right\}, -\infty < z < \infty$$

Minimum Case: The cumulative distribution function of the Gumbel distribution is

$$F(z) = 1 - \exp\left\{-\exp\left(\frac{z-\mu}{\sigma}\right)\right\}, -\infty < z < \infty$$

Maximum Likelihood Estimation

From equation (1.5), the log-likelihood for the GEV parameters when $\xi \neq 0$ is

$$l(\mu,\sigma,\xi) = -n\log\sigma - \left(1+\frac{1}{\xi}\right)\sum_{i=1}^{n}\log\left[1+\xi\left(\frac{z_i-\mu}{\sigma}\right)\right] - \sum_{i=1}^{n}\left[1+\xi\left(\frac{z_i-\mu}{\sigma}\right)\right]^{\frac{-1}{\xi}}, \qquad (1.9)$$

provided that $1+\xi\left(\frac{z_i-\mu}{\sigma}\right) > 0,$ for $i = 1,2,.....,n$

When $\xi \to 0$, (1.9) reduces to the log-likelihood function (23) for Gumbel distribution as

$$l(\mu,\sigma) = -n\log\sigma - \sum_{i=1}^{n}\left(\frac{z_i-\mu}{\sigma}\right) - \sum_{i=1}^{n}\exp\left\{-\left(\frac{z_i-\mu}{\sigma}\right)\right\} \qquad (1.10)$$

In order to find the MLEs of μ and σ, we need to maximize (1.10). We first set

$$\frac{\partial l}{\partial \mu} = 0 \text{ and } \frac{\partial l}{\partial \sigma} = 0.$$

After some mathematical manipulations, we get two equations with two unknowns,

$$\hat{\mu} = -\hat{\sigma}\ln\left[\frac{1}{n}\sum_{i=1}^{n}\exp\left(\frac{-z_i}{\hat{\sigma}}\right)\right] \qquad (1.11)$$

$$\hat{\sigma} - \frac{1}{n}\sum_{i=1}^{n} z_i + \frac{\sum_{i=1}^{n} z_i \exp\left(-\frac{z_i}{\hat{\sigma}}\right)}{\sum_{i=1}^{n}\exp\left(-\frac{z_i}{\hat{\sigma}}\right)} = 0 \qquad (1.12)$$

These equations can be solved numerically to obtain $\hat{\mu}$ and $\hat{\sigma}$.

The corresponding 95% confidence intervals for $\hat{\mu}$ and $\hat{\sigma}$ given by

$$\left(\hat{\mu}-1.96\sqrt{\frac{1.10867\hat{\sigma}^2}{n}},\ \hat{\mu}+1.96\sqrt{\frac{1.10867\hat{\sigma}^2}{n}}\right) \text{ and}$$

$$\left(\hat{\sigma}-1.96\sqrt{\frac{0.60793\hat{\sigma}^2}{n}},\ \hat{\mu}+1.96\sqrt{\frac{0.60793\hat{\sigma}^2}{n}}\right)$$

1.4.2 Quantile

The interest is in the quantile, z_p, for which $F(z_p)=1-p$ where F denotes the cumulative distribution function given by

$$F(z)=\exp\left(-\exp\left(-\frac{z-\mu}{\sigma}\right)\right) \tag{1.13}$$

Inverting $F(z_p)=1-p$ for the above form of F, we get $z_p=\mu-\sigma\log\{-\log(1-p)\}$ and hence, the MLE of $F(z_p)=1-p$ is given by

$$\hat{z}_p=\hat{\mu}-\hat{\sigma}\log\{-\log(1-p\} \tag{1.14}$$

1.4.3 Model Diagnostics

Quantile-Quantile (Q-Q): The Q-Q plot is a graphical tool to assess the fit of the model to the data. For the Gumbel distribution, the Q-Q plot is y_i versus $z_{(i)}$ where $z_{(i)}^s$ are obtained by arranging the data in the ascending order so that $z_{(1)}<z_{(2)}<.....<z_{(n)}$ are known as the order statistics or the observed quantiles, and y_i^s, known as the expected quantiles, are computed using $y_i=\hat{\mu}-\hat{\sigma}\ln\left(\ln\left(\frac{n+1}{i}\right)\right),\ i=1,2,3,....n.$

If the points lie close to the 45 degree straight line it is an indication of an adequate fit of the Gumbel distribution to the data [23].

Gumbel Probability Plot: The Gumbel probability plot is the plot of y_i versus $z_{(i)}$ where $z_{(i)}^s$ are the observed data points that are sorted from the smallest to the largest to form the order statistics, $Z_{(1)} < Z_{(2)} < \ldots\ldots\ldots < Z_{(n)}$. The values of y_i^s are obtained from $y_i = F^{-1}\left(\frac{i}{n+1}\right)$ where n is the number of observed data points, and F^{-1} is the inverse Gumbel cumulative distribution function given by $F^{-1}(z) = \mu - \sigma \ln\left(\ln\left(\frac{1}{z}\right)\right)$.

Note that μ and σ are <u>*not*</u> needed to check the linearity of set of points $\left(z_{(i)}, F^{-1}\left(\frac{i}{n+1}\right)\right)$,

that is, $\left(z_{(i)}, -\ln\left(\ln\left(\frac{1}{\left(\frac{i}{n+1}\right)}\right)\right)\right) = \left(z_{(i)}, -\ln\left(\ln\left(\frac{n+1}{i}\right)\right)\right)$ for $\boldsymbol{i}$ = 1, 2, 3, ……n.

The probability plot for the Gumbel distribution will be linear if and only if the plot of $\left(z_{(i)}, -\ln\left(\ln\left(\frac{n+1}{i}\right)\right)\right)$ is linear for $\boldsymbol{i}$ = *1, 2, 3, ……n*.

If the plot produces points that fall close to a straight line, then the Gumbel distribution is a reasonable model.

1.4.4 Statistical Modeling of Annual Maximum Discharge

Following are the data that consist of $\boldsymbol{n}$ = 59 annual maximum discharges of Feather River in ft^3/sec from 1902 to 1960 [22, 26].

230000, 203000, 185000, 185000, 152000, 140000, 135000, 128000, 122000, 118000, 113000, 110000, 108000, 102000, 102000, 94000, 92100, 85400, 84200, 83100, 81400, 81000, 80400, 80100, 75400, 65900, 64300, 62300, 60100, 59200, 58600, 55700, 54800,

54400, 46400, 45600, 42400, 42400, 42000, 36700, 36400, 34500, 31000, 28200, 24900, 23400, 22600, 22400, 20300, 19200, 16800, 16800, 16400, 16300, 14000, 13000, 11600, 8860, 8080

The probability plot in Figure 1.4.1 seems linear, implying the adequacy of the Gumbel model to fit to the data. The maximum likelihood parameters were computed as $\hat{\mu} = 47309.4$ and $\hat{\sigma} = 37309.1$.

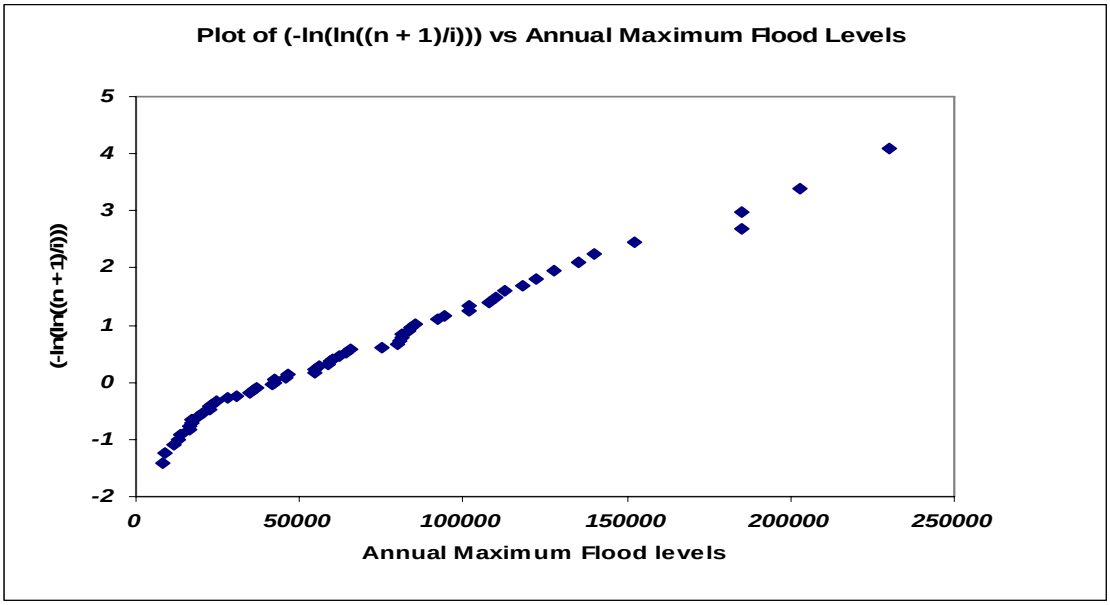


Figure 1.4.1 Probability plot for the annual maximum discharges of Feather River.

The probability that the maximum flood level observed during a given year does not exceed a certain level L is given by $P(Z_j \le L)$. In the data set the largest value observed is 230,000. Therefore, the probability of exceeding this value of 230000 in any given year can be calculated as follows [22]:

$$P(Z_j > 230000) = 1 - F(Z_j \le 230000) = 1 - \exp\left[-\exp\left(-\frac{230000 - 47309.41816}{37309.08825}\right)\right]$$

That is, $P(Z_j > 230000) = 0.00744$.

Thus, the return period is $1/0.00744 = 134$ years.

To compute T-year level u(T) (where u(T) = T year flood level) where T = 100 year

$$P(Z_1 > u(T)) = \frac{1}{T} = 1 - F(u(T)) \Rightarrow \quad F(u(T)) = 1 - \frac{1}{T} \Rightarrow u(T) = F^{-1}(1 - \frac{1}{T})$$

$$u(100) = F^{-1}\left(1 - \frac{1}{100}\right) = F^{-1}\left(\frac{99}{100}\right) = 47309.41816 - 37309.08825 * \ln\left(\ln\left(\frac{1}{99/100}\right)\right) = 218{,}937$$

Therefore, the 100-year return level is 218937. The graph in Figure 1.4.2 displays the predicted return levels as a function of return period.

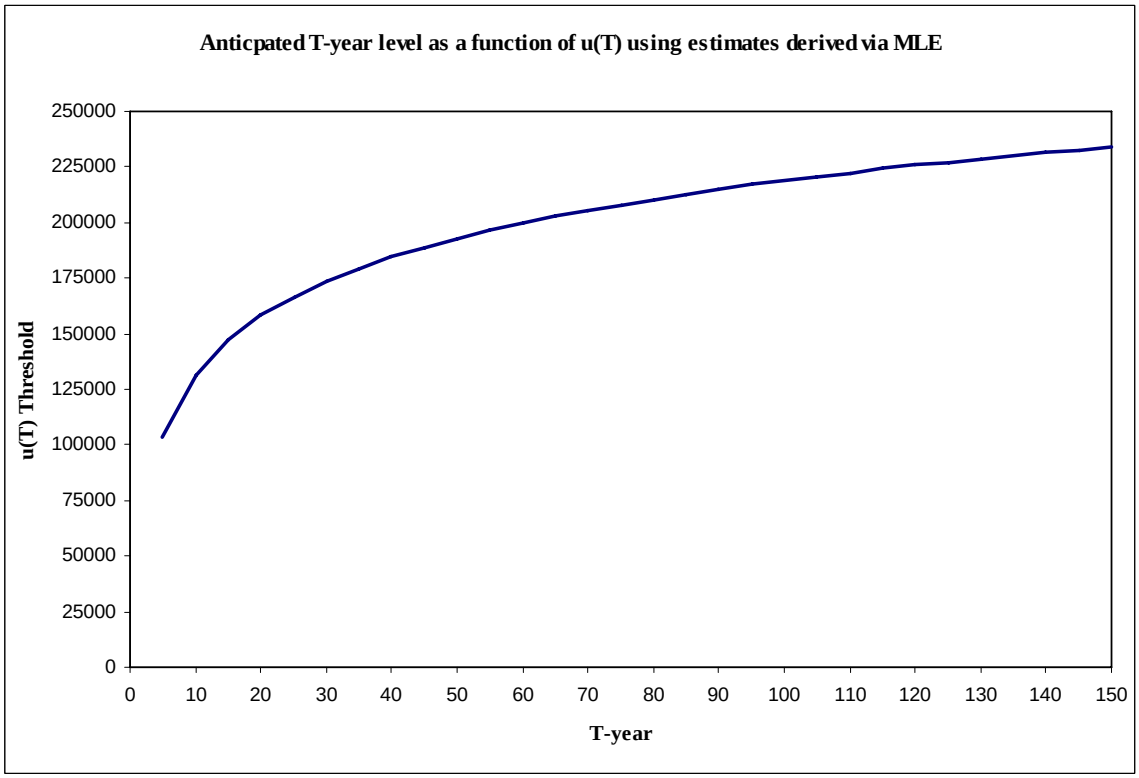


Figure 1.4.2 Graph of u(T) Threshold versus T-year for the annual maximum discharges of Feather River data set.

1.5 Frechet Distribution

The Frechet is the second extreme value distribution, also known as the Type II extreme value distribution. It has wide ranging applications in engineering, environmental modeling, finance and other areas. Recent applications include prediction of solar proton peak fluxes and modeling interfacial damage in microelectronic packages and material properties of constituent particles in an aluminum alloy [96].

1.5.1 Probability Density Function and Maximum Likelihood

When $\xi > 0$, the cumulative distribution function of the GEV distribution reduces to reduces to the cumulative distribution function for Frechet distribution given by

$$F(z) = \exp\left\{-\left(\frac{\sigma}{z-\mu}\right)^{\xi}\right\}$$

The corresponding probability density function is given by

$$f(z) = \frac{\xi}{\sigma}\left(\frac{\sigma}{z-\mu}\right)^{\xi+1} \exp\left\{-\left(\frac{\sigma}{z-\mu}\right)^{\xi}\right\} \tag{1.15}$$

where ξ, μ and σ are the shape, location and scale parameters, respectively.

Maximum Likelihood Estimation

For a random sample $z_1, z_2, ..., z_n$, the likelihood function is given by

$$l(\mu,\sigma,\xi) = \xi^n \sigma^{n\xi}\left\{\prod_{i=1}^{n}(t_i-\mu)\right\}^{-(\xi+1)} \exp\left\{-\sigma^{\xi}\sum_{i=1}^{n}(t_i-\mu)^{-\xi}\right\}$$

The maximum likelihood estimates $\hat{\mu}, \hat{\sigma}$, and $\hat{\xi}$ are obtained by setting [96, 65]

$$\frac{\partial \log L}{\partial \mu} = 0, \frac{\partial \log L}{\partial \sigma} = 0, \text{ and } \frac{\partial \log L}{\partial \xi} = 0 \text{ at } \mu = \hat{\mu}, \sigma = \hat{\sigma}, \text{ and } \xi = \hat{\xi} \text{ respectively.}$$

After some mathematical simplification, we get the equations

$$\frac{n}{\hat{\xi}}+\frac{n\sum_{i=1}^{n}(t_i-\hat{\mu})^{-\hat{\xi}}\log(t_i-\hat{\mu})}{\sum_{i=1}^{n}(t_i-\hat{\mu})^{-\hat{\xi}}}=\sum_{i=1}^{n}\log(t_i-\hat{\mu}) \tag{1.16}$$

$$\frac{n\hat{\xi}\sum_{i=1}^{n}(t_i-\hat{\mu})^{-(\hat{\xi}+1)}}{\sum_{i=1}^{n}(t_i-\hat{\mu})^{-\hat{\xi}}}=(\hat{\xi}+1)\sum_{i=1}^{n}\frac{1}{t_i-\hat{\mu}} \tag{1.17}$$

and $$\hat{\sigma}=\left\{\frac{1}{n}\sum_{i=1}^{n}(t_i-\hat{\mu})^{-\hat{\xi}}\right\}^{-1/\hat{\xi}} \tag{1.18}$$

These equations can be solved numerically to obtain the estimates $\hat{\mu},\hat{\sigma}$, and $\hat{\xi}$.

1.5.2 Quantile

The interest is in the quantile z_p for which $F(z_p)=1-p$, where F denotes the cumulative distribution function [96]. Inverting $F(z_p)=1-p$ for $F(z)$,

$$1-p=\exp\left\{-\left(\frac{\sigma}{z_p-\mu}\right)^{\xi}\right\}\Rightarrow\{-\log(1-p)\}=\left\{\frac{\sigma}{z_p-\mu}\right\}$$

$$\Rightarrow z_p=\mu+\sigma\{-\log(1-p)\}^{-\frac{1}{\xi}} \tag{1.19}$$

By the invariance property of the method of maximum likelihood, the *mle* of (1.19) is given by

$$\hat{z}_p=\hat{\mu}+\hat{\sigma}\{-\log(1-p)\}^{-\frac{1}{\hat{\xi}}} \tag{1.20}$$

For large n, we can assume that equation (1.20) is normally distributed with the mean given by equation (1.19) and the variance given by

$$\operatorname{var}\left(\hat{z}_p\right) \approx \operatorname{var}(\hat{\mu}) + \{-\log(1-p)\}^{-2/\xi} \operatorname{var}(\hat{\sigma}) + \frac{\sigma^2}{\xi^4}\{-\log(1-p)\}^{-2/\xi}\left[-\log\{-\log(1-p)\}\right]^2 \operatorname{var}(\hat{\xi})$$
$$+ 2\{-\log(1-p)\}^{-1/\xi} \operatorname{cov}(\hat{\mu},\hat{\sigma}) + \frac{2\sigma}{\xi^2}\{-\log(1-p)\}^{-1/\xi} \log\{-\log(1-p)\}\operatorname{cov}(\hat{\mu},\hat{\xi})$$
$$+ \frac{2\sigma}{\xi^2}\{-\log(1-p)\}^{-2/\xi} \log\{-\log(1-p)\}\operatorname{cov}(\hat{\sigma},\hat{\xi})$$

The $100(1-\alpha)\%$ confidence interval for (1.19) is given by

$$\hat{z}_p \pm z_{\alpha/2}\sqrt{Var(\hat{z}_p)} \tag{1.21}$$

where $z_{\alpha/2}$ is the $100(1-\frac{\alpha}{2})\%$ percentile of the standard normal distribution [96].

1.5.3 Model Diagnostics

Model diagnostics for Frechet distribution is examined through the P-P Plot and Q-Q Plot in the same way as it was done in the case of Gumbel distribution.

1.6 Weibull Distribution

The Weibull is the third extreme value distribution, also known as the Type III extreme value distribution. It has wide ranging applications in engineering, environmental modeling, finance and other areas. Recent applications include evaluating the magnitude of future earthquakes in the Pacific, Argentina, Japan and in the Indian subcontinent [77].

1.6.1 Probability Density Function and Maximum Likelihood

When $\xi < 0$, the cumulative distribution function of the GEV distribution reduces to the cumulative distribution function for Weibull distribution given by

$$F(z) = 1 - \exp\left\{-\left(\frac{z-\mu}{\sigma}\right)^{\xi}\right\}$$

The probability density function for a 3-parameter Weibull model (61) is given by

$$f(z) = \frac{\xi}{\sigma}\left(\frac{z-\mu}{\sigma}\right)^{\xi-1} \exp\left\{-\left(\frac{z-\mu}{\sigma}\right)^{\xi}\right\} \tag{1.22}$$

where ξ, μ and σ are the shape, location and scale parameters, respectively.

The *pdf* for a 2-parameter Weibull is obtained by setting $\mu = 0$ in equation (1.22). The *pdf* for a 1-parameter Weibull is obtained by setting $\mu = 0$ in equation (1.22) and assuming $\xi = C = \text{constant}$. In the formulation of one-parameter Weibull we assume that the shape parameter ξ is known a priori, from past experience. The advantage of doing this is that data sets with fewer observations can be analyzed.

Maximum Likelihood Estimation

For a random sample $z_1, z_2, ..., z_n$, the likelihood function is given by

$$l(\mu,\sigma,\xi) = \xi^n \sigma^{-n\xi} \left\{ \prod_{i=1}^{n} (t_i - \mu) \right\}^{\xi-1} \exp\left\{ -\sigma^{-\xi} \sum_{i=1}^{n} (t_i - \mu)^{\xi} \right\}$$

The maximum likelihood estimates $\hat{\mu}, \hat{\sigma}$, and $\hat{\xi}$ are obtained by setting

$\frac{\partial \log L}{\partial \mu} = 0, \frac{\partial \log L}{\partial \sigma} = 0,$ and $\frac{\partial \log L}{\partial \xi} = 0$ at $\mu = \hat{\mu}, \sigma = \hat{\sigma},$ and $\xi = \hat{\xi}$ respectively.

After some mathematical simplification, we get the equations

$$\frac{n}{\hat{\xi}} + \sum_{i=1}^{n} \log(t_i - \hat{\mu}) = \frac{n \sum_{i=1}^{n} (t_i - \hat{\mu})^{\hat{\xi}} \log(t_i - \hat{\mu})}{\sum_{i=1}^{n} (t_i - \hat{\mu})^{\hat{\xi}}} \tag{1.23}$$

$$\frac{n\hat{\xi} \sum_{i=1}^{n} (t_i - \hat{\mu})^{\hat{\xi}-1}}{\sum_{i=1}^{n} (t_i - \hat{\mu})^{\hat{\xi}}} = (\hat{\xi} - 1) \sum_{i=1}^{n} \frac{1}{t_i - \hat{\mu}} \tag{1.24}$$

and $$\hat{\sigma} = \left\{ \frac{1}{n} \sum_{i=1}^{n} (t_i - \hat{\mu})^{\hat{\xi}} \right\}^{1/\hat{\xi}} \tag{1.25}$$

These equations can be solved numerically to obtain the estimates $\hat{\mu}, \hat{\sigma}$, and $\hat{\xi}$.

1.6.2 Quantile

The interest is in the quantile z_p for which $F(z_p) = 1 - p$, where F denotes the cumulative distribution function given by

$F(z) = 1 - \exp\left\{ -\left(\frac{z - \mu}{\sigma} \right)^{\xi} \right\}$, where ξ, μ and σ are as described above.

Inverting $F(z_p) = 1 - p$ for the above form of F, the expression for quantile is given by,

$$z_p = \mu + \sigma(-\log p)^{1/\xi} \tag{1.26}$$

By the invariance property of the method of maximum likelihood, the *mle* of (1.26) is

given by $$\hat{z}_p = \hat{\mu} + \hat{\sigma}\{-\log(p)\}^{\frac{1}{\hat{\xi}}} \tag{1.27}$$

For large n, we can assume that equation (1.27) is normally distributed with the mean given by equation (1.26) and the variance given by

$$\begin{aligned}\operatorname{var}(\hat{z}_p) \approx{}& \operatorname{var}(\hat{\mu}) + \{-\log p)\}^{2/\xi} \operatorname{var}(\hat{\sigma}) + \frac{\sigma^2}{\xi^4}\{-\log p\}^{2/\xi}[-\log\{-\log p\}]^2 \operatorname{var}(\hat{\xi}) \\ &+ 2\{-\log p\}^{1/\xi} \operatorname{cov}(\hat{\mu},\hat{\sigma}) - \frac{2\sigma}{\xi^2}\{-\log p\}^{1/\xi} \log\{-\log p\}\operatorname{cov}(\hat{\mu},\hat{\xi}) \\ &- \frac{2\sigma}{\xi^2}\{-\log p\}^{2/\xi} \log\{-\log p\}\operatorname{cov}(\hat{\sigma},\hat{\xi})\end{aligned}$$

The $100(1-\alpha)\%$ confidence interval for (1.26) is given by

$$\hat{z}_p \pm z_{\alpha/2}\sqrt{Var(\hat{z}_p)} \tag{1.28}$$

where $z_{\alpha/2}$ is the $100(1-\frac{\alpha}{2})\%$ percentile of the standard normal distribution.

1.6.3 Model Diagnostics

Model diagnostics for Weibull distribution is examined through the P-P Plot and Q-Q Plot in the same way as it was done in the case of Gumbel distribution.

1.6.4 Statistical Modeling of Annual Maximum Storm Surge

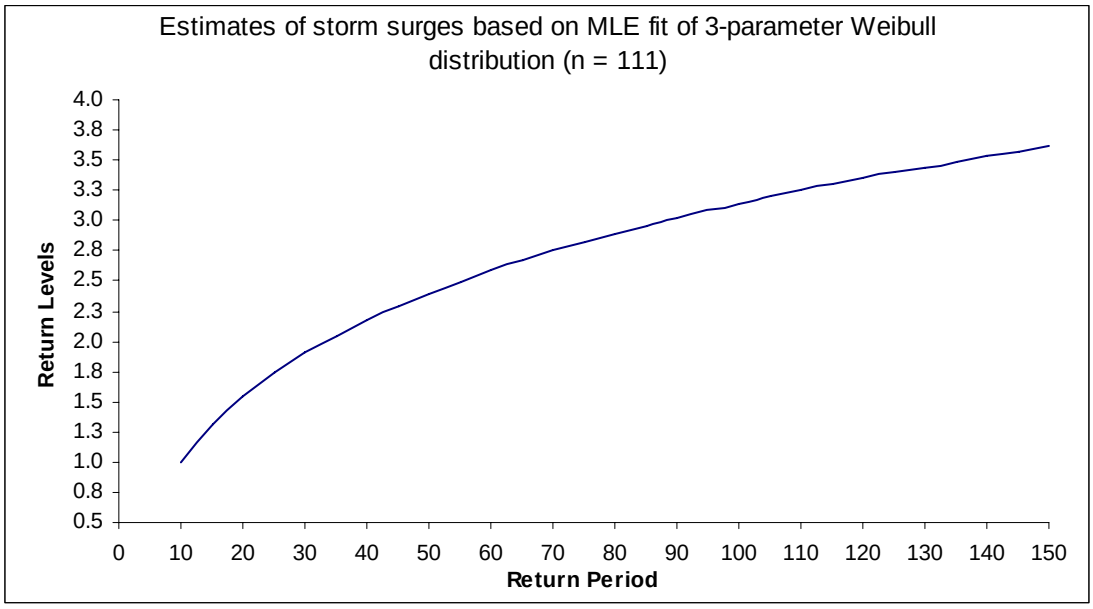
Storm surge return periods are useful to land use planners and emergency managers for assessing the likelihood of extreme water depths associated with tropical cyclones [45, 46]. At a given location, it is desirable to determine sound statistical estimates of return periods (for 20, 25, 50, 100-year epochs) and corresponding assessments of uncertainty due to both limitations in the historical record and effects of parameter estimation. Using the historical storm set consisting of 951 North Atlantic tropical cyclones for the period 1886-1996 for one particular location a total of 111 annual maximum storm surges were obtained. This example is discussed in detail in *Caribbean Storm Surge Return Periods: Final Report by Mark E. Johnson, December 1997* [55]. These 111 annual maximum storm surges are the data for fitting the model.

The parameter estimates for 2-parameter and 3-parameter Weibull fits are obtained using the Statistical Analysis Software (SAS) [91], version 9.1. In the case of 2-parameter Weibull, the location parameter is assumed to be zero whereas in the case of 3-parameter Weibull, it can be either set equal to a value less than the smallest data value or estimated.

The graphs of the return levels for 3-parameter and 2-parameterWeibull distribution fits are displayed in Figure 1.6.1. It can be noted that they are almost similar in the prediction of the return levels. The maximum likelihood estimates of the shape and scale parameters for the two Weibull distributions are given in Table 1.6.1

Table 1.6.1 Estimates of shape and scale parameters with their standard errors from 3-parameter and 2-parameter Weibull fits on the dataset containing 111 annual maximum storm surges.

MODEL	σ (SE)	ξ (SE)
3-Parameter Weibull	0.247296 (0.0412)	0.601609 (0.0449)
2-Parameter Weibull	0.291671 (0.0421)	0.696758 (0.0506)



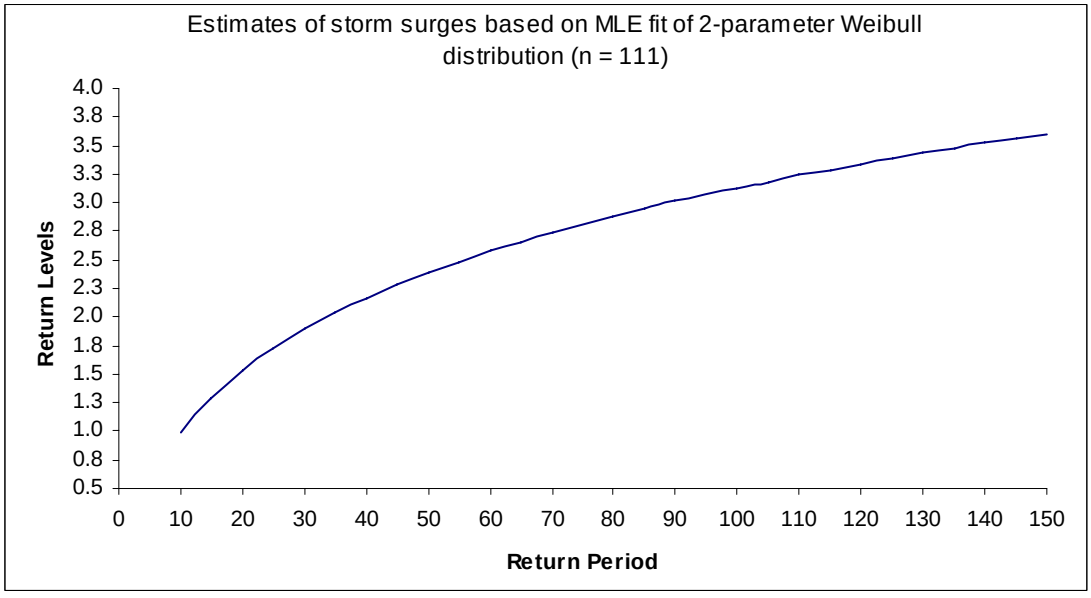


Figure 1.6.1 Estimates of storm surges based on MLE of 2- and 3-parameter Weibull

1.7 Generalized Pareto Distribution

Introduction

The principal drawback to the classical GEV/Gumbel method is that only one value is selected per epoch. This reduces the data available for analysis. To increase the number of cases for analysis, an alternative approach is to obtain "Peak-Over-Threshold (POT)" maxima, extracted from sample data series to produce a series of extreme values above a chosen (high) threshold, and use them with the Generalized Pareto Distribution. This is a common approach to extreme value statistics based on the exceedances over high thresholds [23, 81, 82, 83].

Probability Framework

Let $X_1, X_2, \ldots\ldots$ be a sequence of *iid* random variables, having marginal distribution function F. Let $\boldsymbol{u}$ be a certain high *threshold* value. Any $\boldsymbol{X_i}$ that exceeds $\boldsymbol{u}$ is considered an extreme event. Denoting an arbitrary term in the $\boldsymbol{X_i}$ sequence by $\boldsymbol{X}$, it follows that a description of the stochastic behavior of extreme events is given by the conditional probability:

$$P\left(X > u + y \mid X > u\right) = \frac{1 - F(u+y)}{1 - F(u)}, \quad y > 0.$$

Let $M_n = \text{Max}\{X_1, X_2, \ldots\ldots, X_n\}$. Denote an arbitrary term in the $\boldsymbol{X_i}$ sequence by $\boldsymbol{X}$, and suppose that F satisfies *Extremal Type Theorem* so that for large n, $\Pr\{M_n \le z\} \approx$ G(z) where $G(z) = \exp\left\{-\left[1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}$ for some μ, $\sigma > 0$ and ξ. Then, for large enough **u**, the distribution function of (X – **u**), conditional on X > **u**, is approximately:

$$H(y) = 1 - \left(1 + \frac{\xi y}{\tilde{\sigma}}\right)^{-1/\xi} \tag{1.29}$$

defined on $\left\{y : y > 0 \text{ and } \left(1 + \frac{\xi y}{\tilde{\sigma}}\right) > 0\right\}$, where $\tilde{\sigma} = \sigma + \xi(u - \mu)$ (1.30)

1.7.1 Probability Density Function and Maximum Likelihood

Probability Density Function

The probability density function, corresponding to the cumulative distribution function given in equation (1.29), is

$$h(y) = \frac{1}{\sigma}\left(1 + \frac{\xi y}{\sigma}\right)^{\frac{-1}{\xi}-1} \tag{1.31}$$

where σ and ξ are the scale and shape parameters, respectively.

The shape parameter ξ is dominant in determining the qualitative behavior of the GPD, just as it is for the GEV [23]:

- when ξ< 0 the distribution of excesses has an upper bound of $\left(u - \hat{\sigma} / \xi\right)$
- when ξ > 0 the distribution has no upper limit.
- when ξ = 0, the distribution is also unbounded, which should be interpreted by taking the limit of $H(y) = 1 - \left(1 + \frac{\xi y}{\tilde{\sigma}}\right)^{-1/\xi}$ as $\xi \to 0$, leading to $H(y) = 1 - e^{-y/\hat{\sigma}}, y > 0$ which corresponds to an exponential distribution with parameter $1/\hat{\sigma}$.

Maximum Likelihood Estimation

Suppose that the values $y_1, y_2, ..., y_k$ are the *k* excesses of a threshold *u*.

For $\xi \neq 0$, the log-likelihood is derived as

$$l(\sigma,\xi) = -k\log\sigma - \left(1+\frac{1}{\xi}\right)\sum_{i=1}^{k}\log\left(1+\frac{\xi y_i}{\sigma}\right), \tag{1.32}$$

provided $\left(1+\frac{\xi y_i}{\sigma}\right) > 0$ for $i = 1,2,3,.....,k;$ otherwise, $l(\sigma,\xi) = -\infty$.

In the case $\xi = 0,$ the log-likelihood is obtained from $H(y) = 1 - e^{-y/\hat{\sigma}}$ as

$$l(\sigma) = -k\log\sigma - \frac{1}{\sigma}\sum_{i=1}^{k} y_i$$

Analytical maximization of the log-likelihood is not possible, so numerical techniques are again required, taking care to avoid numerical instabilities when $\xi \approx 0$ in (1.32).

1.7.2 Quantile

It is more convenient to interpret extreme value models in terms of *quantiles* (or *return levels*). Suppose that a generalized Pareto distribution with parameters σ and ξ is a suitable model for exceedances of a threshold *u* by a variable *X* [23].

That is, for X > *u*, $\Pr\{X > x \mid X > u\} = \left[1+\xi\left(\frac{x-u}{\sigma}\right)\right]^{-1/\xi}$

It follows that $\Pr\{X > x\} = \Pr\{X > u\}\left[1+\xi\left(\frac{x-u}{\sigma}\right)\right]^{-1/\xi}$

The level x_m that is exceeded on average once every *m* observations is the solution of $\Pr\{X > u\}\left[1+\xi\left(\frac{x_m-u}{\sigma}\right)\right]^{-1/\xi} = \frac{1}{m}$. Solving for x_m and using $\tau_u = \Pr\{X > u\},$ we get

$$\begin{aligned} x_m &= u + \frac{\sigma}{\xi}\left[(m\tau_u)^{\xi} - 1\right] \text{ for } \xi \neq 0 \\ &\quad u + \sigma\log(m\tau_u) \qquad \text{for } \xi = 0 \end{aligned} \tag{1.33}$$

provided m is sufficiently large to ensure that $x_m > u$.

x_m is the m-observation return level. Plotting x_m against m on a logarithmic scale produces the same qualitative features as return level plot based on GEV model: linearity if $\xi = 0$; convexity if $\xi < 0$; concavity if $\xi > 0$.

It is more convenient to give quantiles or return levels on an annual scale, so that N-year return level is the level expected to be exceeded once every N years. If there are n_y observations per year, this corresponds to the m-observation return level with $m = N*n_y$. Hence, the N-year return level is given by

$$\begin{aligned} z_N &= u + \frac{\sigma}{\xi}\left[\left(N * n_y \tau_u\right)^{\xi} - 1\right] \text{ for } \xi \neq 0 \\ &\quad u + \sigma \log(N * n_y \tau_u) \quad \text{for } \xi = 0 \end{aligned} \tag{1.34}$$

In order to find z_N, we use the maximum likelihood estimates of σ and ξ, and an estimate of τ_u. The probability of an individual observation exceeding the threshold u is τ_u. This has a natural estimator of $\hat{\tau}_u = k/n,$ the sample proportion of observations exceeding u. Since the number of exceedances of u follows the binomial $B(n, \tau_u)$ distribution, $\hat{\tau}_u$ is also the maximum likelihood estimator of τ_u.

The variance of x_m can be derived by the delta method, but the uncertainty in the estimate of τ_u should also be included in the calculation. From the properties of the binomial distribution, $Var(\hat{\tau}_u) \approx \hat{\tau}_u(1-\hat{\tau}_u)/n$.

Then, the complete variance-covariance matrix for $\left(\hat{\tau}_u, \hat{\sigma}, \hat{\xi}\right)$ is approximately

$V = \begin{bmatrix} \text{var}(\hat{\tau}_u) & 0 & 0 \\ 0 & v_{1,1} & v_{1,2} \\ 0 & v_{2,1} & v_{2,2} \end{bmatrix}$, where $v_{i,j}$ denotes the (i, j) term of the variance-covariance

matrix of $\hat{\sigma}$ and $\hat{\xi}$. By the delta method, $Var(\hat{x}_m) \approx \nabla x_m^T \, V \, \nabla x_m$ where

$\nabla x_m^T = \left[\frac{\partial x_m}{\partial \tau_u}, \frac{\partial x_m}{\partial \sigma}, \frac{\partial x_m}{\partial \xi} \right]$ evaluated at $\left(\hat{\tau}_u, \hat{\sigma}, \hat{\xi}\right)$.

1.7.3 Model Diagnostics and Methods of Selecting a Threshold Value

Model Diagnostics

If the generalized Pareto model is reasonable for modeling excesses of *u*, then both *probability* plot and *quantile* plot are approximately linear.

The *return level plot* consists of the locus of points $\{(m, \hat{x}_m)\}$ for large values of *m* where $\hat{x}_m = u + \frac{\hat{\sigma}}{\hat{\xi}}\left[(m\hat{\tau}_u)^{\hat{\xi}} - 1\right]$

The density function of the fitted generalized Pareto model can be compared to a histogram of exceedances.

Methods of Selecting a Threshold Value

Selecting an appropriate threshold value is very critical. Too low a threshold value is likely to violate the asymptotic basis of the model, leading to bias. Too high a threshold value might result in obtaining too few excesses, leading to high variance. The standard practice is to adopt as low a threshold as possible, subject to the limit model providing a reasonable approximation [23].

First Method: This method of selecting a threshold value is an exploratory technique carried out prior to model estimation. The method is based on the mean of the generalized Pareto distribution. This requires plotting the points $\left\{\left(u, \frac{1}{n_u}\sum_{i=1}^{n_u}\left(x_{(i)} - u\right)\right) : u < x_{\max}\right\}$ where $x_{(1)}, x_{(2)}, ..., x_{(n_u)}$ consist of the n_u observations

that exceed u, and $x_{\max}$ is the largest of the X_i. This plot is called *mean residual life plot* or *the mean excess function* [90]. This plot should be approximately linear in u above a threshold value u_0 at which the generalized Pareto distribution provides a valid approximation to the excess distribution. Confidence limits can be added to the plot based on the approximate normality of sample means.

Second Method: This method is an *assessment of the stability of the parameter estimates* based on the fitting of models across a range of different threshold values. Above a level u_0 at which the asymptotic motivation for the generalized Pareto distribution is valid, estimates of the shape parameter ξ should be approximately constant, whereas the estimates of σ_u should be linear in u.

1.7.4 Statistical Modeling of Daily Precipitation Data

The data set used has n = 36524 observations during the time period 1900 – 1999 for a single location in Fort Collins, Colorado, USA [59].

Exceedance rate (per year) = 10.24. The value of negative log likelihood is 109.91. The model diagnostic plots are displayed in Figure 1.7.1.

Table 1.7.1 Estimates of scale and shape parameters with standard errors from the generalized Pareto model on the Fort Collins precipitation data.

PARAMETER	ESTIMATE	STANDARD ERROR
σ (scale)	0.33972	0.0164
ξ (shape)	0.18684	0.0374

The Table 1.7.2 provides the estimates of shape parameter and 100-year return level with 95% profile likelihood confidence intervals. Figures 1.7.2 and 1.7.3 display the same results graphically.

Table 1.7.2 95% Confidence interval estimation using profile likelihood for the GPD shape parameter and 100-year return level.

PARAMETER	ESTIMATE	95% CONFIDENCE INTERVAL
Shape Parameter (ξ)	0.1868	$(0.1181, 0.2650)$
100-year Return level	5.2207	$(4.24, 6.82)$

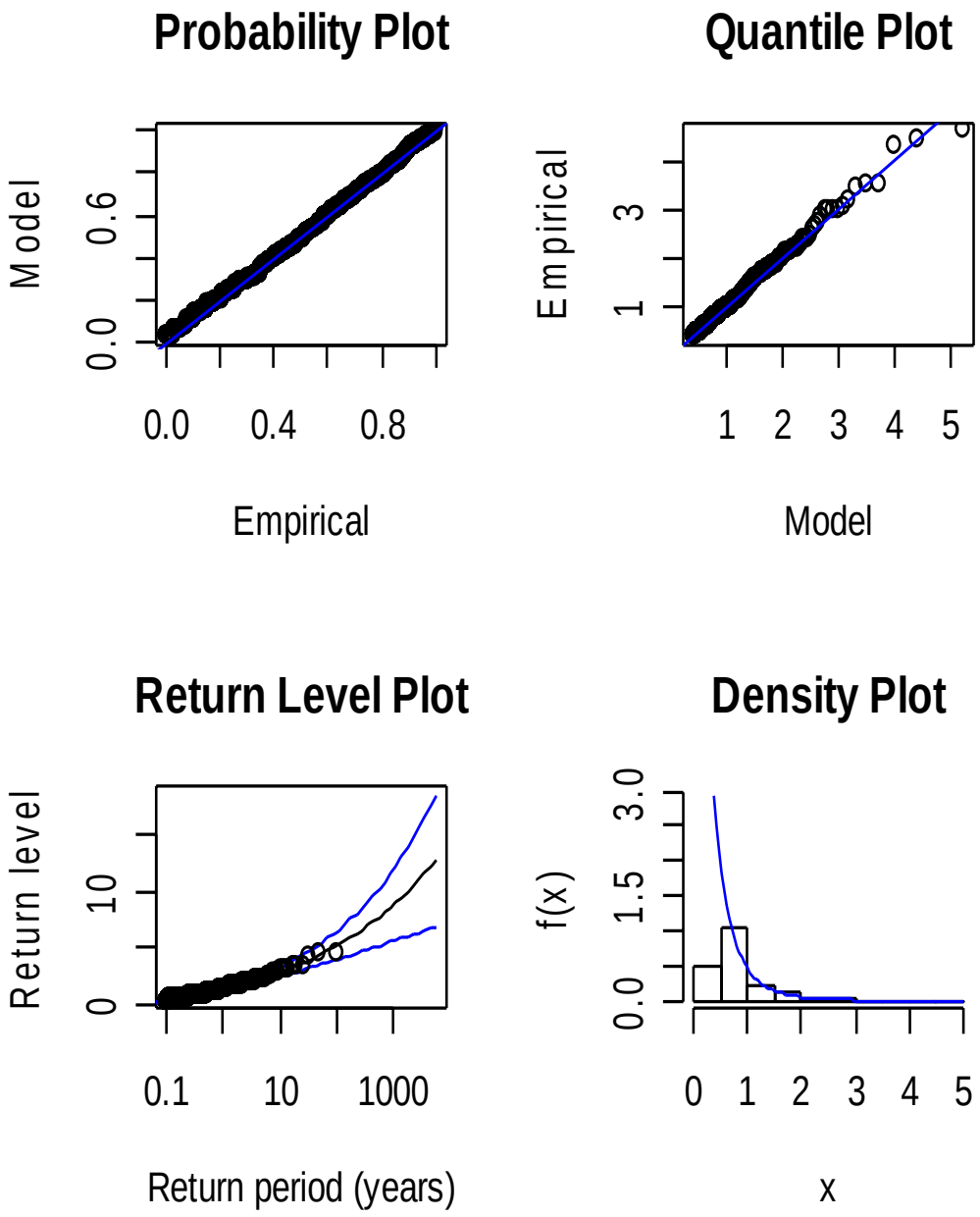


Figure 1.7.1 Diagnostic plots for the GPD fit of the Fort Collins Precipitation Data (threshold value = 0.40)

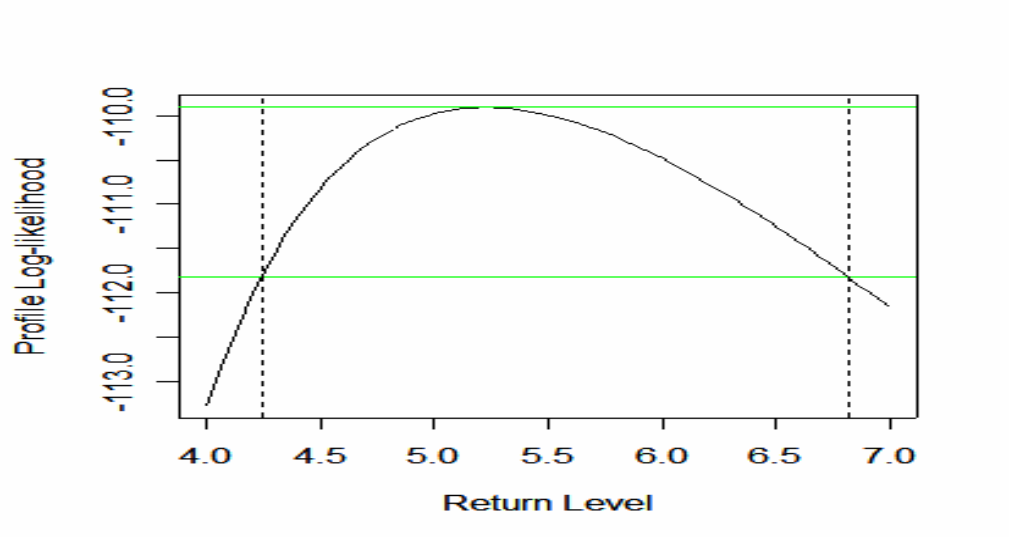


Figure 1.7.2 Profile log-likelihood plot for GPD 100-year return level (inches) for Fort Collins precipitation data.

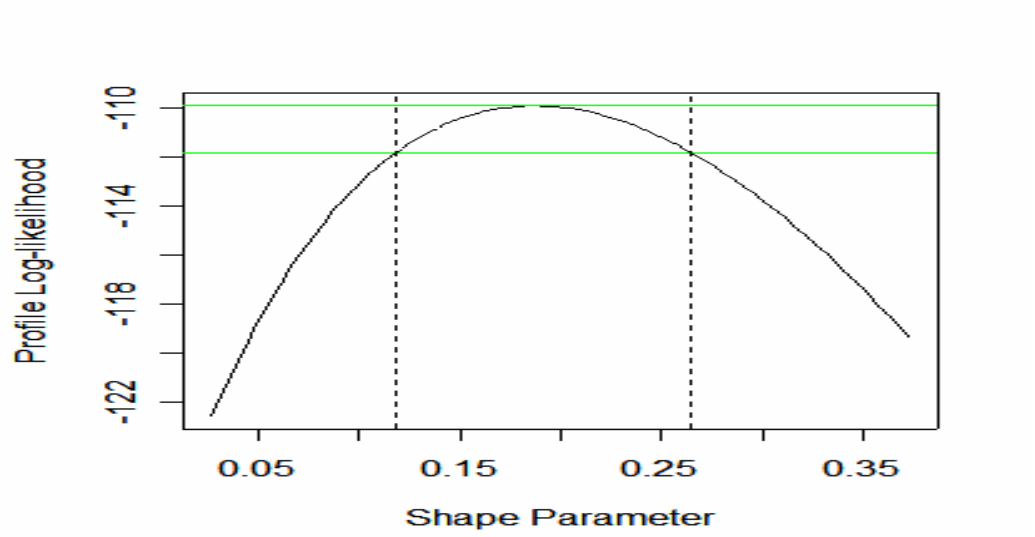


Figure 1.7.3 Profile log-likelihood plot for GPD shape parameter for Fort Collins precipitation data.

This dataset is of special interest due to a flood that occurred on July 28, 1997. The amount recorded for the high precipitation event of July 1997 was 4.63 inches. Note that the confidence interval for the 100-year return level obtained above does include 4.63 inches.

1.8 Aims of the Present Research

The main objective of Chapter 2 is to apply the Generalized Extreme Value (GEV) distribution to statistically model the annual monthly maximum rainfall data for the years 1950 through 2004 for forty four locations, representing central, north and south climatic region of the State of Florida, and use the estimated model parameters for predictive purposes. The time dependence of the rainfall data is incorporated into the GEV model as a non-stationary component in the location parameter. Estimates and confidence intervals are obtained for return levels of return periods of 10, 20, 50, and 100 years, and all 44 locations are classified into clusters based on the similarity profiles.

Chapters 3 and 4 focus on the investigation of the applicability of the theories of a mixture of two extreme value distributions and non-mixture extreme value distributions to statistically model the maximum drug concentration data (C_{max}), a critical pharmacokinetic parameter in the drug development process), the computation of statistical estimates of the model parameters that provide the scientific best prediction of C_{max}, the application of computer algorithms to examine the various model diagnostics and goodness-of-fit tests, and to answer questions regarding the behavior of the data.

In a typical pharmacokinetic experiment, a fixed dose of drug is administered to n subjects, and each subject has blood samples taken at time points $t_1, t_2, ..., t_p$, and $C_{max}(j)$ is the C_{max} observed in subject j. Thus, the data set has n random C_{max} observations.

For the present study, the simulated data values for C_{max} are used with the assumption that these n random C_{max} values are $i.i.d.$ having the extreme value distribution with location, scale and shape parameters.

By careful examination of the histograms of the simulated C_{max} for small samples, a bimodal nature in the distribution of C_{max} was observed. Hence, Chapter 3 focuses on the application of a mixture of two extreme value distributions to statistically model the C_{max} data in small samples. The particular probability density functions used in the mixture model are determined by examining a number of different distributions and evaluating the set of distributions that best fit the data.

Chapter 4 focuses on the application of non-mixture extreme value distributions to statistically model the C_{max} data in large samples. The particular probability distribution functions used are determined by examining a number of extreme value distributions and evaluating the distribution that best fits the data.

An extensive search of the statistical and pharmacological literature has failed to show any research articles relevant to the application of extreme value theory or a mixture of extreme value distributions to statistically model the C_{max} values. This is the major way that the approach presented in this study differs from the pharmacokinetic approach to dealing with C_{max}.

Chapter 5 presents the statistical modeling of a pharmacokinetic system to investigate the drug concentration behavior in an open three-compartment pharmacokinetic model which describes the disposition of an antibiotic drug, coumermycin A_1. We studied a system of random differential equations representing this model. The three rate constants that are used in the system of random differential equations are assumed to have a trivariate distribution. The trivariate statistical distributions that will be used to simulate these rate constants are truncated normal and exponential probability distributions.

More precisely, the numerical solutions for the system will be obtained using the rate constants that are simulated from these distributions for different combinations of covariance structure and initial conditions where the initial conditions are nonrandom.

The effects of different combinations of covariance structures and initials conditions on the deterministic behavior of the drug concentration and the mean behavior of the random solutions as a function of time will be discussed in addition to comparing these two behaviors. Also discussed is the suitability of the use of these two probability distributions to simulate the rate constants.

In Chapter 6 we incorporate the constant time delays into the system of random differential equations considered in Chapter 5, develop extensive numerical solutions to the system of delay random differential equations and study their impact on the rate of decay, absorption and biotransformation of the drug concentration. We begin with the three time delays observed in Chapter 5, that is, a time delay of 3.5 hours for the drug to reach compartment two from compartment one, 6 hours for the drug to reach compartment one from compartment two, and 3 hours for the drug to reach compartment three from compartment two. In addition we consider other sets of constant time delay values to study the behavior of the drug concentration under different time delays.

We discuss the impact of time delays on the overall behavior of the drug concentration in all three compartments, and on the deterministic behavior of the drug concentration and the mean behavior of the random solutions as a function of time. In addition the deterministic behaviors of the drug concentration with and without time delay for each compartment are compared.

Chapter 7 presents the possible extensions of the present research.

Chapter Two

Statistical Modeling of Annual Monthly Maximum Rainfall Using the Generalized Extreme Value Distribution

2.0 Introduction

Events such as extreme rainfalls, earthquakes, hurricanes, and stock market crashes seem to follow no rule, but careful analysis has helped to discover distributions that acceptably model these extreme events [19]. The most important feature of an extreme value analysis (EVA) is the objective to quantify the stochastic behavior of the maximum and the minimum of *i.i.d.* random variables. The EVA requires the estimation of the probability of events that are more extreme than any that have already been observed [23]. Extreme value theory is a unique discipline that develops statistical techniques for describing the unusual phenomena such as extreme rainfalls, floods, wind gusts, earthquakes, risk management, and insurance.

E.J. Gumbel, in his *statistics of extremes* [45, 46] in 1958, discussed many applications that dealt with hydrology or climatology. As part of the global climate change, an accelerated hydrologic cycle including an increase in heavy precipitation is anticipated on a theoretical basis [94, 95], is predicted by numerical models of the climate system [24], and has been detected in observed precipitation [32, 58]. Katz et al. [47] used extreme value distributions to fit to the precipitation data at a single location (Fort Collins, Co, USA) for the time period 1900-1999. A review of applications of extreme value distributions to climate data can be found in Farago and Katz [29]. *An Introduction*

to Statistical Modeling of Extreme Values is a fairly recent book which can serve as an excellent reference [23].

2.1 Focus of the Chapter Two

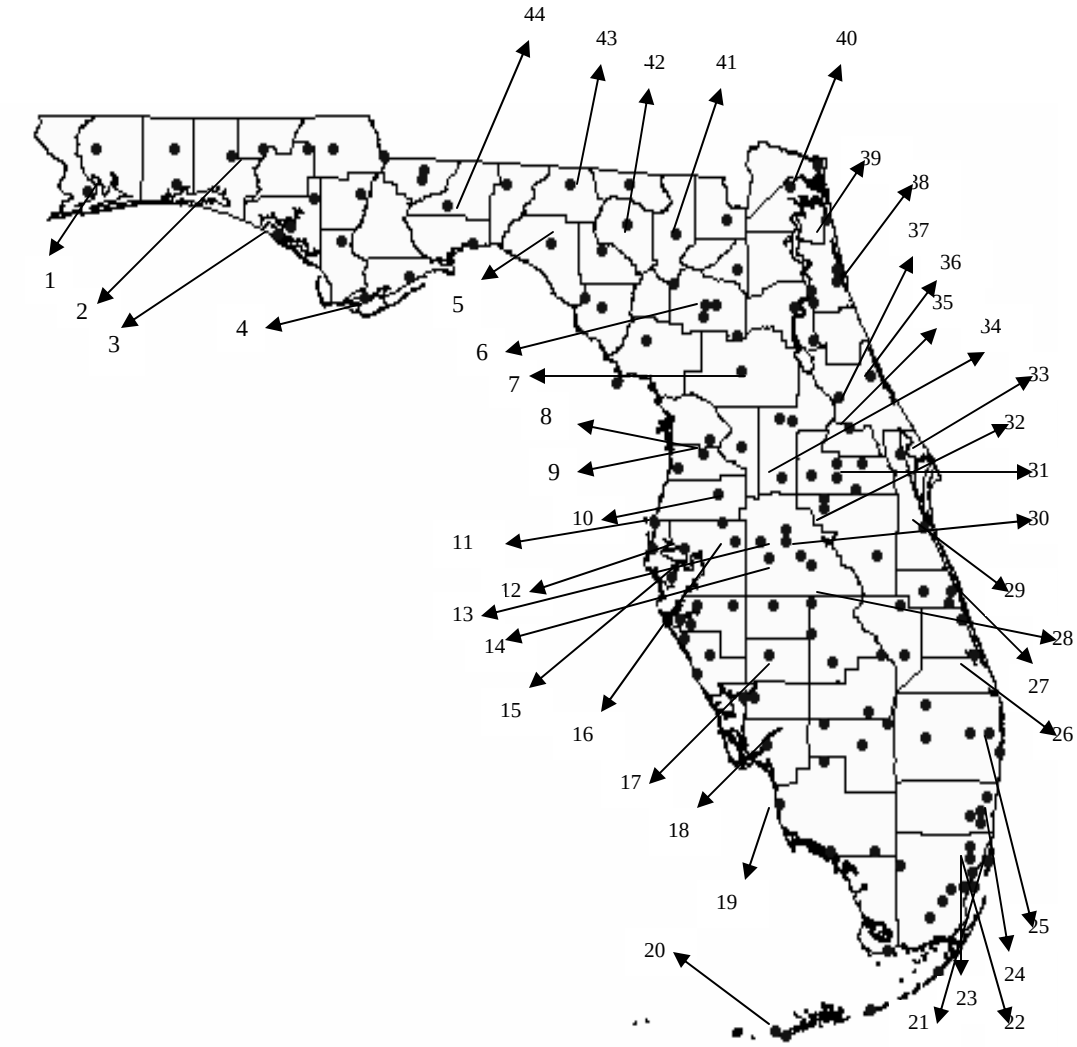
The main focus of this chapter is the application of the theory of Generalized Extreme Value (GEV) distribution to derive statistical models to monitor the rainfall data of forty four locations. These locations are selected from different climatic zones such as central, north and south climatic regions of the State of Florida. The rainfall data is a function of time and thus the subject probability density function is to be adjusted to inherit the time dependence. This is done by adjusting the location parameter (μ) to be time dependent. Thus, our statistical model will incorporate both stationary and non-stationary components to obtain the predictions of median rainfall with profile confidence intervals for each location in addition to providing similarity profiles for rainfall locations and identifying clusters.

2.1.1 Data

The data consist of total monthly rainfall records from 1950 to 2004 for these forty four locations. The data were obtained through two sources: National Oceanic and Atmospheric Administration National Data Center [72] through annual subscription, and Southeast Regional Climate Center [89]. We have used what is known as the block-maxima method to define the extreme rainfall as the maximum of monthly rainfalls within each year. The data for any month in a given year is considered missing if that month has more than five days of data missing.

Table 2.1 Tabulation of the location, climate zone (central, north, or south), years of data, latitude and longitude for all forty four locations

LOCATION	CLIMATE ZONE	YEARS OF DATA	LATITUDE	LONGITUDE
Apalachicola	North	1950 - 2004	29°44'N	85°01'W
Arcadia	South	1950 - 2004	27°13'N	81°52'W
Avon Park	South	1950 - 2004	27°36'N	81°32'W
Bartow	South	1950 - 2004	27°54'N	81°51'W
Brooksville	Central	1950 - 2004	28°37'N	82°22'W
Clermont	South	1950 - 2004	28°27'N	81°45'W
Daytona Beach	Central	1950 - 2004	29°11'N	81°03'W
De Funiak Springs	North	1950 - 2004	30°45'N	86°05'W
Deland	Central	1950 - 2004	29°01'N	81°19'W
Ft. Lauderdale	South	1950 - 2004	26°06'N	80°12'W
Ft. Myers	South	1950 - 2004	26°35'N	81°52'W
Gainesville	Central	1950 - 2004	29°41'N	82°30'W
Hialeah	South	1950 - 2004	25°49'N	80°17'W
Inverness	Central	1950 - 2004	28°48'N	82°19'W
Jacksonville AP	North	1950 - 2004	30°30'N	81°42'W
Jacksonville Beach	North	1950 - 2004	30°17'N	81°24'W
Key West AP	South	1950 - 2004	24°33'N	81°45'W
Kissimmee	Central	1950 - 2004	28°17'N	81°25'W
Lake City	North	1950 - 2004	30°11'N	82°36'W
Lakeland	Central	1950 - 2004	28°01'N	81°55'W
Live Oak	North	1950 - 2004	30°17'N	82°58'W
Madison	North	1950 - 2004	30°27'N	83°25'W
Melbourne	Central	1950 - 2004	28°06'N	80°39'W
Miami Airport	South	1950 - 2004	25°47'N	80°19'W
Miami Beach	South	1950 - 2004	25°47'N	80°08'W
Plant City	Central	1950 - 2004	28°01'N	82°09'W
Saint Augustine	North	1950 - 2004	29°53'N	81°20'W
Saint Leo	Central	1950 - 2004	28°20'N	82°16'W
Saint Petersburg	Central	1950 - 2004	27°46'N	82°38'W
Sanford	Central	1950 - 2004	28°48'N	81°16'W
Stuart	South	1950 - 2004	27°12'N	80°10'W
Tallahassee	North	1950 - 2004	30°24'N	84°21'W
Tampa Airport	Central	1950 - 2004	27°58'N	82°32'W
Tarpon Springs	Central	1950 - 2004	28°09'N	82°45'W
Titusville	Central	1950 - 2004	28°38'N	80°50'W
Vero Beach	Central	1950 - 2004	27°39'N	80°25'W
West Palm Beach	South	1950 - 2004	26°41'N	80°06'W
Winter Haven	Central	1950 - 2004	28°01'N	81°44'W
Naples	South	1950 - 2004	26°10'N	81°43'W
Ocala	Central	1950 - 2004	29°05'N	82°05'W
Orlando	Central	1950 - 2004	28°26'N	81°20'W
Panama City	North	1950 - 2004	30°10'N	85°42'W
Pensacola	North	1950 - 2004	30°29'N	87°11'W
Perry	North	1950 - 2004	30°06'N	83°34'W



1=Pensacola	12=Tampa AP	23=Hialeah	34=Clermont
2=DeFuniak	13=Lakeland	24=Ft. Lauderdale	35=Sanford
3=Panama City	14=Bartow	25=West Palm	36=Daytona Beach
4=Apalachicola	15=St. Petersburg	26=Stuart	37=Deland
5=Perry	16=Plant City	27=Vero Beach	38=Saint Augustine
6=Gainesville	17=Arcadia	28=Avon Park	39=Jacksonville BCH
7=Ocala	18=Ft. Myers	29=Melbourne	40=Jacksonville AP
8=Brooksville	19=Naples	30=Winter Haven	41=Lake City
9=Inverness	20=Key West AP	31=Orlando	42=Live Oak
10=Saint Leo	21=Miami Beach	32=Kissimmee	43= Madison
11=Tarpon Springs	22=Miami Airport	33= itusville	44= Tallahassee

Figure 2.0 Map of the State of Florida [89] identifying all 44 locations that are used in the statistical modeling.

2.1.2 Methodology

To achieve our objective we need to perform a statistical goodness-of-fit test and this requires estimates of the parameters that are inherent in the GEV. To obtain such estimates we utilize the maximum likelihood functional form of the GEV given by

$$L(\mu,\sigma,\xi)=\frac{1}{\sigma^n}\prod_{i=1}^{n}\left(1+\xi\frac{x_i-\mu}{\sigma}\right)^{-\frac{1}{\xi+1}}\times\exp\left\{-\sum_{i=1}^{n}\left(1+\xi\frac{x_i-\mu}{\sigma}\right)^{-\frac{1}{\xi}}\right\} \qquad (2.1)$$

The estimates of μ, σ, and ξ, that is, $\hat{\mu},\hat{\sigma}$, and $\hat{\xi}$, are taken to be those values which maximize the likelihood function L. Furthermore, we calculated the *return level of rainfall*, X_T, for the GEV pdf, given by the following expression

$$x_T=\mu-\frac{\sigma}{\xi}\left[1-\left\{-\log\left(1-\frac{1}{T}\right)\right\}^{-\xi}\right] \qquad (2.2)$$

where T is defined as the *rainfall return period*.

Under the assumption that X_T can be approximated by the χ^2 probability distribution, we can obtain the confidence limits for the return level of rainfall, that is, profile deviance of X_T. The $(1-\alpha)\%$ confidence interval for X_T can be obtained using the following expression

$$\left\{X_T : 2\ln\left(\frac{L(\mu,\sigma,\xi)}{L_p(X_T)}\right)<\chi^2_{1,1-\alpha}\right\} \text{ where } \underset{\sigma,\xi}{L_p(X_T)=\sup L(X_T,\sigma,\xi)}$$

2.1.3 Models

The models fitted to the rainfall data are the GEV model with the following variations in the location parameter, μ:

Model 1: Stationary, where μ, σ, and ξ are constants, implying that the maximum rainfall data is stationary with respect to time.

Model 2: Linear trend in which the location parameter, μ, possesses a linear behavior with time and is numerically approximated with its slope indicating a decrease or increase in rainfall maximum. The scale (σ) and shape (ξ) parameters are constants.

Model 3: Quadratic trend in which the location parameter, μ, is treated as time-dependent with non-linear behavior with time and is numerically approximated. The scale (σ) and shape (ξ) parameters are constants.

When once the most appropriate of the models has been identified, we proceed to estimate numerically the median of the rainfall data using the following analytical expression,

$$\text{Median} = \mu - \frac{\sigma}{\xi}\left\{1 - \left(\log(2)\right)^{-\xi}\right\} \tag{2.3}$$

In addition, we obtain approximate numerical estimates of 90%, 95%, and 99% confidence limits using the following equations, respectively,

$$\left[\hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}}\left\{1 - \left(\log(20)\right)^{-\xi}\right\},\ \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}}\left\{1 - \left(-\log(0.95)\right)^{-\xi}\right\}\right] \tag{2.4}$$

$$\left[\hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}}\left\{1 - \left(\log(40)\right)^{-\xi}\right\},\ \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}}\left\{1 - \left(-\log(0.975)\right)^{-\xi}\right\}\right] \tag{2.5}$$

$$\left[\hat{\mu}-\frac{\hat{\sigma}}{\hat{\xi}}\left\{1-\left(\log(200)\right)^{-\xi}\right\},\ \hat{\mu}-\frac{\hat{\sigma}}{\hat{\xi}}\left\{1-\left(-\log(0.995)\right)^{-\xi}\right\}\right] \tag{2.6}$$

2.2 Results and Discussion

Models 1 through 3 were fitted to the annual maximum monthly rainfall data for each of the forty four locations listed in Table 2.1. The method of maximum likelihood was used.

For Plant City data, the stationary GEV model (Model 1) leads to a maximized log-likelihood of -139.97. The GEV model with a linear trend in μ (Model 2) has a maximized log-likelihood of -137.85, and the GEV model with a quadratic trend in μ (Model 3) resulted in a maximized likelihood of -136.44. The deviance statistic for comparing Model 1 and Model 2 is greater than $\chi^2_{1,0\,95} = 3.84$ suggesting that a model with linear trend is preferable over the stationary model (μ is constant). However, the deviance statistic for comparing Model 1 and Model 3 is greater than $\chi^2_{2,0\,95} = 5.99$. This suggests that the model with a quadratic trend is preferable. The two model diagnostic plots, quantile and the density plots that are given in Figure 2.1 support the fit of the model. Thus, we conclude that the best model for extreme rainfall in Plant City is Model 3.

For Hialeah data, the stationary GEV model (Model 1) leads to a maximized log-likelihood of -157.5. The GEV model with a linear trend in μ (Model 2) has a maximized log-likelihood of -155.34, and the GEV model with a quadratic trend in μ (Model 3) resulted in a maximized likelihood of -155.19. The deviance statistic for comparing Model 1 and Model 2 is greater than $\chi^2_{1,0\,95} = 3.84$ suggesting that a model with linear trend is preferable over the stationary model (μ is constant). However, the deviance

statistic for comparing Model 1 and Model 3 is less than $\chi^2_{2,0\,95} = 5.99$, and for comparing Model 2 and Model 3 is less than $\chi^2_{1,0\,95} = 3.84$. Neither of them is statistically significant. Thus, we conclude that the best model for extreme rainfall in Hialeah is Model 2.

The above analyses were repeated for the remaining forty two locations.

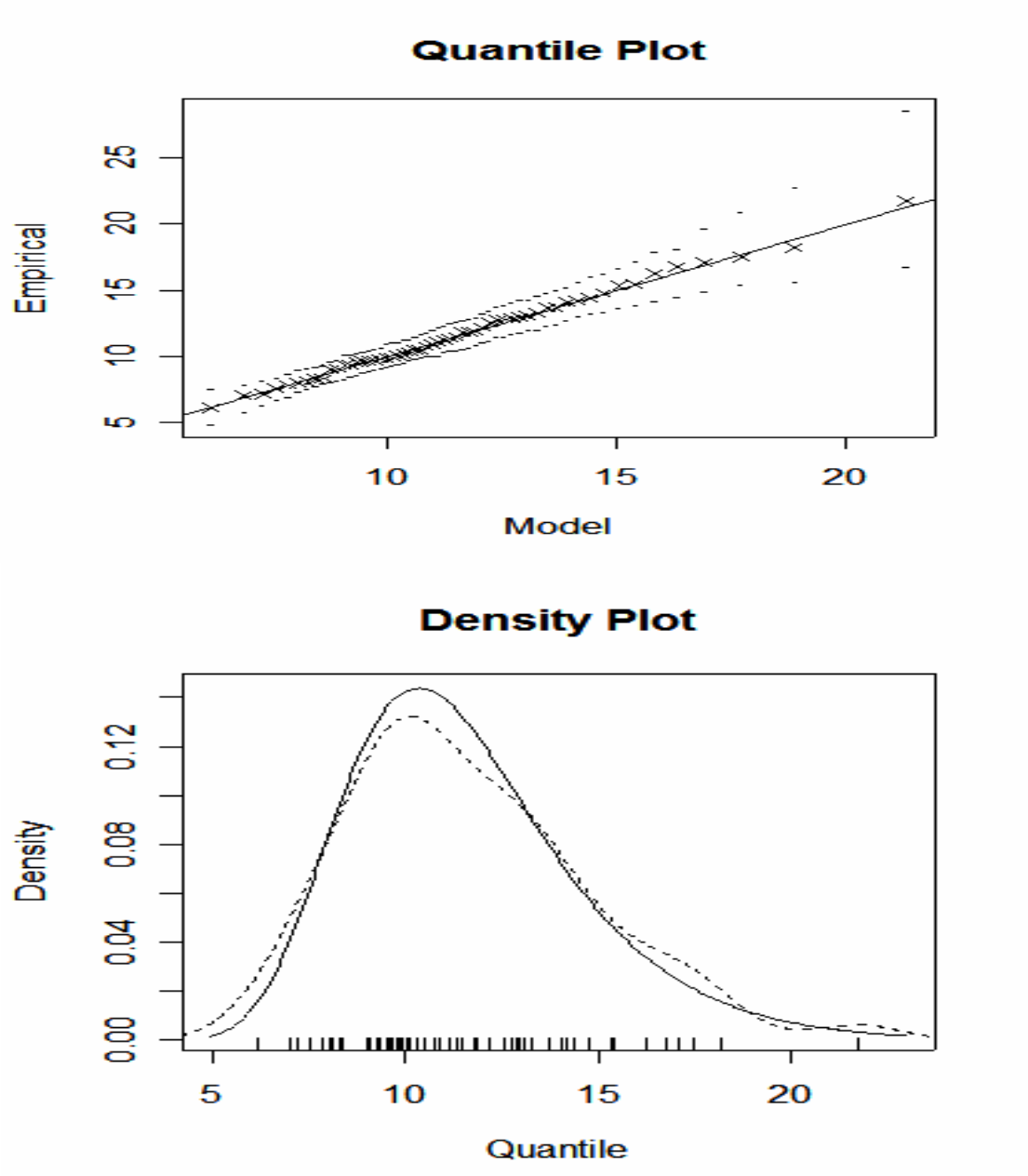


Figure 2.1 Diagnostics plots for Plant City: Quantile and density plots

Table 2.2 Summary of the stationary or non-stationary form of the maximum rainfall data for 56 years of data for the forty four locations

LOCATION	CLIMATE ZONE	TYPE OF NONSTATIONARITY	ESTIMATE OF SLOPE ASSUMING LINEAR TREND
Brooksville	Central	Stationary	0.2354
Daytona Beach	Central	Quadratic	2.2908
Deland	Central	Stationary	2.192
Gainesville	Central	Stationary	-1.2524
Kissimmee	Central	Stationary	3.9379
Lakeland	Central	Linear	5.5988
Melbourne	Central	Stationary	4.4886
Ocala	Central	Stationary	-1.7752
Orlando	Central	Stationary	1.0776
Plant City	Central	Quadratic	4.7798
Saint Leo	Central	Quadratic	5.475
Saint Petersburg	Central	Stationary	-0.5477
Sanford	Central	Stationary	0.148
Tampa Airport	Central	Stationary	0.9241
Tarpon Springs	Central	Linear	4.4139
Titusville	Central	Stationary	-0.8291
Vero Beach	Central	Stationary	4.1174
Winter Haven	Central	Stationary	2.7802
Apalachicola	North	Stationary	-1.773
De Funiak Springs	North	Stationary	1.6623
Jacksonville AP	North	Stationary	0.4221
Jacksonville Beach	North	Stationary	0.9628
Lake City	North	Stationary	1.089
Live Oak	North	Stationary	-1.2083
Madison	North	Stationary	1.3547
Panama City	North	Stationary	0.9075
Pensacola	North	Linear	5.6868
Perry	North	Stationary	0.5131
Saint Augustine	North	Stationary	1.1906
Tallahassee	North	Stationary	2.576
Arcadia	South	Stationary	4.5838
Avon Park	South	Stationary	-0.8828
Bartow	South	Quadratic	4.546
Clermont	South	Stationary	1.5075
Ft. Lauderdale	South	Stationary	5.149
Ft. Myers	South	Stationary	4.3141
Hialeah	South	Linear	5.621
Inverness	South	Stationary	-1.749
Key West AP	South	Stationary	-0.2484
Miami Airport	South	Stationary	3.2126
Miami Beach	South	Stationary	2.3161
Naples	South	Stationary	1.3354
Stuart	South	Stationary	2.173
West Palm Beach	South	Stationary	0.034

Table 2.2 also displays the estimate of the slope assuming the linear trend for the forty four locations. Thirty six locations are identified as stationary. Four locations are identified as non-stationary, and the non-stationarity is due to linear trend. Four locations are identified as non-stationary, and the non-stationarity is due to quadratic trend.

Table 2.3 displays return levels of the extreme rainfall for the return period 10, 20, 50 and 100 years along with the profile deviance of quantiles for a few locations.

Figure 2.2 displays the graphs of the return level X_T for T= 10, 20, 50, and 100 for all forty four locations. From these graphs of return levels, these forty four locations can be grouped into five different clusters. Hialeah, Panama City and Jacksonville Beach form a cluster with Hialeah showing the highest return levels of all forty four stations. Gainesville, Lakeland, Daytona Beach and Madison form another cluster with Gainesville showing the lowest return levels of all forty four locations. West Palm Beach, Miami Airport and St. Petersburg form another cluster. The rest of the locations can be categorized into two large clusters.

Table 2.3 Return levels of the extreme rainfall with profile deviance of quantiles

LOCATION	PERIOD (YEARS)	RETURN LEVELS	90% CONFIDENCE INTERVAL	95% CONFIDENCE INTERVAL	99% CONFIDENCE INTERVAL
Brooksville	10	16.28	(14.9, 18.57)	(14.67, 19.22)	(14.26, 20.79)
	20	18.43	(16.53, 22.19)	(16.27, 23.34)	(15.77, 26.21)
	50	21.34	(18.52, 28.03)	(18.14, 30.2)	(17.52, 35.87)
	100	23.62	(19.87, 33.42)	(19.42, 36.75)	(18.72, 45.72)
Daytona Beach	10	13.28	(13.67, 16.24)	(12.49, 16.65)	(13.16, 17.63)
	20	14.61	(14.98, 18.5)	(13.21, 19.16)	(14.41, 20.79)
	50	16.14	(16.42, 21.63)	(15.2, 22.74)	(15.78, 25.57)
	100	17.16	(17.32, 24.11)	(16.21, 25.67)	(16.62, 29.71)
Gainesville	10	13.53	(12.83, 14.49)	(12.71, 14.75)	(12.47, 15.4)
	20	14.49	(13.71, 15.88)	(13.58, 16.31)	(13.34, 17.39)
	50	15.55	(14.62, 17.69)	(14.48, 18.4)	(14.23, 20.28)
	100	16.23	(15.15, 19.05)	(15, 20.02)	(14.75, 22.66)
Lakeland	10	13.72	(12.71, 15.11)	(12.55, 15.53)	(12.25, 16.56)
	20	14.89	(13.73, 17.09)	(13.56, 17.77)	(13.24, 19.51)
	50	16.26	(14.82, 19.87)	(14.61, 21.04)	(14.27, 24.11)
	100	17.19	(15.45, 22.12)	(15.23, 23.77)	(14.89, 28.23)
Melbourne	10	14.54	(13.54, 15.99)	(13.37, 16.38)	(13.05, 17.31)
	20	15.98	(14.78, 18.11)	(14.59, 18.73)	(14.24, 20.26)
	50	17.67	(16.13, 21.02)	(15.91, 22.06)	(15.53, 24.7)
	100	18.83	(16.97, 23.31)	(16.71, 24.76)	(16.3, 28.51)
St. Petersburg	10	17.65	(16.04, 20.4)	(15.77, 21.2)	(15.28, 23.19)
	20	20.19	(17.98, 24.87)	(17.68, 26.35)	(17.1, 30.13)
	50	23.60	(20.29, 32.19)	(19.85, 35.1)	(19.17, 42.91)
	100	26.25	(21.8, 39.04)	(21.35, 43.62)	(37.91, 56.27)
Vero Beach	10	16.52	(14.98, 19.15)	(14.73, 19.91)	(14.27, 21.84)
	20	18.95	(16.81, 23.41)	(16.51, 24.83)	(15.95, 28.53)
	50	22.31	(19.07, 30.52)	(18.62, 33.36)	(17.92, 41.16)
	100	24.98	(20.62, 37.3)	(20.13, 41.83)	(19.32, 54.72)
Ft. Myers	10	17.11	(16.22, 18.37)	(16.06, 18.74)	(15.75, 19.66)
	20	18.27	(17.31, 20.15)	(17.15, 20.76)	(16.84, 22.32)
	50	19.49	(18.35, 22.45)	(18.19, 23.47)	(17.9, 26.18)
	100	20.23	(18.91, 24.15)	(18.77, 25.55)	(18.49, 29.36)
Lake City	10	14.54	(13.37, 16.33)	(13.18, 16.82)	(12.83, 18.01)
	20	16.44	(14.87, 19.3)	(14.62, 20.15)	(14.2, 22.3)
	50	19.05	(16.78, 24)	(16.47, 25.61)	(15.9, 29.8)
	100	21.13	(18.19, 28.31)	(17.79, 30.74)	(17.13, 37.33)

Table 2.3 (Continued)

Pensacola	10	17.50	(16.27, 20.29)	(16.02, 21.01)	(15.54, 22.81)
	20	19.66	(18.09, 24.14)	(17.81, 25.42)	(17.26, 28.72)
	50	22.38	(20.16, 30.04)	(19.76, 32.45)	(19.16, 38.92)
	100	24.36	(21.44, 35.24)	(21.06, 38.87)	(20.38, 48.97)
Tallahassee	10	16.33	(15.04, 18.41)	(14.83, 18.99)	(14.43, 20.37)
	20	18.33	(16.58, 21.67)	(16.34, 22.66)	(15.87, 25.12)
	50	20.96	(18.41, 26.74)	(18.07, 28.56)	(17.51, 33.26)
	100	22.96	(19.6, 31.26)	(19.16, 34)	(18.55, 41.24)
Ft. Lauderdale	10	18.76	(17.64, 20.27)	(17.44, 20.67)	(17.06, 21.6)
	20	20.32	(19.06, 22.4)	(18.84, 23)	(18.45, 24.48)
	50	22.07	(20.55, 25.13)	(20.34, 26.08)	(19.91, 28.49)
	100	23.20	(21.49, 27.15)	(21.25, 28.41)	(20.82, 31.69)
Hialeah	10	20.15	(18.66, 23.2)	(18.38, 23.99)	(17.85, 25.95)
	20	22.99	(20.78, 27.78)	(20.45, 29.2)	(19.82, 32.84)
	50	26.92	(23.4, 35.15)	(22.93, 37.9)	(22.15, 45.22)
	100	30.08	(25.21, 41.95)	(24.65, 46.21)	(23.78, 57.9)
Naples	10	16.57	(15.52, 18.1)	(15.34, 18.52)	(14.99, 19.52)
	20	18.09	(16.82, 20.34)	(16.61, 21.01)	(16.24, 22.69)
	50	19.90	(18.25, 23.47)	(18.03, 24.6)	(17.61, 27.56)
	100	21.15	(19.18, 25.96)	(18.9, 27.56)	(18.46, 31.82)

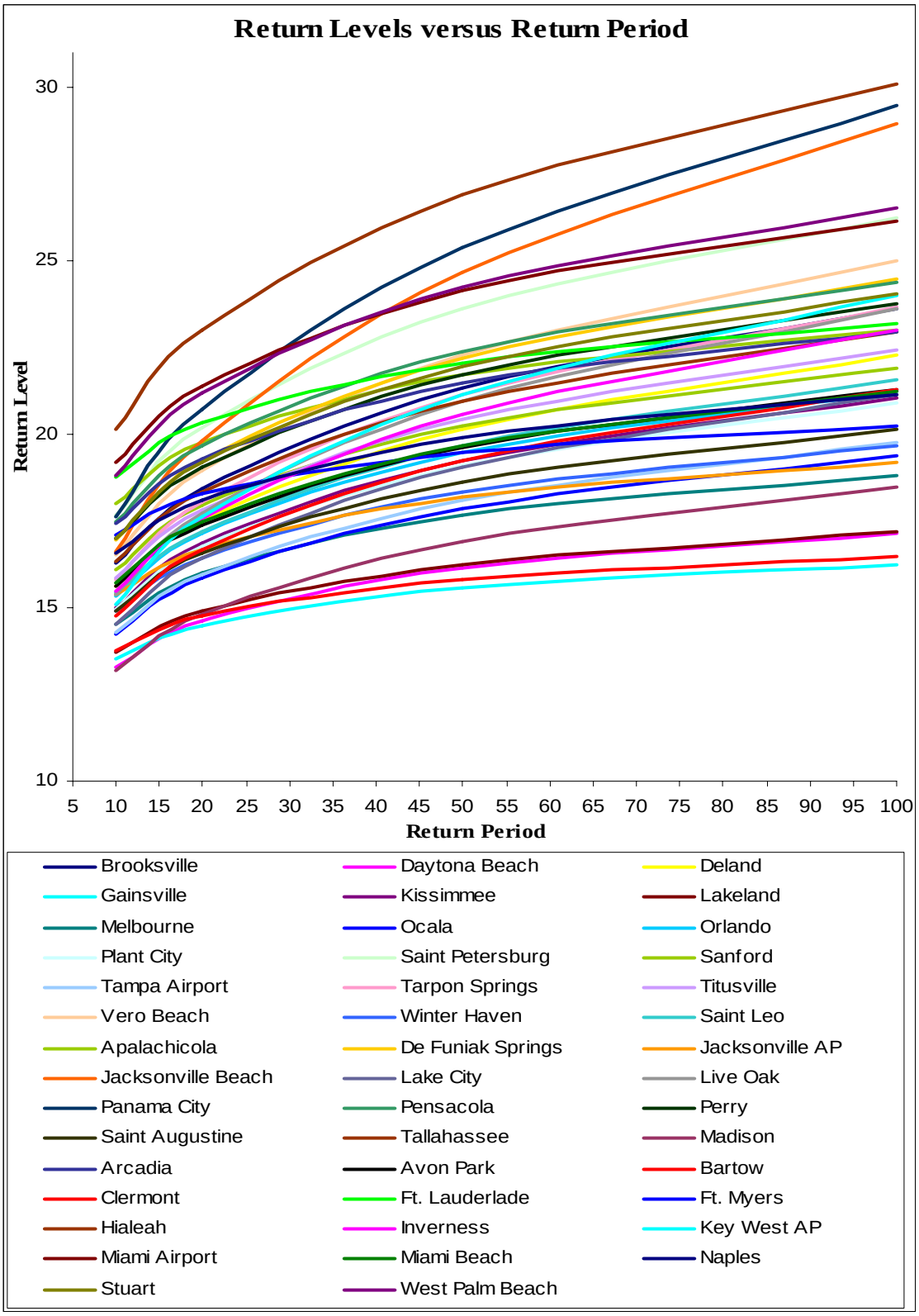


Figure 2.2: Return Level versus Return Periods for all 44 locations

2.3 Similarity Profiles by Each of the Three Climatic Zones

In the central climatic zone 18 locations are selected. From the similarities in the return level profiles shown in the graphs of Figure 2.3A, these 18 locations can be grouped into 5 clusters. Vero Beach and St. Petersburg form one group with St. Petersburg showing the highest return levels of all eighteen locations. Lakeland, Daytona Beach and Gainesville form another cluster with Gainesville showing the lowest return levels of all eighteen locations. While Brooksville and Tarpon Springs form on cluster, another cluster is formed by Winter Haven, Ocala, Tampa Airport and Melbourne. The remaining seven locations turn out to be another cluster.

In the north climatic zone 12 locations are selected. From the similarities in the return level profiles shown in the graphs of Figure 2.3B, these 12 locations can be grouped into 5 clusters. Madison has the lowest return level and seems to stand by itself. While Lake City, St. Augustine and Jacksonville Airport form one cluster, Jacksonville Beach and Panama City form another cluster. The remaining six locations make up another cluster.

In the south climatic zone 14 locations are selected. From the similarities in the return level profiles shown in the graphs of Figure 2.3C, these 14 locations can be grouped into 5 clusters. Hialeah stands by itself and has the highest return level. Bartow stands by itself and has the lowest return level. While Miami Airport and West Palm Beach form a cluster, Naples, Miami Beach, Avon Park, Ft. Myers and Clermont make another cluster. The remaining five locations make up another cluster.

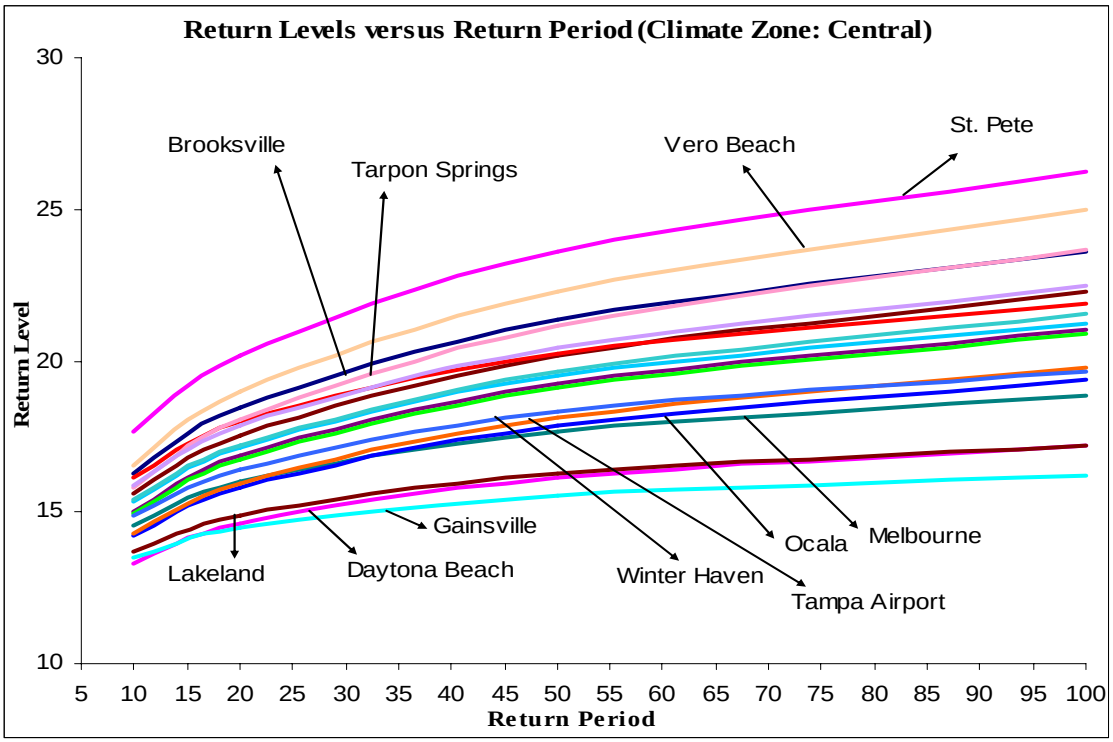


Figure 2.3A Return level profiles for all 18 locations in Central Climatic Zone

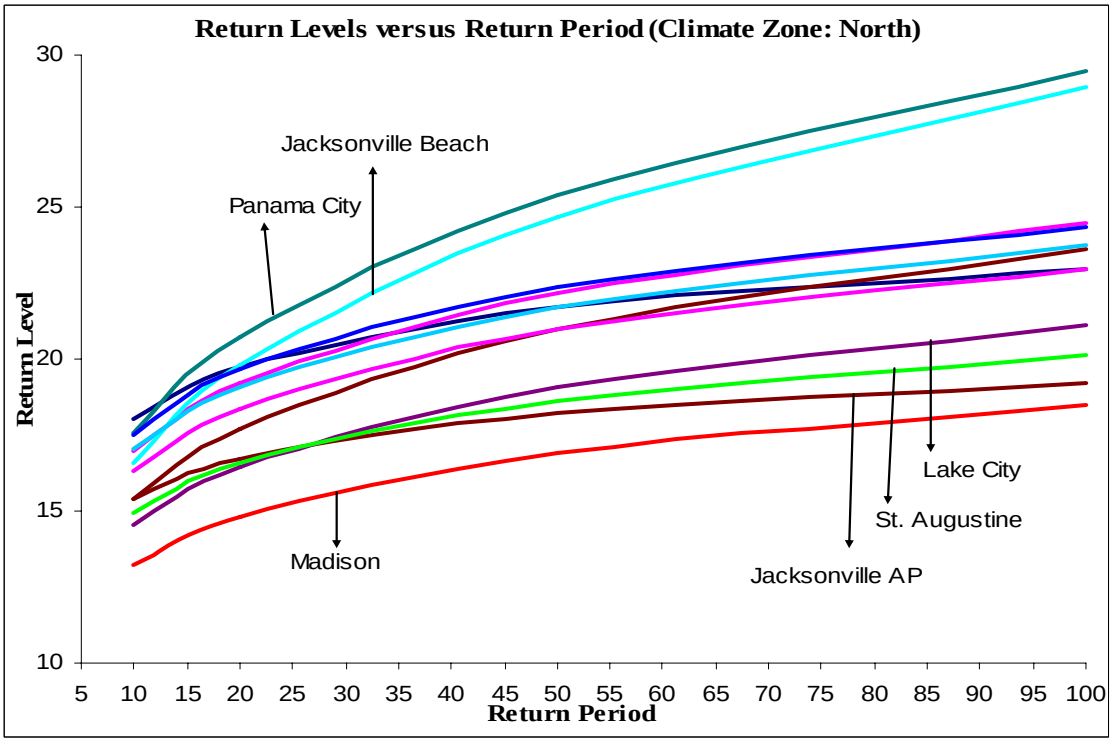


Figure 2.3B Return level profiles for all 12 locations in North Climatic Zone

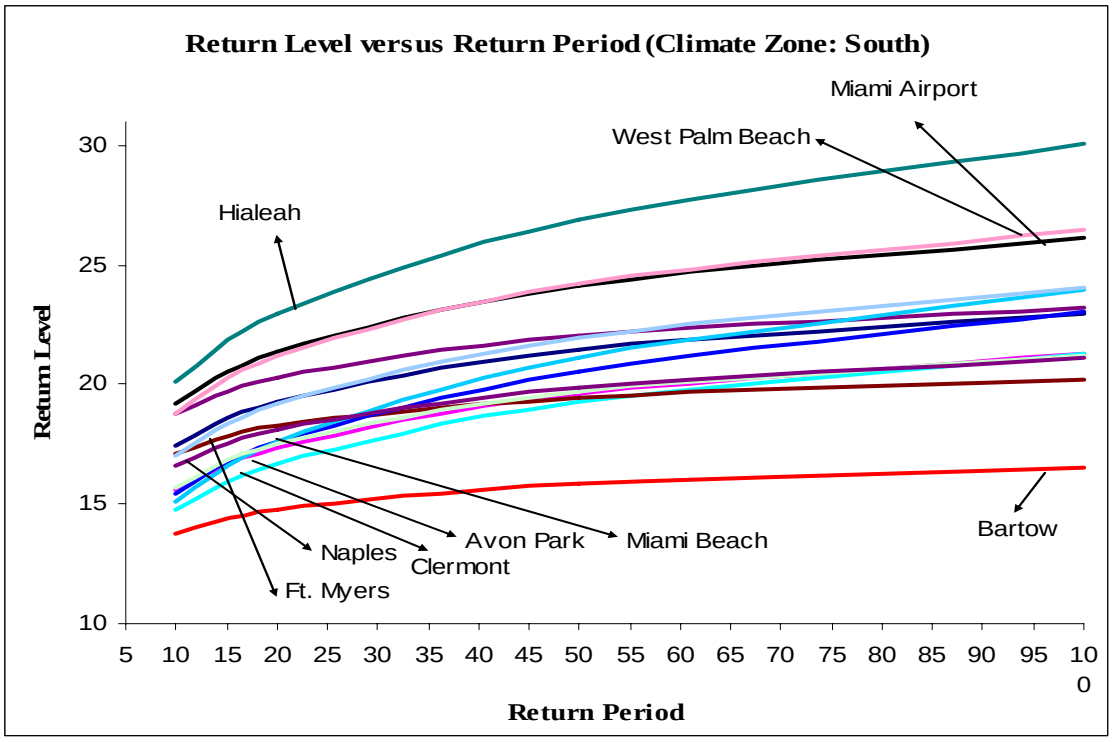


Figure 2.3C Return level profiles for all 14 locations in South Climatic Zone

The plots of 10-year, 20-year, 50-year, and 100-year return levels along with the 95% profile likelihood confidence intervals for all forty four locations provide further insight into the similarities in these profiles. Figures 2.3D through 2.3G provide these plots for 10-year, 20-year, 50-year, and 100-year return levels, respectively

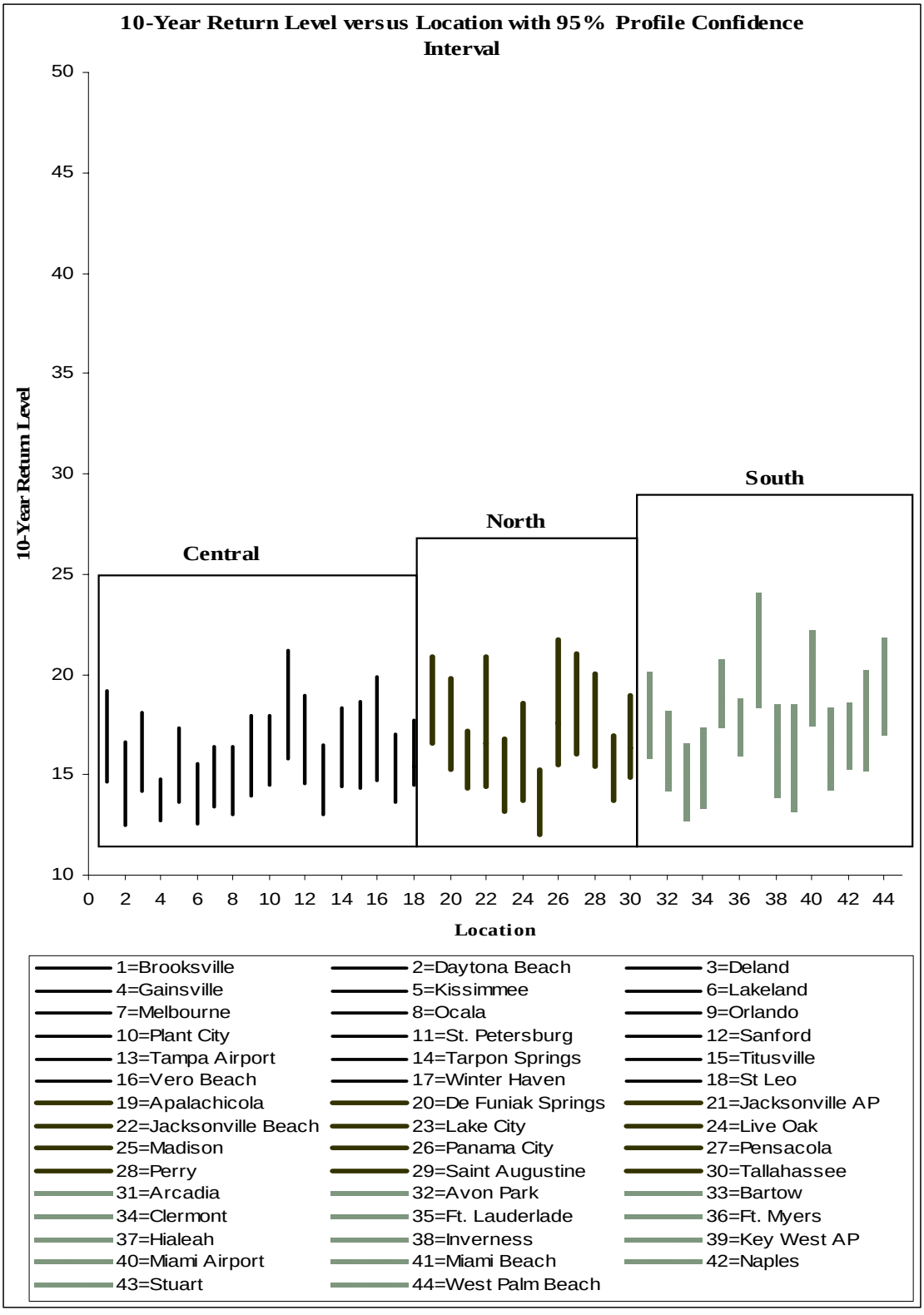


Figure 2.3D 10-Year return level versus location with 95% CI for all 44 locations

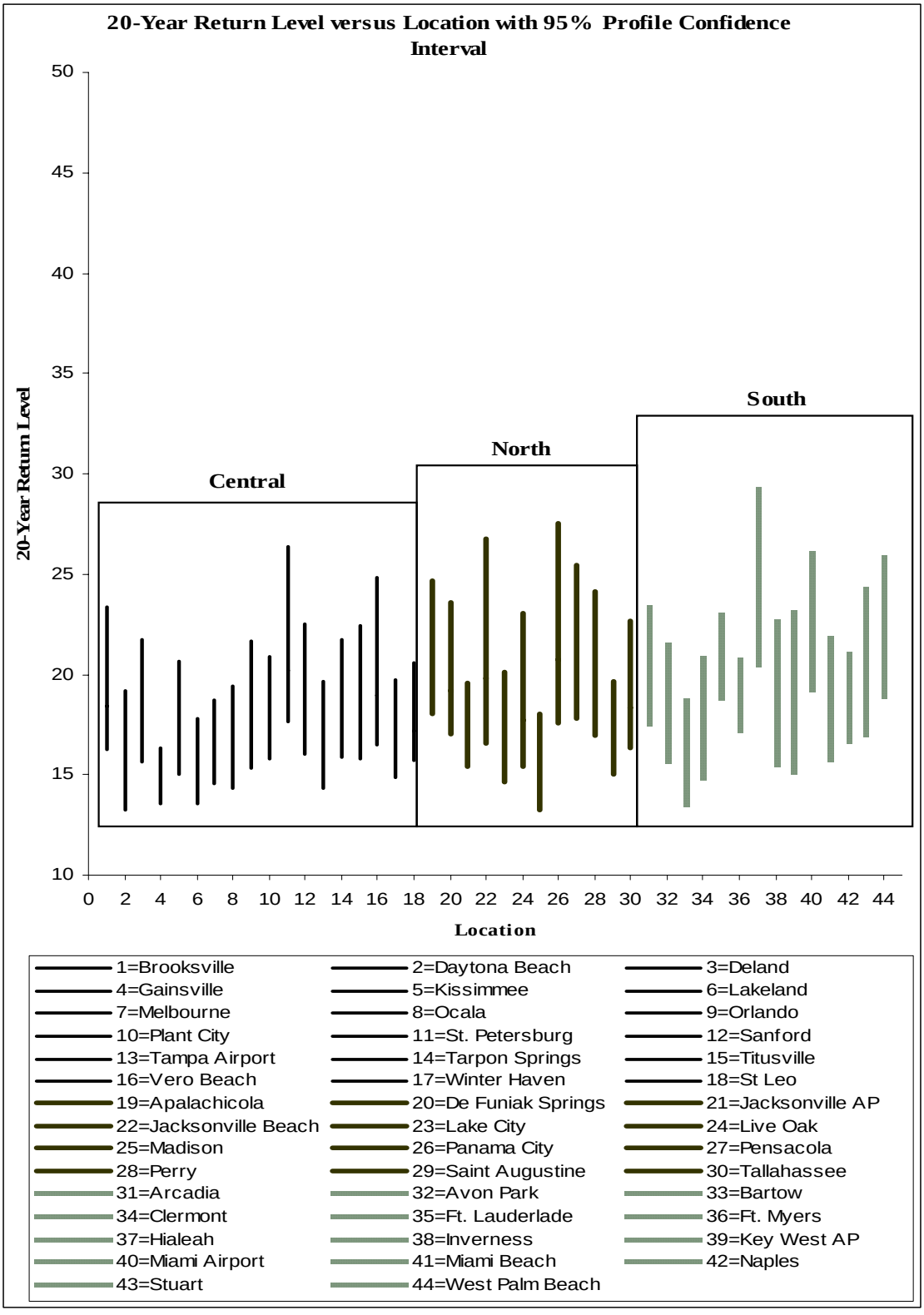


Figure 2.3E 20-Year return level versus location with 95% CI for all 44 locations.

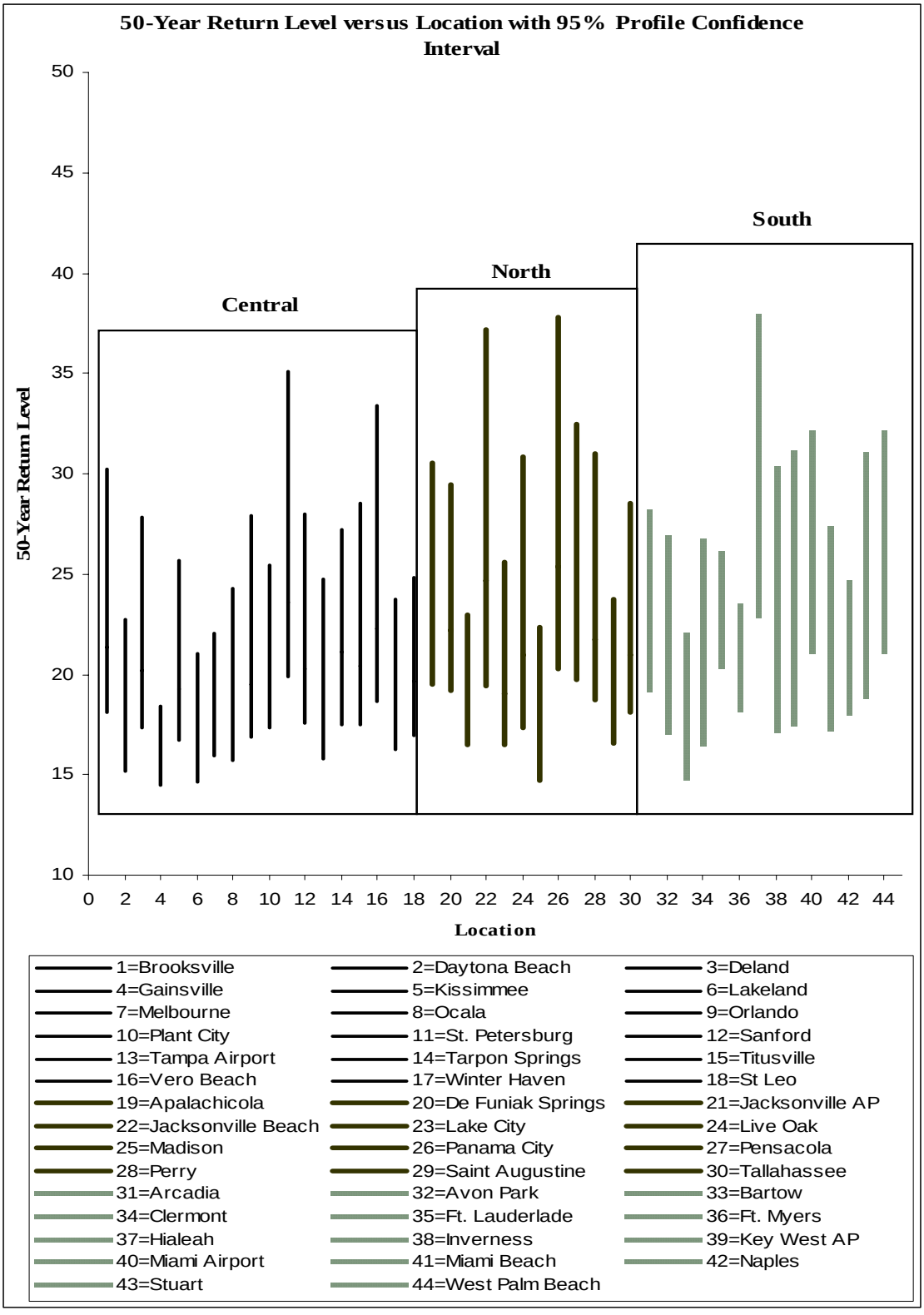


Figure 2.3F 50-Year return level versus location with 95% CI for all 44 locations

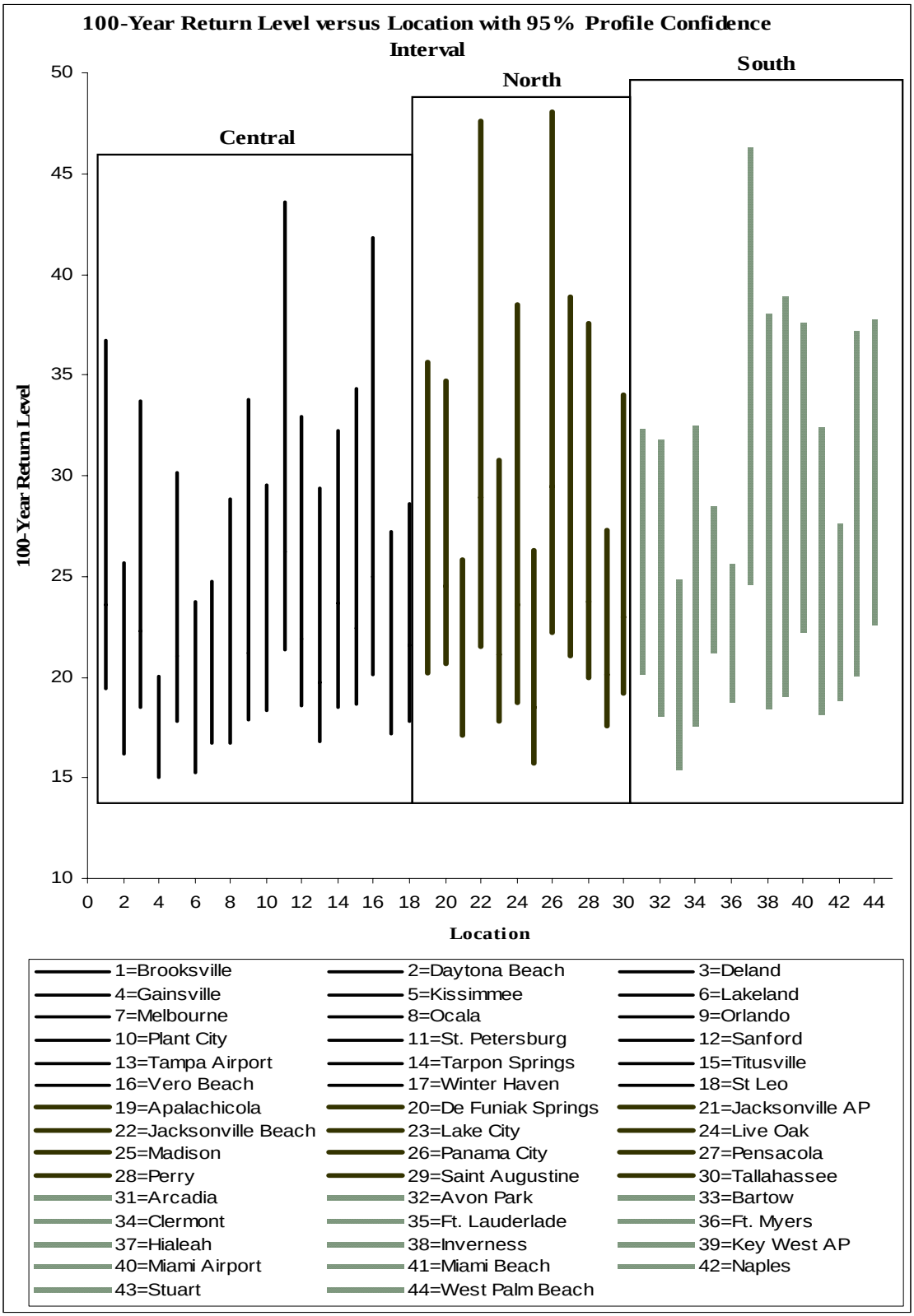


Figure 2.3G 100-Year return level versus location with 95% CI for all 44 locations.

2.4 Conclusion

We have modeled the annual maximum monthly rainfall data for forty four locations from the central, north and south climatic zones in the State of Florida using the Generalized Extreme Value distribution. Since the rainfall data is a function of time, we utilized the time-dependence structure in the location parameter to build non-stationary model with the non-stationarity in the form of a linear trend as well as a quadratic trend. Extreme rainfalls for some locations do display an evidence of non-stationarity in the form of either a linear trend or a quadratic trend. The return level estimates can be used to determine ways to protect people and the property by taking appropriate measures against extreme rainfalls that may lead to floods and other water-related natural disasters. Based on the magnitude of these estimates, the forty four locations can be classified into five clusters as described in the results and discussion section.

Chapter Three

Mixed Statistical Model for the Pharmacokinetic Parameter, Maximum Drug Concentration(C_{max}): Small Samples

3.0 Introduction

The maximum drug concentration in blood or plasma after the administration of a certain drug in an individual is considered as an extreme value. Extreme maximum drug concentration levels are associated with undesirable clinical outcomes such as toxicity, especially in cases of those drugs that have narrow therapeutic windows. A drug is identified as having a narrow therapeutic index if small changes in systemic concentration can lead to either sub-therapeutic or toxic results in patients [9, 67]. Warfarin is among those drugs considered to have a narrow therapeutic index. As these maximum drug concentration levels vary in patients due to varying factors, such as age, body weight, genetic polymorphism and drug-drug interactions, analysis of such data using appropriate statistical techniques is crucial in understanding the behavior of the drug [52].

Thus, the purpose of the present study is to formulate a necessary theory that includes the computation of statistical estimates of the model parameters that provide the scientific best prediction of maximum drug concentration, and application of algorithms to answer key questions concerning the behavior of the data.

3.1 Outline of Each Section of Chapter III

Section 3.2 introduces the concept of pharmacokinetics, bioavailability and its assessment. In addition, it discusses Warfarin drug, its application for the prevention and treatment of thromboembolic disorders, side effects, and interaction with other drugs. The section ends with the description of the simulation of the maximum drug concentration (C_{max}) data. Section 3.3 presents the focus of chapter three.

Section 3.4 introduces the statistical distributions that are considered to model the bimodal data.

Section 3.5 presents a brief introduction to the concept of a mixture of two extreme value distributions. This section also presents an analytical basis for not using the mixture of two Gumbel models using a simulation study. It further discusses the mixture models for Pareto and Weibull, and the derivations of the normal equations.

Section 3.6 has a detailed discussion of the results of the analysis on small samples from the mixture models. Section 3.7 presents the conclusion.

3.2 Pharmacokinetics

Pharmacokinetics is a mathematical quantification of the amounts of drug in the body over time. It is the study of the processes of bodily absorption, distribution, metabolism, and excretion of a drug [36, 38]. The pharmacokinetic parameters required in the Food and Drug Administration regulations for an in-vivo bioavailability study are the maximum drug concentration, C_{max}, the area under the plasma or blood concentration-time curve, AUC_0^{∞}, time to achieve the maximum concentration, T_{max}, elimination half-life, $t_{1/2}$, and the rate constant, k_e, of the therapeutic moiety [30].

3.2.1 Bioavailability and Assessment of Bioavailability

Bioavailability

According to the Food and Drug Administration guidelines, bioavailability is defined as *the rate and extent to which the active ingredient or active moiety is absorbed from a drug product and becomes available at the site of action.* The maximum drug concentration is a function of both the *rate* and *extent of absorption* [30].

Before a medication can have any effect on the body, it must enter the bloodstream. If the medication is administered intravenously, then 100% of the medication is free in the bloodstream immediately after dosing. The simulated data that are used in this chapter are from a medication that is administered orally.

Bioavailability of orally administered drugs depends on absorption from the gastrointestinal tract (GI) and few factors that affect this process are the enzymes in the gastrointestinal tract that metabolize the drug and thus prevent certain percentage of it from getting into the bloodstream; first-pass metabolism of the drug after it is absorbed into the blood; and some portion of the drug may bind with protein in the plasma and hence may not be free in the bloodstream for the body to use it.

Assessment of Bioavailability

One of the three main pharmacokinetic parameters that are used to assess the bioavailability is $C_{\max}$, the maximum drug concentration. The $C_{\max}$ increases with an increase in the dose as well as with an increase in the absorption rate. The $T_{\max}$, the time at which the $C_{\max}$ occurs reflects the rate of drug absorption, and decreases as the absorption rate increases [30].

3.2.2 Introduction to Warfarin

Warfarin (Brand name: Coumadin) is the most widely prescribed anticoagulant drug for the prevention and treatment of thromboembolic disorders. Because of large interpatient variability in dose-anticoagulant effect relationship and a narrow therapeutic index, dosing is a challenge. The goal is to minimize the risk of serious bleeding events without compromising the anticoagulant effect. Haemorrhage is the major complication (~ 5 - 30% of patients). Haematuria, epistaxis, uterine bleeding, petechial, or simple bruising are common side effects but GI, intra-cranial, and retroperitoneal bleeding may also occur [52]. An enormous variation in pharmacological response is a characteristic feature of the clinical use of Warfarin. The average patient requires a daily dose of about 5 mg to maintain blood hypocoagulability but dosage requirements can vary from 1 - 25% daily. Differences in both pharmacokinetics and pharmacodynamics contribute to this variation. Some well documented factors that influence the response to Warfarin are shown below.

Factors influencing the response to Warfarin

Factors	Effect	Mechanism
Age	↑ with increasing age	Enhanced receptor site sensitivity
Pregnancy	↑	Increased blood coagulability
Liver disease	↓	Defective synthesis of clotting factors
Heart failure	↓	Reduced clotting factor synthesis
Hyperthyroidism	↓	Increased clotting factor degradation
Concomitant drugs	↑↓	Induced or inhibited metabolism

Warfarin shows a bimodal plasma concentration distribution. In other words, the same dose of drug gives high levels in some patients, low levels in others. Warfarin is eliminated almost entirely (> 99%) by metabolism. Cytochrome P-450 (CYP2C9) is the enzyme primarily responsible for the metabolism of warfarin and it has been suggested that genetic variation in the gene coding for this enzyme contributes significantly to the large inter-patient variability in warfarin-dose requirements. Of all the covariates tested, CYP2C9 genotype was the only one identified as having a significant effect, with a 46% (14%) reduction in clearance for warfarin as compared to wild-type [4]. Patient's nutritional status is another factor to be considered before the Warfarin therapy. Those who are malnourished should receive lower doses of Warfarin because they probably have low vitamin K intake and decreased serum albumin concentrations. Women generally require lower doses than men, and thus, gender is another factor. Major pharmacokinetic drug-drug interactions can occur with warfarin and drugs that interfere with the metabolism of warfarin resulting in variable plasma concentrations of warfarin. Examples of major pharmacokinetics drug interactions with warfarin are shown below [52].

Major Pharmacokinetic drug interactions with warfarin

Drug	Effect	Mechanism
Cholestyramine	↓	Decreased absorption
Amiodarone	↑	Decreased metabolism
Co-trimoxazole and Cimetidine	↑	Decreased metabolism
Barbiturates, carbamazepine, Rifampicin	↓	Increased metabolism

A system, commonly called the International Normalized Ratio (INR), was established by the World Health Organization and the International Committee on Thrombosis and Hemostasis for reporting the results of blood coagulation tests. All results are standardized using the International Sensitivity Index (ISI) for the particular thromboplastin reagent and instrument combination utilized to perform the test. For example, a person taking the anticoagulant warfarin might optimally maintain a Prothrombin Time (PT) of 2 to 3 INR. The purpose of administering warfarin is to give enough dosage so as to achieve the patient's PT of 2 to 3 INR. The INR is the ratio of a patient's PT to a normal (control) sample, raised to the power of the ISI [98]. That is,

$$INR = \left(\frac{PT_{test}}{PT_{normal}} \right)^{ISI}$$

Knowledge of the pharmacokinetics of warfarin is helpful in understanding the initial response to therapy. Warfarin can be detected in the plasma one hour after oral administration, and peak concentrations occur within two to eight hours of administration.

3.2.3 Maximum Drug Concentration (C_{max}) Data

A Monte Carlo simulation method was used to generate maximum drug concentration data. Monte Carlo simulation relies on pseudo random numbers to generate random times to C_{max} based on C_{max}. The analysis tool pack of Microsoft Excel is utilized to perform the simulation. The simulation was based on an assumption that the distribution of logarithm of pharmacokinetic parameter (that is, $\ln[C_{max}]$) is normally distributed. The $\ln[C_{max}]$ of warfarin were used in the simulation. In this case, random simulation of $\ln[C_{max}]$ was performed using published mean and coefficient of variation of $\ln[C_{max}]$. The obtained $\ln[C_{max}]$ was transformed to C_{max}, anti-log value. As a result,

we got the random simulated C_{max} values. The sample size of this simulated data set is 1000. The random samples of varying sample sizes, ranging from 50 to 500, were generated. The histograms of the data for sample sizes n = 50, n = 100, n = 200, n = 300, n = 400, n = 500 were examined.

3.3 Focus of the Chapter Three

For sample size $N \leq 100$, Warfarin C_{max} shows a bimodal distribution. In other words, the same dose of drug gives high levels in some patients, low levels in others. Because of the bimodal nature in the maximum drug concentration data when sample size is small, a mixture of two extreme value distributions is more appropriate in adequately modeling such bimodal data.

Thus, the focus of the present chapter is the statistical modeling of the simulated C_{max} , the maximum drug concentration data for warfarin, for small samples using a mixture of two extreme value distributions. The particular probability density functions used in the mixture model are determined by examining a number of different distributions and evaluating the set of distributions that best fit the data. In each of these distributions the maximum likelihood methods were used to estimate the model parameters and their standard errors. The criterion used to select the model is the Akaike Information Criterion (AIC). The AIC for a model is computed as $-2\ln L + 2M$, where M is the number of parameters estimated for the model, and L is the maximized likelihood for that model. The smaller the value of AIC the better is the fit of the model to the observed data. These models are examined below in separate sections, followed by an overall summary of results.

3.4 Statistical Methods for Small Samples

3.4.1 Generalized Extreme Value and Gumbel Distributions

For small sample of sizes 50 and 100, the Gumbel distribution was fit. The p-value for all three goodness-of-fit tests (Chi-square, Anderson-Darling, and Kolmogorov-Smirnoff) was non-significant (> 0.10), leading to the failure of rejecting the null hypothesis that the shape parameter is zero. The same conclusion was drawn based on the comparison of the value of Akaike Information Criterion (AIC) from this model to the value of AIC from the generalized extreme value model. But a non-mixture model with the shape parameter being zero is not a suitable model to fit to the maximum drug concentration data that exhibit bimodality in small samples. This leads to the exploration of a mixture model to adequately model such data.

3.4.2 Weibull Distribution

Although the three-parameter Weibull model seems to fit the maximum drug concentration data for small samples based on the χ^2 goodness-of-fit test, the bimodal nature of the data are better fit by a mixture model. The two-parameter Weibull does not provide a good fit of the data. This was assessed by using the likelihood ratio test, and also through Anderson-Darling and Cramer-von Mises goodness-of-fit tests. The following probability plots display the absence of linearity in the graph suggesting the inappropriateness of the 2-parameter non-mixture Weibull model. The points on the plot could be modeled by two straight lines of different slopes. This may be an indicative of a bimodal data set.

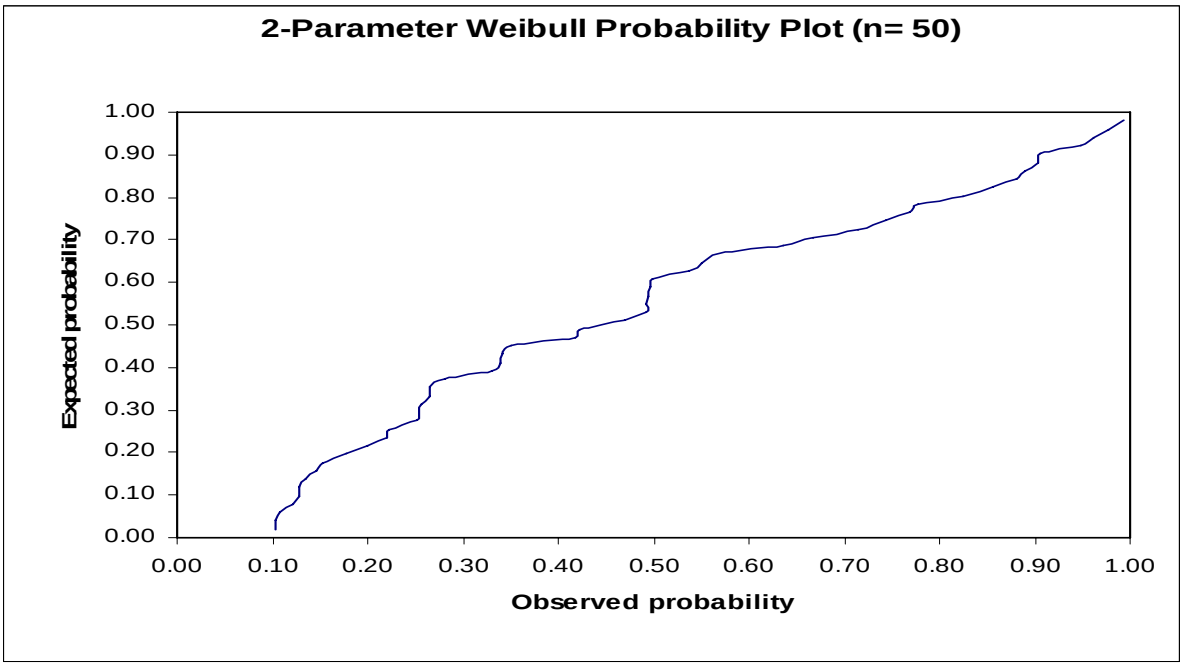


Figure 3.1 Probability Plot from the 2-parameter Weibull fit for n = 50.

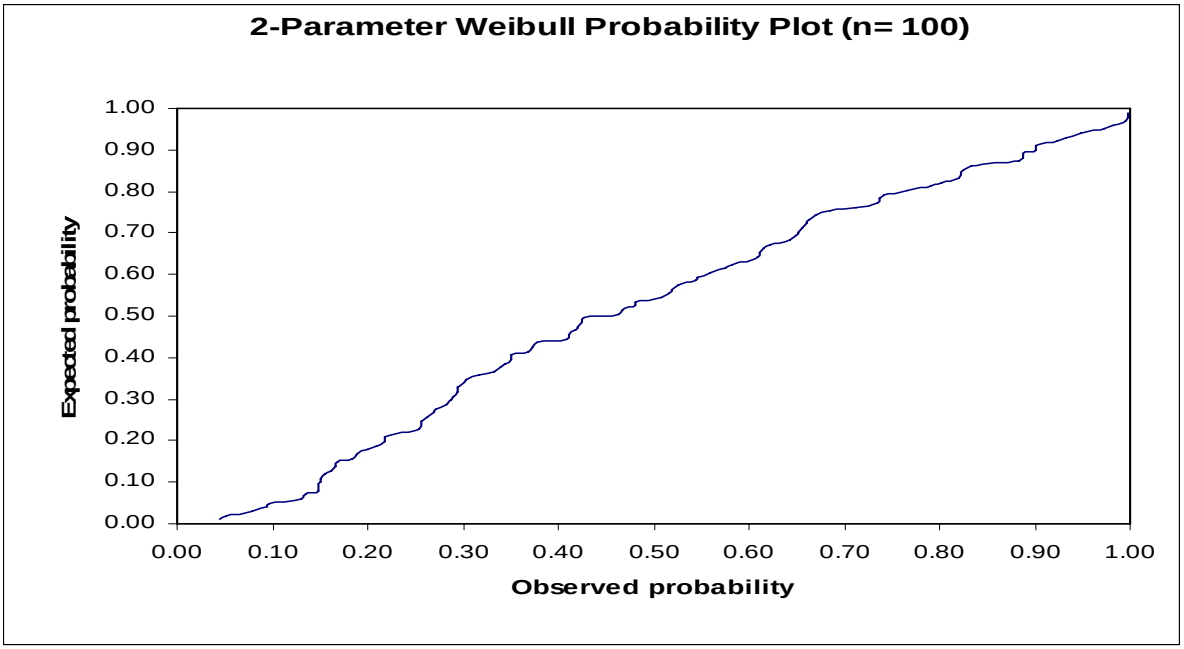


Figure 3.2 Probability Plot from the 2-parameter Weibull fit for n = 100.

Hence, the modeling of the data for small samples of size 50 and 100 was carried out using the mixture of two Weibull extreme value distributions (for which the location parameter μ = 0) with a mixing parameter $p(0 < p < 1)$.

3.4.3 Pareto Distribution

The Pareto distribution was not a good fit to the maximum concentration data for small samples. This was assessed by using the χ^2 goodness-of-fit test. The probability plots shown below lack the linearity in their appearances. Hence, a mixture of two Pareto distributions will be examined to measure its adequacy to model the bimodal data when sample size is small.

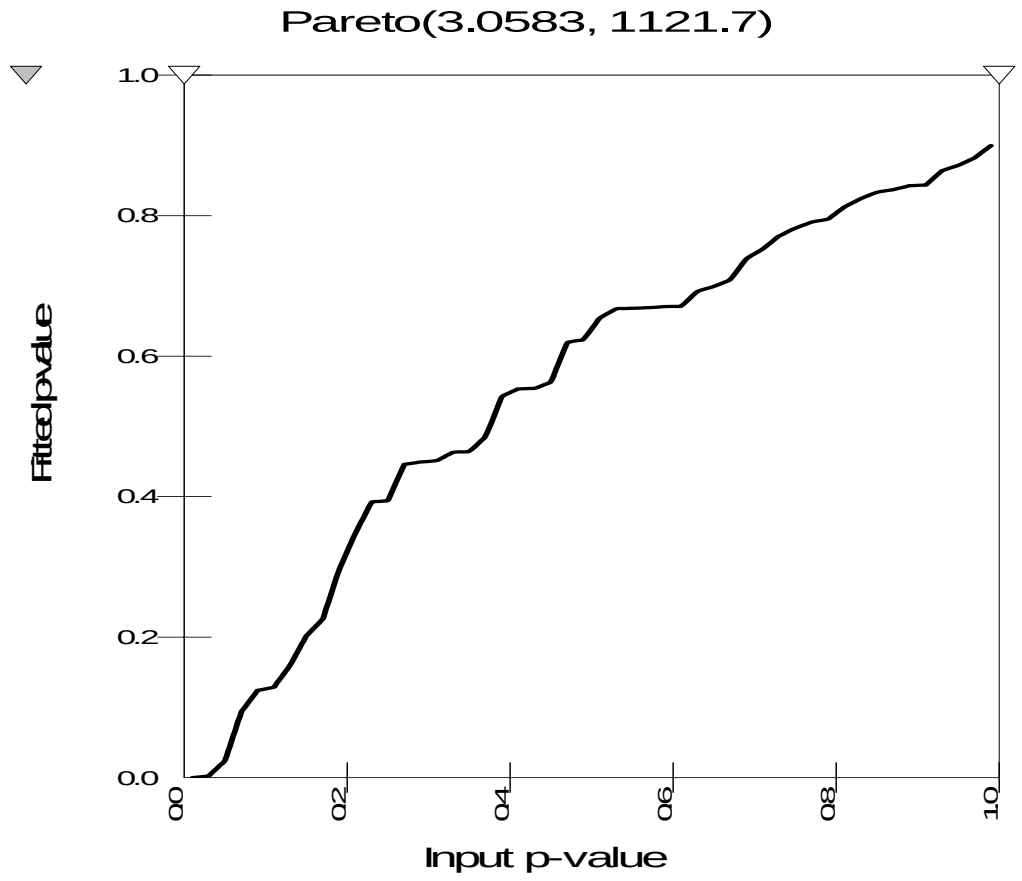


Figure 3.3 Probability plot from the Pareto fit for n = 50.

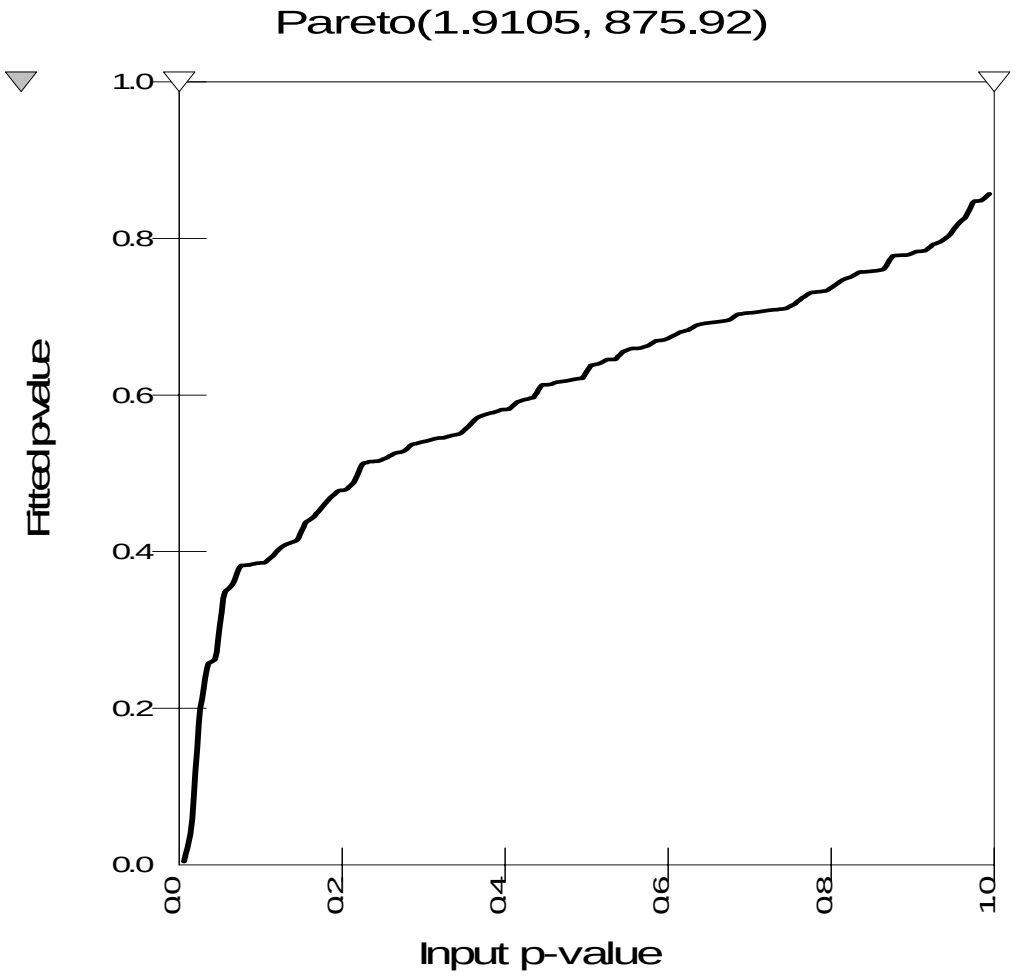


Figure 3.4 Probability plot from the Pareto fit for n = 100.

3.5 Mixture of Two Extreme Value Distributions

3.5.1 Introduction

Because of the bi-modal nature of the maximum concentration data, these observations were used to estimate the parameters of a two-component mixture model. The two-component mixture model has the general form [61],

$$f(x;\alpha,\beta,p) = pf_1(x;\alpha) + (1-p)f_2(x;\beta), \tag{3.1}$$

where $f_1(x;\alpha)$ is the probability density function with parameters α, for one subgroup, and $f_2(x;\beta)$ is the probability density function with parameters β for the other group. The mixing parameter $p(0 \le p \le 1)$ is the fraction of observations in one group and $(1-p)$ is the rest of the observations in the other group [61].

Suppose p is known: In a sample of *n* one would expect approximately np observations from the first component in the above probability density function and $n(1-p)$ observations from the second. If the two distributions are fairly well separated and p is not too close to 0 or 1, the smallest observations will be from one component and the largest observations are from the other. Then, one can treat the k_1 smallest observations as the first k_1 in a sample of size np from the first distribution, and the k_2 largest as the largest observations in a sample of size $n(1-p)$ from the second distribution. We can estimate the parameters of the two distributions separately from these two parts of the data.

Suppose p is unknown: Being unsure about how sensitive the procedure is to the exact value of *p* being used, it is desirable to estimate *p* from the mixture model using different start values for *p* so as to examine the consistency in the estimation of the model parameters and their standard errors

The particular probability density functions used for $f_1(x;\alpha)$ and $f_2(x;\beta)$ were selected after examining the extreme value type I (or Gumbel), Pareto, and Weibull distributions for each component of the model.

3.5.2 Mixture of Two Gumbel Distributions

Using the probability density function for Gumbel, equation (3.1) can now be written as,

$$f_i(x_i;\mu_1,\mu_2,\sigma_1,\sigma_2,p) = p\left(\frac{1}{\sigma_1}\right)\exp\left(-e^{-\frac{x_i-\mu_1}{\sigma_1}}\right)e^{-\frac{x_i-\mu_1}{\sigma_1}} + (1-p)\left(\frac{1}{\sigma_2}\right)\exp\left(-e^{-\frac{x_i-\mu_2}{\sigma_2}}\right)e^{-\frac{x_i-\mu_2}{\sigma_2}} \quad (3.2)$$

where μ_1 and σ_1 are the location and scale parameters respectively for the first probability density function in the mixture model, and μ_2 and σ_2 are for the second probability density function. The parameter *p* is the mixing parameter, and *i* = 1, 2,…….*n* observations. The Gumbel distribution is unimodal, and its shape parameter has a value of zero. Since the shape parameter determines the shape, and depending on its value, the function can change drastically, using a distribution such as Gumbel for which the shape parameter is zero does not adequately model the data that exhibits bimodality. The simulation results described below serves as analytical basis for not using a mixture of two Gumbels to model the maximum concentration data.

Simulation

Three Gumbel density functions were simulated using a set of pre-chosen parameters. The first Gumbel y_1 was generated using $\mu_1 = 200$ and $\sigma_1 = 0.2 * \mu_1$. The second Gumbel y_2 was generated using $\mu_2 = \mu_1 + \mu_1 * d$ and $\sigma_2 = \sigma_1$, where the value of *d* was fixed. The third Gumbel y_3 is a mixture of y_1 and y_2 defined as

$$y_3 = p * y_1 + (1-p) * y_2$$

where *p* is a mixture parameter $(0 \le p \le 1)$. The value of $p = 0.40$ was used.

The above simulation was performed for three different sample sizes 1000, 500, and 100, respectively. Each simulation was conducted for each of the five values of the parameter $d = 1.5, 1.0, 0.5, 0.25, 0.01$.

In the next step, a mixture of two Gumbel data sets was simulated using the same set of parameters as described above. A mixture of two Gumbel models was then fit to the simulated data and the maximum likelihood estimators were obtained. Table 3.1 displays the results of the mixture model fit.

Table 3.1 Maximum likelihood estimators of the parameters with standard errors from the mixture of two Gumbel models fit to the mixture of two Gumbel data sets simulated using $n = 1000, \mu_1 = 200, \sigma_1 = \sigma_2 = 40, \ \mu_2 = 200 * d, \ p = 0.40$

n	d	μ_1(SE)	σ_1(SE)	μ_2(SE)	σ_2(SE)	p(SE)
1000	1.50	203.9 (2.23)	41.5 (1.71)	499.1 (1.64)	37.9 (1.27)	0.40 (0.02)
1000	1.00	203.1 (2.38)	40.9 (2.06)	403.1 (1.84)	39.1 (1.34)	0.40 (0.02)
1000	0.50	202.4 (4.61)	40.8 (3.55)	308.5 (3.10)	39.4 (1.86)	0.44 (0.04)
1000	0.25	**225.5** (1.67)	**48.4** (1.18)	**599.5** $(-\infty)$	**1.0** $(-\infty)$	**0.99** (0.03)
1000	0.05	208.3 (1.37)	41.0 (1.04)	**421.3** (0.64)	**1.0** (1.18)	**0.99** (0.00)

The mixture of two Gumbel model turns out to be a reasonable fit to the data when $d = 1.5, 1.0,$ and 0.5. The maximum likelihood parameter estimates are very close to the pre-set values that were used to simulate the data. However, for values of $d = 0.25$ and 0.05, the maximum likelihood estimators and the standard errors are unreliable, leading to the unsuitability of a mixture of two Gumbel models as d gets smaller and smaller.

A similar simulation analysis was done when the sample sizes are 500 and 100. Table 3.2 and Table 3.3 present the results from the analysis using the mixture of two Gumble models when sample sizes n = 500 and n = 100, respectively.

Table 3.2 Maximum likelihood estimators of the parameters with standard errors from the mixture of two Gumbel models fit to the mixture of two Gumbel data sets simulated using $n = 500, \mu_1 = 200, \sigma_1 = \sigma_2 = 40, \ \mu_2 = 200 * d, \ p = 0.40$

n	d	μ_1(SE)	σ_1(SE)	μ_2(SE)	σ_2(SE)	p(SE)
500	1.50	192.9 (3.02)	40.4 (2.53)	500.2 (2.39)	39.1 (1.80)	0.40 (0.02)
500	1.00	195.1 (3.28)	39.1 (2.52)	399.8 (2.59)	39.9 (2.14)	0.39 (0.02)
500	0.50	**244.1** (3.35)	**62.4** (2.80)	**330.8** (2.71)	8.91 (3.06)	**0.92** (0.02)
500	0.25	**223.9** (2.52)	**60.0** (1.88)	**270.0** (0.84)	**1.90** (0.85)	**0.96** (0.02)
500	0.05	205.2 (1.90)	40.0 (1.48)	**316.0** (1.09)	**1.20** (1.22)	**0.99** (0.01)

Table 3.3 Maximum likelihood estimators of the parameters with standard errors from the mixture of two Gumbel models fit to the mixture of two Gumbel data sets simulated using $n = 100, \mu_1 = 100, \sigma_1 = \sigma_2 = 40, \ \mu_2 = 200 * d, \ p = 0.40$

n	d	μ_1(SE)	σ_1(SE)	μ_2(SE)	σ_2(SE)	p(SE)
100	1.50	205.4 (6.40)	35.3 (5.17)	507.6(2.39)	44.4 (4.85)	0.40 (0.05)
100	1.00	203.1 (5.57)	**32.8** (5.16)	405.8 (6.26)	44.1 (5.45)	0.41 (0.05)
100	0.50	**241.4** (7.34)	**68.4** (7.25)	**303.9** (0.91)	**1.00** (0.88)	**0.95** (0.03)
100	0.25	**226.8** (5.20)	**45.0** (3.86)	**448.6** (6.55)	**1.00** (17.5)	**0.98** (0.10)
100	0.05	202.9 (4.00)	**36.9** (3.26)	**160.5** (2.88)	**1.00** (4.10)	**0.98** (0.08)

From the values that are shaded in the above tables, it is clear that the maximum likelihood estimators for the parameters of a mixture of two Gumbel models are unreliable. This becomes even more evident when the sample size gets smaller. These analyses results clearly indicate the non-suitability of a mixture of two Gumbel models, especially, for sample sizes ≤ 500; moreover, for sample sizes as large as 1000 the mixture of two Gumbels is not suitable when the difference in the two Gumbel location parameters is $\leq 0.25\sigma$ units. Hence, a mixture of two Gumble models will not be further investigated.

3.5.3 Mixture of Two Pareto Distributions

Using the probability density function for Pareto, equation (3.1) can now be written as,

$$f_i(x_i;\sigma_1,\sigma_2,\lambda_1,\lambda_2,p)=\frac{p}{\sigma_1}\left(1-\frac{\lambda_1 x_i}{\sigma_1}\right)^{\frac{1}{\lambda_1}-1}+\frac{1-p}{\sigma_2}\left(1-\frac{\lambda_2 x_i}{\sigma_2}\right)^{\frac{1}{\lambda_2}-1} \tag{3.3}$$

where λ_1 and σ_1 are the shape and scale parameters respectively for the first probability density function in the mixture model, and λ_2 and σ_2 are for the second probability density function. The parameter *p* is the mixing parameter as previously defined, and *i* = 1, 2,…….*n* observations.

Maximum Likelihood Estimation of the Parameters

The log-likelihood function for a sample of *n* observations is given by

$$\ln L(x,\theta)=\sum_{i=1}^{n}\ln\left[\frac{p}{\sigma_1}\left(1-\frac{\lambda_1 x_i}{\sigma_1}\right)^{\frac{1}{\lambda_1}-1}+\frac{1-p}{\sigma_2}\left(1-\frac{\lambda_2 x_i}{\sigma_2}\right)^{\frac{1}{\lambda_2}-1}\right] \tag{3.4}$$

The vector of partial derivatives of $\ln L(x,\theta)$ with respect to $\theta = \{\sigma_1, \sigma_2, \lambda_1, \lambda_2, p\}$ gives the score function as $\sum_{i=1}^{n} \frac{\partial \ln f_i(x_i;\theta)}{\partial \theta_j}$, where j = 1, 2, 3, 4, 5.

Further simplification of (3.4) yields the following equation

$$\ln L = \sum_x \ln \left[\frac{p\left(\frac{\sigma_1 - \lambda_1 x_i}{\sigma_1}\right)^{-\frac{-1+\lambda_1}{\lambda_1}} \sigma_2 + (1-p)\sigma_1 \left(-\frac{-\sigma_2 + \lambda_2 x_i}{\sigma_2}\right)^{-\frac{-1+\lambda_2}{\lambda_2}}}{\sigma_1 \sigma_2} \right] \tag{3.5}$$

The partial derivative [66] of (3.5) with respect to the parameter λ_1 is equal to

$$\frac{\partial \ln L}{\partial \lambda_1} = \sum_x \left[-\left[\frac{p\left(\frac{\sigma_1 - \lambda_1 x_i}{\sigma_1}\right)^{-\frac{-1+\lambda_1}{\lambda_1}} \left\{ \left(-\frac{1}{\lambda_1} + \frac{-1+\lambda_1}{\lambda_1^2}\right) \ln\left(\frac{\sigma_1 - \lambda_1 x_i}{\sigma_1}\right) + \frac{(-1+\lambda_1)x_i}{\lambda_1(\sigma_1 - \lambda_1 x_i)} \right\} \sigma_2}{-\sigma_2 p\left(\frac{\sigma_1 - \lambda_1 x_i}{\sigma_1}\right)^{-\frac{-1+\lambda_1}{\lambda_1}} - (1-p)\left(-\frac{-\sigma_2 + \lambda_2 x_i}{\sigma_2}\right)^{-\frac{-1+\lambda_2}{\lambda_2}}} \right] \right]$$

Simplifying and setting it equal to zero gives the following equation:

$$\frac{1}{\lambda_1^2} \left[p\sigma_2 \left[\sum_x \frac{\left(\frac{\sigma_1 - \lambda_1 x_i}{\sigma_1}\right)^{-\frac{-1+\lambda_1}{\lambda_1}} \left\{ \ln\left(\frac{\sigma_1 - \lambda_1 x_i}{\sigma_1}\right)\sigma_1 - \lambda_1 x_i \ln\left(\frac{\sigma_1 - \lambda_1 x_i}{\sigma_1}\right) + \lambda_1 x_i - \lambda_1^2 x_i \right\}}{(\sigma_1 - \lambda_1 x_i)\left\{ -p\sigma_2 \left(\frac{\sigma_1 - \lambda_1 x_i}{\sigma_1}\right)^{-\frac{-1+\lambda_1}{\lambda_1}} - (1-p)\sigma_1 \left(-\frac{-\sigma_2 + \lambda_2 x_i}{\sigma_2}\right)^{-\frac{-1+\lambda_2}{\lambda_2}} \right\}} \right] \right] = 0 \tag{3.6}$$

Taking the partial derivative of log-likelihood with respect to λ_2 and setting it equal to 0, the following equation is obtained $\frac{\partial \ln L}{\partial \lambda_2} = -\frac{1}{\lambda_2^2} \sum_x \frac{a}{b}$ (3.7)

where *a* and *b* are given by the following equations:

$$a = \left(-\frac{-\sigma_2 + \lambda_2 x_i}{\sigma_2} \right)^{\frac{1}{\lambda_2}} (\sigma_1 - \lambda_1 x_i) \left\{ \begin{array}{l} (\sigma_2 - \lambda_2 x_i) \ln\left(-\frac{-\sigma_2 + \lambda_2 x_i}{\sigma_2} \right) + \lambda_2 x_i - \lambda_2^2 x_i \\ -(\sigma_2 - \lambda_2 x_i)\, p \ln\left(-\frac{-\sigma_2 + \lambda_2 x_i}{\sigma_2} \right) - p x_i \lambda_2 + p x_i \lambda_2^2 \end{array} \right\}$$

$$b = \left[\left(-p\sigma_2 + p\lambda_2 x_i \right) \left(\frac{\sigma_1 - \lambda_1 x_i}{\sigma_1} \right)^{\frac{1}{\lambda_1}} + \left\{ \left(-\sigma_1 + \lambda_1 x_i \right) + p\left(\sigma_1 - \lambda_1 x_i \right) \right\} \left(-\frac{-\sigma_2 + \lambda_2 x_i}{\sigma_2} \right)^{\frac{1}{\lambda_2}} \right] * \left[-\sigma_2 + \lambda_2 x_i \right]$$

The derivative with respect to σ_1 is given by

$$\frac{\partial \ln L}{\partial \sigma_1} = -\frac{p}{\sigma_1} \sum_x \frac{c}{d} \tag{3.8}$$

where *c* and *d* are given by the following equations:

$$c = \left(\frac{\sigma_1 - \lambda_1 x_i}{\sigma_1} \right)^{\frac{1}{\lambda_1}} \left(-\sigma_2 + \lambda_2 x_i \right) \left(-x_i + \sigma_1 \right)$$

$$d = \left[p\left(-\sigma_2 + \lambda_2 x_i \right) \left(\frac{\sigma_1 - \lambda_1 x_i}{\sigma_1} \right)^{\frac{1}{\lambda_1}} + \left(-\sigma_1 + \lambda_1 x_i + p\sigma_1 - p\lambda_1 x_i \right) \left(-\frac{-\sigma_2 + \lambda_2 x_i}{\sigma_2} \right)^{\frac{1}{\lambda_2}} \right] * (\sigma_1 - \lambda_1 x_i)$$

The partial derivative [66] with respect to σ_2 is given by

$$\frac{\partial \ln L}{\partial \sigma_2} = -\frac{1}{\sigma_2} \sum_x \frac{e}{f} \tag{3.9}$$

where *e* and *f* are given by the following equations:

$$e = \left(-\frac{-\sigma_2 + \lambda_2 x_i}{\sigma_2} \right)^{\frac{1}{\lambda_2}} \left(\sigma_1 - \lambda_1 x_1 \right) \left(-x_i + p x_i + \sigma_2 - p\sigma_2 \right)$$

$$f = \left\{ p\left(-\sigma_2 + \lambda_2 x_i\right)\left(\frac{\sigma_1 - \lambda_1 x_i}{\sigma_1}\right)^{\frac{1}{\lambda_1}} + \left(-\sigma_1 + \lambda_1 x_i + p\sigma_1 - p\lambda_1 x_i\right)\left(-\frac{-\sigma_2 + \lambda_2 x_i}{\sigma_2}\right)^{\frac{1}{\lambda_2}} \right\} * \left(-\sigma_2 + \lambda_2 x_i\right)$$

$$\frac{\partial \ln L}{\partial p} = 0 \Rightarrow \frac{-\sigma_2\left(\frac{\sigma_1 - \lambda_1 x_i}{\sigma_1}\right)^{-\frac{-1+\lambda_1}{\lambda_1}} + \sigma_1\left(-\frac{-\sigma_2 + \lambda_2 x_i}{\sigma_2}\right)^{-\frac{-1+\lambda_2}{\lambda_2}}}{-p\sigma_2\left(\frac{\sigma_1 - \lambda_1 x_i}{\sigma_1}\right)^{-\frac{-1+\lambda_1}{\lambda_1}} - (1-p)\sigma_1\left(-\frac{-\sigma_2 + \lambda_2 x_i}{\sigma_2}\right)^{-\frac{-1+\lambda_2}{\lambda_2}}} = 0 \qquad (3.10)$$

Equations (3.6) through (3.10) are the normal equations that cannot be solved explicitly to get closed form solutions for the maximum likelihood estimates $\hat{\sigma}_1, \hat{\sigma}_2, \hat{\lambda}_1, \hat{\lambda}_2$ and $\hat{p}$. Instead, for each sample $x_1, x_2, ..., x_n$, the equations must be solved using an iterative numerical procedure. The maximum likelihood estimates are those values of $\hat{\sigma}_1, \hat{\sigma}_2, \hat{\lambda}_1, \hat{\lambda}_2$ and $\hat{p}$ that maximize the likelihood. The parameter estimates were calculated numerically using the *mle* version 2.2 statistical programming language [51].

3.5.4 Mixture of Two Weibull Distributions

Using the probability density function for a two-parameter Weibull, equation (3.1) can now be written as,

$$f_i\left(x_i; \sigma_1, \sigma_2, \lambda_1, \lambda_2, p\right) = p\left(\frac{\lambda_1}{\sigma_1}\right)\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1 - 1} \exp\left(-\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}\right) + (1-p)\left(\frac{\lambda_2}{\sigma_2}\right)\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2 - 1} \exp\left(-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}\right) \qquad (3.11)$$

where λ_1 and σ_1 are the shape and scale parameters respectively for the first probability density function in the mixture model, and λ_2 and σ_2 are for the second probability density function. The parameter *p* is the mixing parameter, and *i* = 1, 2,…….*n* observations.

Maximum Likelihood Estimation of the Parameters

Maximum likelihood was used to estimate model parameters. The likelihood function for a sample of *n* observations is given by

$$L(x,\theta)=\prod_{i=1}^{n}\left[p\left(\frac{\lambda_1}{\sigma_1}\right)\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1-1}\exp\left(-\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}\right)+(1-p)\left(\frac{\lambda_2}{\sigma_2}\right)\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2-1}\exp\left(-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}\right)\right]$$

The log likelihood function is given by

$$\ln L(x,\theta)=\sum_{i=1}^{n}\ln\left[p\left(\frac{\lambda_1}{\sigma_1}\right)\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1-1}\exp\left(-\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}\right)+(1-p)\left(\frac{\lambda_2}{\sigma_2}\right)\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2-1}\exp\left(-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}\right)\right] \tag{3.12}$$

The vector of partial derivatives of $\ln L(x,\theta)$ with respect to $\theta=\{\sigma_1,\sigma_2,\lambda_1,\lambda_2,p\}$ gives the score function as:

$$\sum_{i=1}^{n}\frac{\partial \ln f_i(x_i;\theta)}{\partial\theta_j}\text{, where } j = 1, 2, 3, 4, 5$$

We now find the partial derivatives [66] with respect to each of the five unknown parameters and set each one of them equal to zero to get a set of normal equations.

$$\ln L=\sum_{i=1}^{n}\ln\left[\frac{1}{\sigma_1\sigma_2}\left[\begin{array}{l}p\lambda_1\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1-1}\exp\left(-\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}\right)\sigma_2+\lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2-1}\exp\left(-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}\right)\sigma_1\\ -\lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2-1}\exp\left(-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}\right)\sigma_1 p\end{array}\right]\right]$$

$$\ln L=\sum_{i=1}^{n}\ln\left[\frac{1}{x_i}\left[\begin{array}{l}p\lambda_1\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}\exp\left(-\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}\right)+\lambda_2\left(\frac{x_1}{\sigma_2}\right)^{\lambda_2}\exp\left(-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}\right)\\ -\lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}\exp\left(-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}\right)p\end{array}\right]\right]$$

$$\frac{\partial \ln L}{\partial \lambda_1} =$$

$$\sum_{i=1}^{n}\left[\frac{\left[p\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1} e^{\left(-\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}\right)} + p\lambda_1\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}\ln\left(\frac{x_i}{\sigma_1}\right)e^{\left(-\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}\right)} - p\lambda_1\left(\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}\right)^2\ln\left(\frac{x_i}{\sigma_1}\right)e^{\left(-\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}\right)}\right]}{\left[p\lambda_1\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1} e^{\left(-\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}\right)} + \lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2} e^{\left(-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}\right)} - \lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2} e^{\left(-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}\right)} p\right]}\right]$$

simplifying and setting it equal to zero results in the following equation:

$$p\left[\sum_{i=1}^{n}\left[\frac{\left[e^{\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}}\left[\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1} + \lambda_1\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}\ln\left(\frac{x_i}{\sigma_1}\right) - \lambda_1\left(\frac{x_i}{\sigma_1}\right)^{2\lambda_1}\ln\left(\frac{x_i}{\sigma_1}\right)\right]\right]}{\left[p\lambda_1\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1} e^{\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}} + \lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2} e^{\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}} - \lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2} pe^{\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}}\right]}\right]\right] = 0 \qquad (3.13)$$

$$\frac{\partial \ln L}{\partial \lambda_2} = \sum_{i=1}^{n}\left[\frac{\text{numerator}}{\text{denominator}}\right]$$

where the denominator term is given by the following expression:

$$p\lambda_1\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1} e^{-\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}} + \lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2} e^{-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}} - \lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2} e^{-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}} p$$

and the numerator term is given by the following expression:

$$\left[\begin{array}{l}\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2} e^{-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}} + \lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}\ln\left(\frac{x_i}{\sigma_2}\right)e^{-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}} - \lambda_2\left(\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}\right)^2\ln\left(\frac{x_i}{\sigma_2}\right)e^{-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}} \\ -\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2} e^{-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}} p - \lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}\ln\left(\frac{x_i}{\sigma_2}\right)e^{-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}} p + \lambda_2\left(\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}\right)^2\ln\left(\frac{x_i}{\sigma_2}\right)e^{-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}} p\end{array}\right]$$

$$\frac{\partial \ln L}{\partial \lambda_2} = 0 \Rightarrow$$

$$-\sum_{i=1}^{n}\left[-\left[\frac{e^{\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}}\left[\begin{array}{l}\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}+\lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}\ln\left(\frac{x_i}{\sigma_2}\right)-\lambda_2\left(\frac{x_i}{\sigma_2}\right)^{2\lambda_2}\ln\left(\frac{x_i}{\sigma_2}\right)-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}p\\ -\lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}\ln\left(\frac{x_i}{\sigma_2}\right)p+\lambda_2\left(\frac{x_i}{\sigma_2}\right)^{2\lambda_2}\ln\left(\frac{x_i}{\sigma_2}\right)p\end{array}\right]}{\left[p\lambda_1\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}e^{\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}}+\lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}e^{\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}}-\lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}pe^{\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}}\right]}\right]\right]=0$$

(3.14)

$$\frac{\partial \ln L}{\partial \sigma_1}=\sum_{i=1}^{n}\left[\left[-\frac{p\frac{\lambda_1^2}{\sigma_1}\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}e^{-\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}}+p\frac{\lambda_1^2}{\sigma_1}\left(\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}\right)^2e^{-\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}}}{p\lambda_1\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}e^{-\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}}+\lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}e^{-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}}-\lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}e^{-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}}p}\right]\right]$$

Simplifying and setting equal to zero gives

$$-\frac{1}{\sigma_1}\left[p\lambda_1^2\left[\sum_{i=1}^{n}\frac{e^{\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}}\left[\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}-\left(\frac{x_i}{\sigma_1}\right)^{2\lambda_1}\right]}{p\lambda_1\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}e^{\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}}+\lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}e^{\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}}-\lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}e^{-\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}}p}\right]\right]=0 \qquad (3.15)$$

$$\frac{\partial \ln L}{\partial \sigma_2} = \sum_{i=1}^{n} \left[\frac{\left[-\frac{\lambda_2^2}{\sigma_2}\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2} e^{-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}} + \frac{\lambda_2^2}{\sigma_2}\left[\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}\right]^2 e^{-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}} + p\frac{\lambda_2^2}{\sigma_2}\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2} e^{-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}} - p\frac{\lambda_2^2}{\sigma_2}\left[\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}\right]^2 e^{-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}} \right]}{\left[p\lambda_1\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1} e^{-\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}} + \lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2} e^{-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}} - p\lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2} e^{-\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}} \right]} \right]$$

Further simplification of this equation gives the following equation:

$$-\frac{1}{\sigma_2}\left[p\lambda_2^2 \left[\sum_{i=1}^{n} \frac{e^{\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}} \left[\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2} - \left(\frac{x_i}{\sigma_2}\right)^{2\lambda_2} - \left(\frac{x_i}{\sigma_2}\right)^{\lambda_2} p + p\left(\frac{x_i}{\sigma_2}\right)^{2\lambda_2} \right]}{p\lambda_1\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1} e^{\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}} + \lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2} e^{\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}} - \lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2} e^{-\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}} p} \right] \right] = 0 \qquad (3.16)$$

$$\frac{\partial \ln L}{\partial p} = 0 \Rightarrow \sum_{i=1}^{n} \left[\frac{\lambda_1\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1} e^{\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}} - \lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2} e^{\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}}}{p\lambda_1\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1} e^{\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2}} + \lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2} e^{\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}} - p\lambda_2\left(\frac{x_i}{\sigma_2}\right)^{\lambda_2} e^{\left(\frac{x_i}{\sigma_1}\right)^{\lambda_1}}} \right] = 0 \qquad (3.17)$$

These normal equations (3.13) through (3.17) cannot be solved explicitly to give general formulas for the maximum likelihood estimates $\hat{\sigma}_1, \hat{\sigma}_2, \hat{\lambda}_1, \hat{\lambda}_2$ and $\hat{\mathrm{p}}$. Instead, for each sample $x_1, x_2, ..., x_n$, the equations must be solved using an iterative numerical procedure. The maximum likelihood estimates are those values of $\hat{\sigma}_1, \hat{\sigma}_2, \hat{\lambda}_1, \hat{\lambda}_2$ and $\hat{\mathrm{p}}$ that maximize the likelihood. The parameter estimates were calculated numerically using the *mle* version 2.2 statistical programming language [51].

3.6 Results of the Modeling on Sample Sizes 50 and 100

3.6.1 Mixture of Two Pareto Distributions

The Pareto mixture model was fit into the data to obtain all 5 parameter estimates using the maximum likelihood method for various start values of the mixture parameter p . In the model, p is the estimated mixture parameter, σ_1 and ξ_1 are the scale and shape parameters of the Pareto distribution for the first subgroup, and σ_2 and ξ_1 are the scale and shape parameters of the Pareto distribution for the second subgroup.

For n = 50 and for any start value of p greater or equal to 0.5, with an increment of 0.05, the model resulted in a singular matrix and the results obtained with start value of p being 0.05, 0.10, 0.15, 0.25, and 0.45 are displayed in Table 3.4

Table 3.4 Parameter estimates and the standard errors from the mixture Pareto model (n = 50). The Akaike Information Criterion (AIC) for this model is 720.01.

PARAMETER	p	σ_1	σ_2	ξ_1	ξ_2
ESTIMATE	0.63	1121.7	1587.5	4.40	4.99
STANDARD ERROR	0.16	0	134.86	2.12	2.46

For n = 100, and for any start value of p with an increment of 0.05, Table 3.5 displays the results. All models with different start values for p gave the same parameter estimates but the standard errors of the parameter estimates of the second component of the model are extremely high. Moreover, 99.1% of the data values were put in one subgroup for each of these models.

Table 3.5 Parameter estimates and the standard errors from the mixture Pareto model (n = 100). The Akaike Information Criterion (AIC) for this model is 1537.22

PARAMETER	p	σ_1	σ_2	ξ_1	ξ_2
ESTIMATE	0.99	875.9	1331.4	1.91	0.50
STANDARD ERROR	0.10	0.00	1167843201	0.49	986008150

3.6.2 Mixture of Two Weibull Distributions

The Weibull mixture model was fit into the data to obtain all 5 parameter estimates using the maximum likelihood method for various start values of the mixture parameter p . The estimated parameters p , σ_1, ξ_1, σ_2 and ξ_1 are as defined before but for the Weibull distributions.

For n = 50, and for any start value of p less than or equal to 0.45 with an increment of 0.05 Table 3.6A displays the results. The density graphs corresponding to the parameter estimates in Table 3.6A, obtained from the mixture of two Weibull models, are presented in Figure 3.5A.

Table 3.6A Parameter estimates and the standard errors from the mixture Weibull model (n = 50). The Akaike Information Criterion (AIC) for this model is 727.63.

PARAMETER	p	σ_1	σ_2	ξ_1	ξ_2
ESTIMATE	0.44	1981.66	1452.70	7.81	8.57
STANDARD ERROR	0.23	169.8	60.41	4.17	2.62

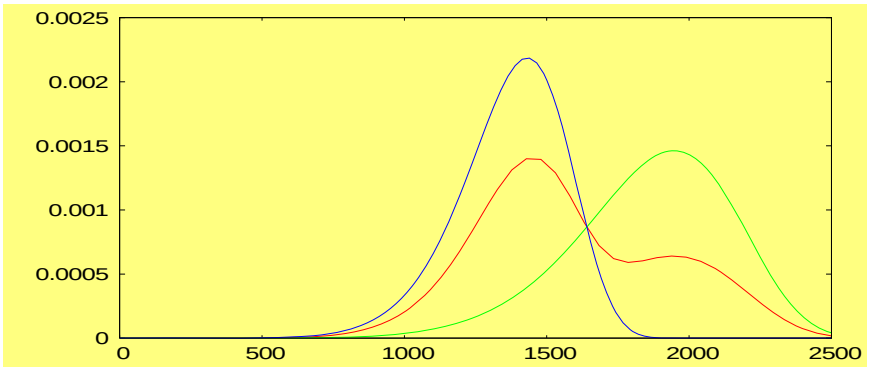


Figure 3.5A Fitted distribution of maximum drug concentration for each subgroup and combined subgroups, based on the parameter estimates in Table 3.6A

For n = 50, and for start value of *p* greater or equal to 0.55, results are displayed in Table 3.6B. The start value of *p* is 0.5 resulted in a singular matrix. The density graphs corresponding to the parameter estimates in Table 3.6B are presented in Figure 3.5B.

Table 3.6B Parameter estimates and the standard errors from the mixture Weibull model (n = 50). The AIC value for this model is 727.93.

PARAMETER	p	σ_1	σ_2	ξ_1	ξ_2
ESTIMATE	0.31	1331.37	1857.15	11.99	6.25
STANDARD ERROR	0.15	45.64	102.29	5.38	1.84

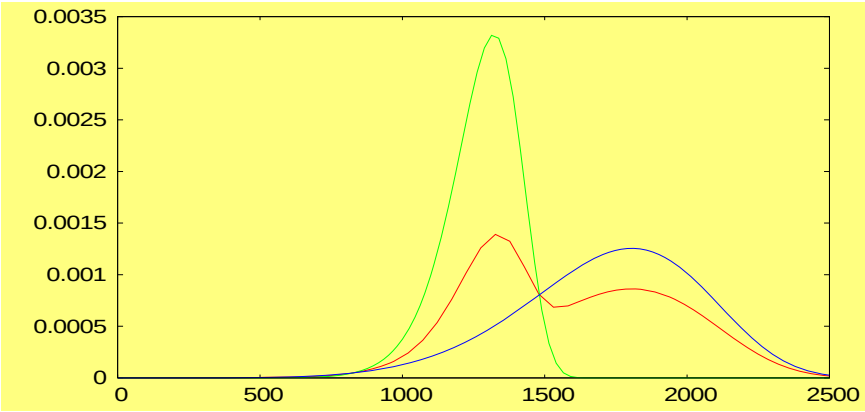


Figure 3.5B Fitted distribution of maximum drug concentration for each subgroup and combined subgroups, based on the parameter estimates in Table 3.6B

For n = 100, and for start value of *p* less than or equal to 0.45 with an increment of 0.05, Table 3.7A displays the results. The density graphs corresponding to the parameter estimates in Table 3.7A are presented in Figure 3.6A.

Table 3.7A Parameter estimates and the standard errors from the mixture Weibull model (n = 100). The AIC value for this model is 1445.41

PARAMETER	p	σ_1	σ_2	ξ_1	ξ_2
ESTIMATE	0.65	1472.56	1893.03	7.06	5.87
STANDARD ERROR	0.55	57.9	558.58	2.01	6.49

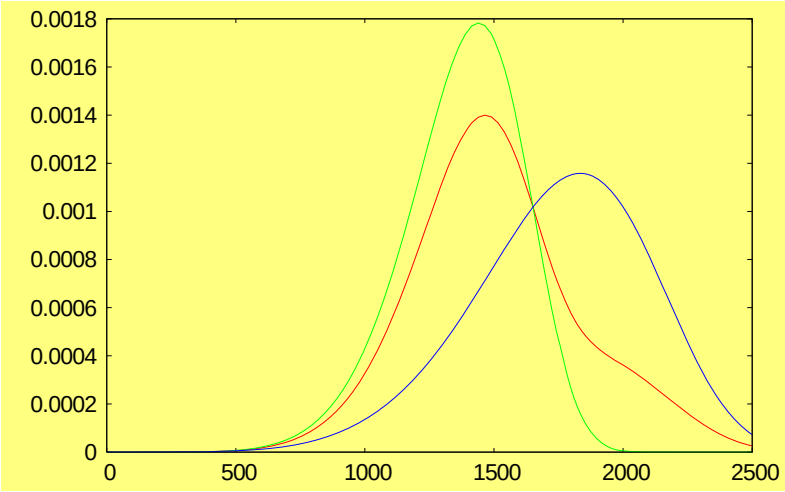


Figure 3.6A Fitted distribution of maximum drug concentration for each subgroup and combined subgroups, based on the parameter estimates in Table 3.7A

For n = 100, the results, when start value of *p* is greater or equal to 0.50, are displayed in Table 3.7B. The density graphs corresponding to the parameter estimates in Table 3.7B are presented in Figure 3.6B.

Table 3.7B Parameter estimates and the standard errors from the mixture Weibull model (n = 100). The AIC value for this model is 1445.41

PARAMETER	p	σ_1	σ_2	ξ_1	ξ_2
ESTIMATE	0.28	1326.38	1737.35	11.23	5.31
STANDARD ERROR	0.17	43.58	106.56	4.95	1.09

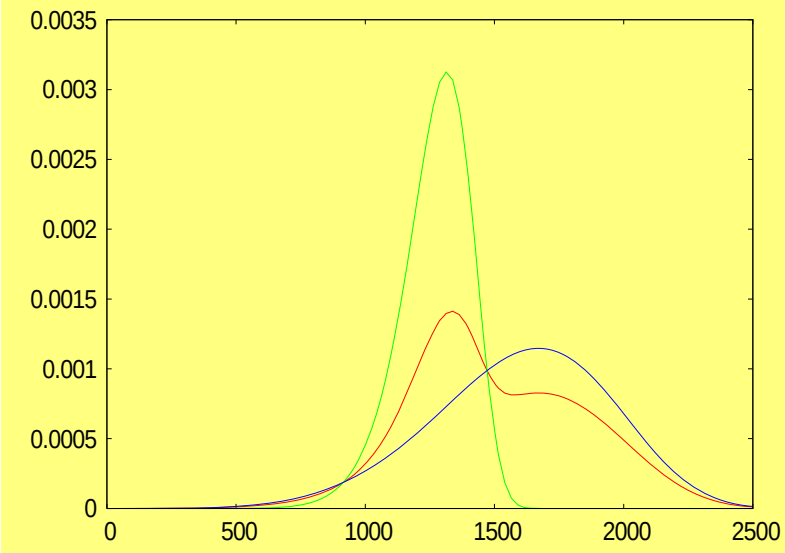


Figure 3.6B Fitted distribution of maximum drug concentration for each subgroup and combined subgroups, based on the parameter estimates in Table 3.7B

The graphs in Figures 3.7A and 3.7B show the predicted maximum drug concentration based on the mixture model parameters displayed in Table 3.6A and Table 3.6B, respectively, when the sample size is 50. Similarly, graphs in Figures 3.8A and 3.8B show the predicted maximum drug concentration based on the mixture model parameters displayed in Table 3.7A and Table 3.7B, respectively, when the sample size is 100.

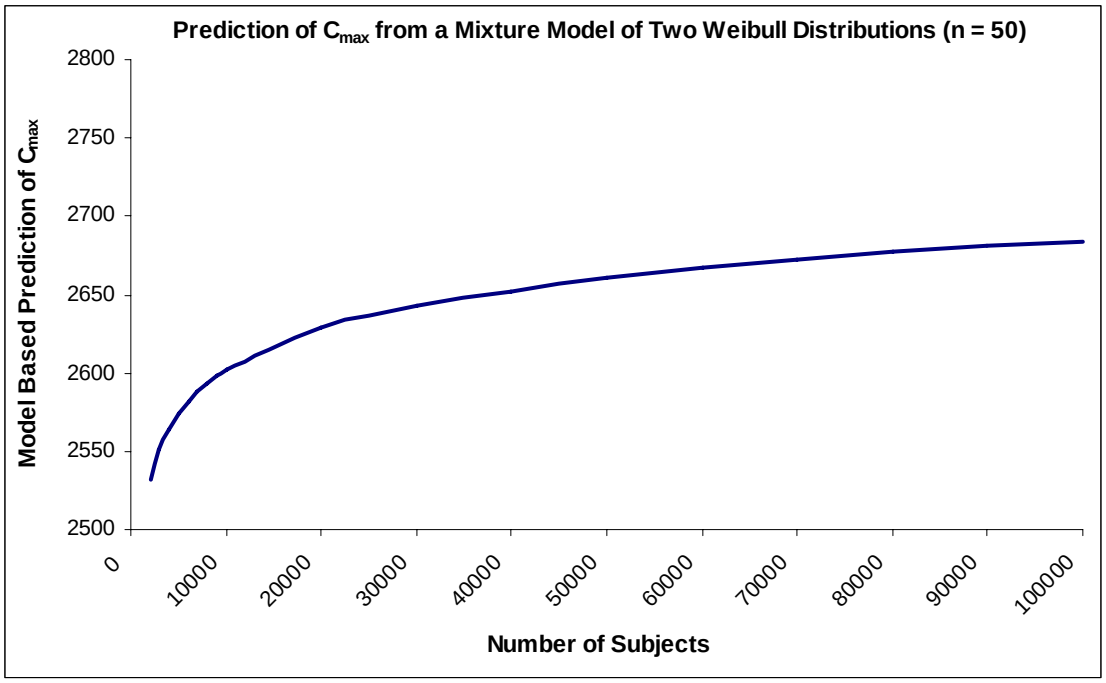


Figure 3.7A Prediction of C_{max} based on the model parameters displayed in Table 3.6A.

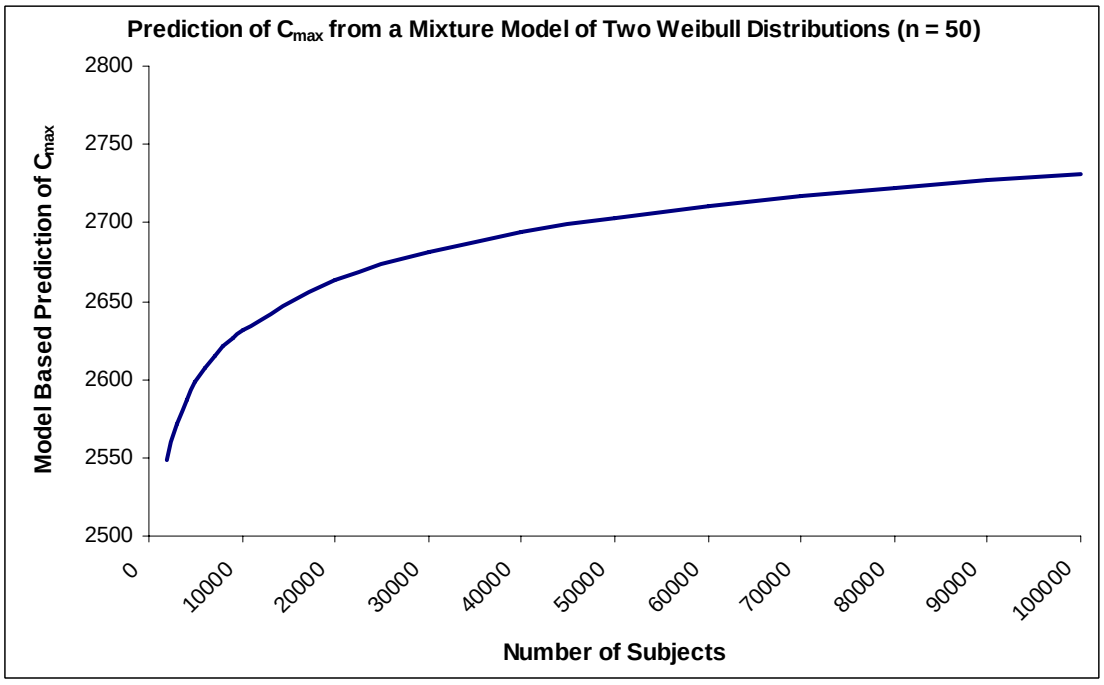


Figure 3.7B Prediction of C_{max} based on the model parameters displayed in Table 3.6B.

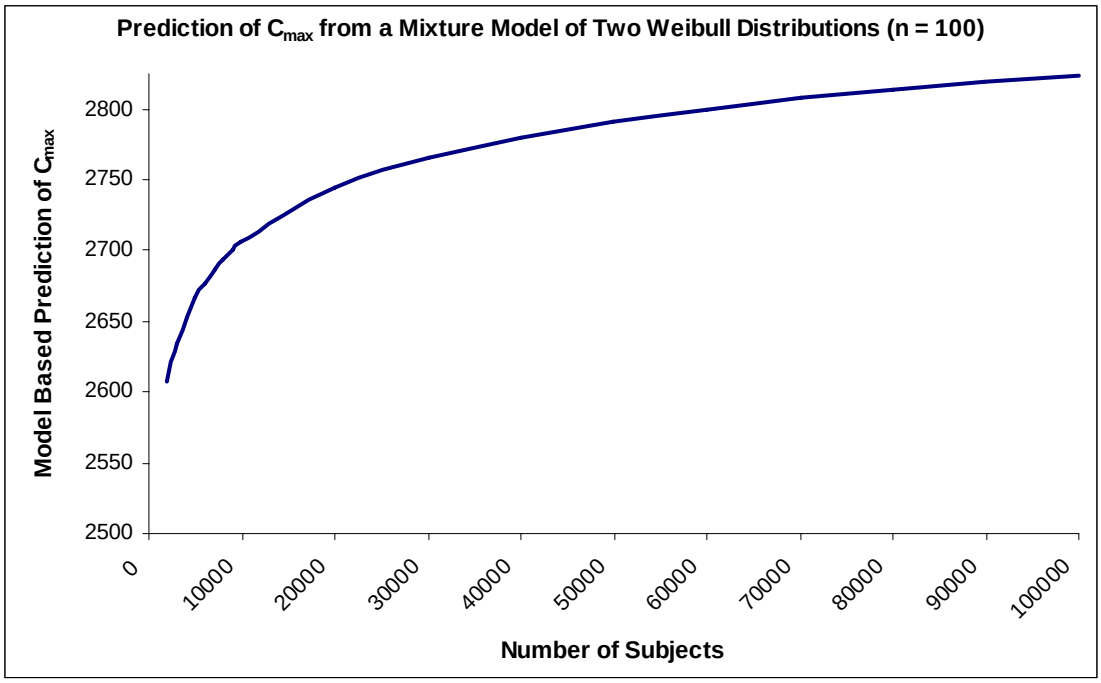


Figure 3.8A Prediction of C_{max} based on the model parameters displayed in Table 3.7A

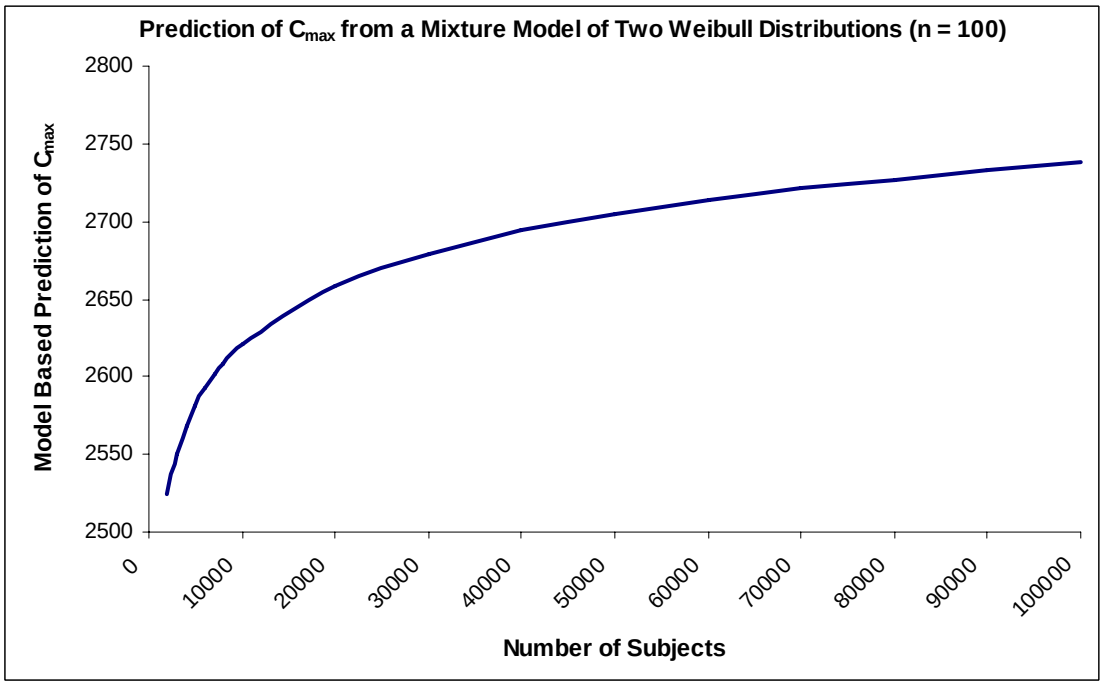


Figure 3.8B Prediction of C_{max} based on the model parameters displayed in Table 3.7B

3.7 Summary and Conclusions

For n = 50, the Pareto model results in a singular matrix when the start value of p is greater or equal to 0.50 with an increment of 0.05. For other start values of p , the model parameters have fluctuating standard errors. The inadequacy of the Pareto model stems from this sensitivity of the model to the start value of p in estimating the model parameters and their standard errors, although the AIC value in this case is slightly smaller in comparison with the AIC value for the Weibull mixture model.

For n = 100, although the values of the estimated model parameters from the Pareto distribution remain almost the same for any start value of the mixture parameter p , its inadequacy comes from the fact that it places 99.1% of the data in one subgroup, and the standard errors for the estimated parameters for the second probability density component of the distribution are extremely high. The value of Akaike Information Criterion (AIC) is much larger in comparison with the AIC value for Weibull mixture model.

This leads to the examination of the results of Weibull mixture model. For n = 50, the parameter estimates and their standard errors from the Weibull model are consistent for any start value of p less than or equal to 0.45 with an increment of 0.05. The results are tabulated in the Table 3.6A. Another set of consistent parameter estimates was obtained for any start value of p greater or equal to 0.55 with an increment of 0.05. These results are presented in Table 3.6B. The AIC value remains the same in both cases.

For n = 100, and any start value of p less than or equal to 0.45, the results are tabulated in Table 3.7A. Similarly, for n = 100 and any start value of p greater or equal to 0.5, the results are tabulated in Table 3.7B. One can observe that the estimated value of

the mixture parameter in Table 3.7B is equal to (1 – estimated p from Table 3.7A), and the estimates of the scale parameters get interchanged as well as estimates of the shape parameters.

Based on the comparison of the AIC values between Pareto and Weibull for both sample sizes n = 50 and n = 100, and taking into account the inconsistency in the estimation of the Pareto model parameters, a mixture of two Weibull models is the only suitable model to fit to the data that exhibit bimodality when sample size is small.

Chapter Four

Statistical Model for the Pharmacokinetic Parameter, Maximum Drug Concentration (C_{max}): Large Samples

4.0 Introduction

In chapter three, we examined a mixture of two extreme value distributions to statistically model the maximum drug concentration data that exhibited a bimodal distribution in small samples ($N \leq 100$). The focus of the present chapter is to examine various statistical models appropriate for large-sample extreme model distributions: generalized extreme value distribution, Gumbel distribution, generalized Pareto distribution, the three-parameter and the two-parameter Weibull distributions. In each of these distributions the maximum likelihood methods were used to estimate the model parameters and their standard errors, and using these estimates, extreme quantiles of the maximum drug concentration distribution were obtained. These models are examined below in separate sections, followed by an overall summary of results.

4.1 Outline of Each Section of Chapter Four

Section 4.2 presents the derivations of the normal equations for the generalized extreme value distribution, and discusses the validity of the model.

Sections 4.3 and 4.4 have the same information as in section 4.3 but for Gumbel and Weibull distributions, respectively. Section 4.5 gives a description of generalized Pareto distribution pertaining to modeling maximum drug concentration data.

Section 4.6 presents a detailed discussion of the results of the statistical modeling of the maximum drug concentration data for large samples using the extreme value models. Section 4.7 contains the return level profiles based on the quartiles, followed by the summary and conclusions in section 4.8.

4.2 Generalized Extreme Value (GEV) Distribution

The Generalized Extreme Value distribution asymptotically models maxima from any distribution with a stable maximum value distribution. The modeling of maxima is used to answer questions about how often, in the future, values larger than a certain value may occur. In the present problem, the interest lies in the estimation of the return level of the maximum drug concentration for any given patient in the population. For example, the interest is to estimate the return level of the maximum drug concentration for an n^{th} patient with $100(1-\alpha)\%$ confidence interval where $\alpha = 0.01, 0.05,$ and 0.10 .

The data set consists of maximum drug concentration values. Each data value can be represented as $M_n = \max\{X_1, X_2, ..., X_n\},$ where $X_1, X_2, ..., X_n$, is a sequence of independent random variables having a common distribution function F. In the present application, the $X_i's$ represent values of drug concentration measured at *n* pre-set time points so that M_n represents the maximum drug concentration of these *n* values for each individual. The distribution of M_n can be derived as [23]

$$\Pr\{M_n \le z\} = \Pr\{X_1 \le z\} * \Pr\{X_2 \le z\} * ..... * \Pr\{X_n \le z\} = \{F(z)\}^n$$

In practice, this is not helpful since F is not known, and a small discrepancy in the estimate of F can lead to substantial discrepancies for F^n . In real-life applications, one has to be able to statistically estimate the probability of predicting the maximum drug

concentration for an individual after the administration of the drug with confidence bounds, and thus characterize the stochastic distribution of maximum drug concentrations.

Probability Density Function

The statistical modeling of the maximum drug concentration data starts with the family of models having the distribution function of the form [23]

$$F_X(x) = \begin{cases} \exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} & -\infty < x \le \mu - \frac{\sigma}{\xi} \text{ for } \xi < 0; \\ & \mu - \frac{\sigma}{\xi} \le x < \infty \text{ for } \xi > 0; \\ \exp\left\{-\exp\left\{-\frac{x-\mu}{\sigma}\right\}\right\} & -\infty < x < \infty \text{ for } \xi = 0; \end{cases} \quad (4.1)$$

This is the generalized extreme value (GEV) family of distributions. When $\xi = 0$, the above distribution reduces to Type 1 extreme value distribution known as the Gumbel distribution. When $\xi > 0$, it reduces to Type 2 extreme value distribution known as the Frechet distribution, and finally, when $\xi < 0$, it reduces to Type 3 extreme value distribution referred to as the Weibull distribution. The probability density function corresponding to (4.1) is

$$f_X(x) = \begin{bmatrix} \exp -\left\{1+\xi\left(\frac{x-\mu}{\sigma}\right)\right\}^{-\frac{1}{\xi}} \frac{1}{\sigma}\left\{1+\xi\left(\frac{x-\mu}{\sigma}\right)\right\}^{-\frac{1}{\xi}-1} & -\infty < x \le \mu - \frac{\sigma}{\xi} \text{ for } \xi < 0; \\ & \mu - \frac{\sigma}{\xi} \le x < \infty \text{ for } \xi > 0; \\ \exp\left\{-e^{-\frac{x-\mu}{\sigma}}\right\}\frac{1}{\sigma}e^{-\frac{x-\mu}{\sigma}} & -\infty \le x < \infty \text{ for } \xi = 0 \end{bmatrix} \quad (4.2)$$

where ξ, μ, and σ are the shape, location and scale parameters, respectively.

4.2.1 Derivation of Normal Equations

Suppose we have *n* observations $X_1, X_2, ..., X_n$ for which the GEV distribution is appropriate. The log-likelihood for the GEV parameters when $\xi \neq 0$ is given by [23]

$$l(\mu,\sigma,\xi) = -n\log\sigma - \left(1+\frac{1}{\xi}\right)\sum_{i=1}^{n}\log\left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right] - \sum_{i=1}^{n}\left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{\frac{-1}{\xi}} \tag{4.3}$$

The partial derivatives [66] of (4.3) with respect to each of the three unknown parameters $\xi, \mu,$ and σ are obtained and set each one of them equal to zero to get a set of normal equations.

$$\frac{\partial l}{\partial \xi} = 0 \Rightarrow$$

$$-\frac{1}{\xi^2}\left[\sum_x\left[\frac{1}{\sigma+\xi x_i-\xi\mu}\left[\begin{array}{l} x_i\xi^2+\xi x_i-\ln\left(\frac{\sigma+\xi x_i-\xi\mu}{\sigma}\right)x_i\xi-x_i\xi\left(\frac{\sigma+\xi x_i-\xi\mu}{\sigma}\right)^{\left(-1/\xi\right)}+ \\ \left(\frac{\sigma+\xi x_i-\xi\mu}{\sigma}\right)^{\left(-1/\xi\right)}\ln\left(\frac{\sigma+\xi x_i-\xi\mu}{\sigma}\right)x_i\xi-\mu\xi^2-\xi\mu+\ln\left(\frac{\sigma+\xi x_i-\xi\mu}{\sigma}\right)\mu\xi \\ -\left(\frac{\sigma+\xi x_i-\xi\mu}{\sigma}\right)^{\left(-1/\xi\right)}\ln\left(\frac{\sigma+\xi x_i-\xi\mu}{\sigma}\right)\mu\xi+\xi\left(\frac{\sigma+\xi x_i-\xi\mu}{\sigma}\right)^{\left(-1/\xi\right)}\mu \\ -\ln\left(\frac{\sigma+\xi x_i-\xi\mu}{\sigma}\right)\sigma+\left(\frac{\sigma+\xi x_i-\xi\mu}{\sigma}\right)^{\left(-1/\xi\right)}\ln\left(\frac{\sigma+\xi x_i-\xi\mu}{\sigma}\right)\sigma \end{array}\right]\right]\right] = 0 \tag{4.4}$$

$$\frac{\partial l}{\partial \mu} = 0 \Rightarrow \quad \sum_x\left[\frac{-\left(\frac{\sigma+\xi x_i-\xi\mu}{\sigma}\right)^{\left(-1/\xi\right)}+1+\xi}{\sigma+\xi x_i-\xi\mu}\right] = 0 \tag{4.5}$$

$$\frac{\partial l}{\partial \sigma} = 0 \;\Rightarrow$$

$$-\frac{\sum_x \left[\frac{\left(\frac{\sigma + \xi x_i - \xi\mu}{\sigma} \right)^{(-1/\xi)} x_i - x_i - \left(\frac{\sigma + \xi x_i - \xi\mu}{\sigma} \right)^{(-1/\xi)} \mu + \sigma + \mu}{\sigma + \xi x_i - \xi\mu} \right]}{\sigma} = 0 \qquad (4.6)$$

These normal equations (4.4) through (4.6) cannot be solved explicitly to give general formulas for the maximum likelihood estimates $\hat{\xi}, \hat{\mu}$, and $\hat{\sigma}$. Instead, for each sample $x_1, x_2, ..., x_n$, the equations must be solved using an iterative numerical procedure. The maximum likelihood estimates are those values of $\hat{\xi}, \hat{\mu}$, and $\hat{\sigma}$ that maximize the likelihood. The parameter estimates and standard errors were calculated numerically using the *evd* (extreme value distribution) package for the R statistical computing system [75, 53].

4.2.2 Validity of the GEV model

The validity of the GEV model to model the maximum drug concentration data is examined through probability, quantile, and return level plots, and each of these plots is based on a comparison of model-based and empirical estimates of the distribution function. Another method to examine the validity of the model is the comparison of the probability density function of a fitted model with a histogram of the data. .

Any departure from linearity in the probability or quantile plot indicates the inadequacy of the model to fit the maximum drug concentration data. The return level plots enable us to assess the adequacy of the model. The model-based curve and

empirical estimates should be in reasonable agreement. The presence of substantial disagreement indicates the inadequacy of the GEV model.

4.3 Gumbel Distribution (Type I Extreme Value Distribution)

Probability Density Function

The shape parameter ξ determines the shape, and depending on its value, the distribution function can change drastically. The standard asymptotic results of consistency, asymptotic efficiency and asymptotic normality hold when $\xi > -0.5$. For $-1 < \xi < -0.5$, maximum likelihood estimators do not have the asymptotic properties and for $\xi < -1$, they are unlikely to be obtainable. It is often of interest to test the hypothesis that the shape parameter ξ is 0, leading to modeling the maximum drug concentration data using Gumbel distribution.

When $\xi = 0$, the probability density function for GEV reduces to Gumbel probability density function given by

$$f_x(x) = \frac{1}{\sigma}\exp\left(-e^{-\frac{x-\mu}{\sigma}}\right)e^{-\frac{x-\mu}{\sigma}}, \quad -\infty < x < \infty \tag{4.7}$$

4.3.1 Derivation of Normal Equations

As in the case of GEV, the log-likelihood for the Gumbel parameters [23] is

$$l(\mu,\sigma) = -n\log\sigma - \sum_{i=1}^{n}\frac{x_i-\mu}{\sigma} - \sum_{i=1}^{n}\exp\left(-\frac{x_i-\mu}{\sigma}\right) \tag{4.8}$$

Taking the partial derivatives of $l(\mu,\sigma)$ with respect to the parameters μ and σ the result is the following:

$$\frac{\partial l}{\partial \mu} = -\frac{\left(\sum_{i=1}^{n} e^{-\frac{x_i-\mu}{\sigma}}\right) - n}{\sigma} \tag{4.9}$$

$$\frac{\partial l}{\partial \sigma} = -\frac{i = \sum_{i=1}^{n}\left[\sigma + e^{-\frac{x_i-\mu}{\sigma}} x_i - e^{-\frac{x_i-\mu}{\sigma}} \mu - x_i + \mu\right]}{\sigma^2} \tag{4.10}$$

Setting $\frac{\partial l}{\partial \mu} = 0$, and solving for μ, we have

$$\mu = \sigma\left[\ln(n) - \ln\left(\sum_{i=1}^{n} e^{-\frac{x_i}{\sigma}}\right)\right] \tag{4.11}$$

Setting $\frac{\partial l}{\partial \sigma} = 0$, and solving for σ, we get

$$\sigma = \bar{x} - \frac{\sum_{i=1}^{n} e^{-\frac{x_i}{\sigma}} x_i}{\sum_{i=1}^{n} e^{-\frac{x_i}{\sigma}}} \tag{4.12}$$

The parameter estimates and standard errors were calculated numerically using the BestFit 4.5.4 [11], a probability distribution fitting software package for Microsoft Windows from Palisade Corporation.

4.3.2 Validity of the Gumbel Model

The validity of the Gumbel model to model the maximum drug concentration data is examined through probability and quantile plots, and through the comparison of the probability density function of a fitted model with a histogram of the data.

The suitability of any particular member of the GEV family can be assessed by comparison with the maximum likelihood estimate within the entire family. The

appropriateness of replacing the GEV family with the Gumbel family, corresponding to the $\xi = 0$ subset of the GEV family is also assessed using the likelihood ratio test statistic.

4.4 Weibull Distribution (Type III Extreme Value Distribution)

Probability Density Function

Since the fitted shape from the generalized extreme value distribution is negative, Weibull distribution was used to model the maximum drug concentration data to examine if it is an appropriate distribution to model the data. The three-parameter probability density function is given by

$$f(x) = \frac{\xi}{\sigma}\left(\frac{x-\mu}{\sigma}\right)^{\xi-1} e^{-\left(\frac{x-\mu}{\sigma}\right)^{\xi}}, \quad \mu, \xi > 0, \ \sigma > 0, \ -\infty < \mu < \infty \tag{4.13}$$

where $\xi, \mu,$ and σ are the shape (also called Weibull slope), location, and scale parameters, respectively. Different values of ξ can have marked effects on the behavior of the distribution.

4.4.1 Derivation of Normal Equations

For a random sample $x_1, x_2, ..., x_n$, the log-likelihood function is given by

$$\ln L = n \ln \xi - n \ln \sigma + (\xi - 1)\sum_{i=1}^{n} \ln\left(\frac{x_i - \mu}{\sigma}\right) - \sum_{i=1}^{n}\left(\frac{x_i - \mu}{\sigma}\right)^{\xi} \tag{4.14}$$

The maximum likelihood estimates $\hat{\mu}, \hat{\sigma},$ and $\hat{\xi}$ are obtained by setting

$\frac{\partial \ln L}{\partial \mu} = 0, \frac{\partial \ln L}{\partial \sigma} = 0,$ and $\frac{\partial \ln L}{\partial \xi} = 0$ at $\mu = \hat{\mu}, \sigma = \hat{\sigma},$ and $\xi = \hat{\xi}$ respectively.

After some mathematical simplification, we get the equations

$$\frac{n}{\hat{\xi}} + \sum_{i=1}^{n} \log(t_i - \hat{\mu}) = \frac{n \sum_{i=1}^{n} (t_i - \hat{\mu})^{\hat{\xi}} \log(t_i - \hat{\mu})}{\sum_{i=1}^{n} (t_i - \hat{\mu})^{\hat{\xi}}} \tag{4.15}$$

$$\frac{n\hat{\xi} \sum_{i=1}^{n} (t_i - \hat{\mu})^{\hat{\xi}-1}}{\sum_{i=1}^{n} (t_i - \hat{\mu})^{\hat{\xi}}} = (\hat{\xi} - 1) \sum_{i=1}^{n} \frac{1}{t_i - \hat{\mu}} \tag{4.16}$$

and $$\hat{\sigma} = \left\{ \frac{1}{n} \sum_{i=1}^{n} (t_i - \hat{\mu})^{\hat{\xi}} \right\}^{1/\hat{\xi}} \tag{4.17}$$

These equations can be solved numerically to obtain the estimates $\hat{\mu}, \hat{\sigma}$, and $\hat{\xi}$. The parameter estimates and standard errors were calculated numerically using the Statistical Analysis Software (SAS), Version 9.1, Cary, N.C. [91]

4.4.2 Validity of the Weibull Model

The validity of the Weibull model to model the maximum drug concentration was examined using the goodness-of-fit tests by using Cramer-von-Mises and Anderson-Darling tests. The appropriateness of replacing the three-parameter Weibull with the two-parameter Weibull, corresponding to the location parameter $\mu = 0$, was also assessed using the likelihood ratio test statistic.

4.5 Generalized Pareto Distribution

The generalized Pareto distribution (GPD) is a two-parameter family of distributions which can be used to model exceedances over a threshold. The most common estimators for the shape and scale parameters of the GPD are maximum likelihood estimators. The probability density function for the GPD is given by

$$f(x;\lambda,\sigma)=\frac{1}{\sigma}\left(1-\frac{\lambda x}{\sigma}\right)^{\frac{1}{\lambda}-1} \tag{4.18}$$

where λ and σ are the shape and the scale parameters, respectively.

One of the crucial factors is to select an appropriate threshold value. The two techniques that are used to select a threshold value are the use of mean residual life plots, or fit the generalized Pareto distribution at a range of thresholds and look for stability of parameter estimates. Above a certain value of the threshold u_0 at which the generalized Pareto distribution provides a valid approximation to the excess distribution, the mean residual plot should be approximately linear in u. By arbitrarily selecting a range of threshold values, the mean residual plots drawn did not display linearity, indicating the inappropriateness of the use of the generalized Pareto distribution to model the maximum drug concentration data.

4.6 Results of the Modeling on Large Samples

4.6.1 Assessment of Generalized Extreme Value (GEV) Distribution

For large samples of sizes 300, 400, 500 and 1000, the generalized extreme value distribution was fit. The various diagnostic plots for assessing the accuracy of the GEV model fitted to the maximum drug concentration data were carefully examined. Both the probability and quantile plots were linear or near-linear indicating no evidence of lack of fit to the GEV distribution. The estimated shape parameter ξ was negative, and as a consequence of this, the return level curve asymptotes to a finite level. The fitted density seems a reasonable fit to the histogram of the maximum drug concentrations. Thus, all four diagnostic plots supported the fitted GEV model. The Table 4.1 displays maximum likelihood estimates for the GEV model along with their standard errors, and Deviance statistic for each of the samples of sizes 300, 400, 500, and 1000. The Figure 4.1 displays the diagnostic plots for the case of n = 1000. Figures 4.2 through 4.4 display the return level graphs with 90%, 95%, and 99% confidence intervals, respectively.

Table 4.1 Parameter Estimates (SE) for the GEV Model

n	μ (SE)	σ (SE)	ξ (SE)	DEVIANCE
1000	1360.81 (10.6)	303.45 (7.5)	-0.095 (0.020)	14468.9
500	1363.00 (15.2)	305.27 (10.7)	-0.082 (0.029)	7250.02
400	1346.68 (16.2)	291.44 (11.5)	-0.075 (0.033)	5765.8
300	1355.65 (21.4)	330.03 (15.2)	-0.112 (0.041)	4388.8

In the above table $\mu, \sigma,$ and ξ are the location, the scale and the shape parameters of the GEV. Using the parameter estimates and their corresponding standard errors one can construct 95% (90% or 99%) confidence interval. A greater accuracy in the computation of confidence intervals can be obtained by the use of profile likelihood.

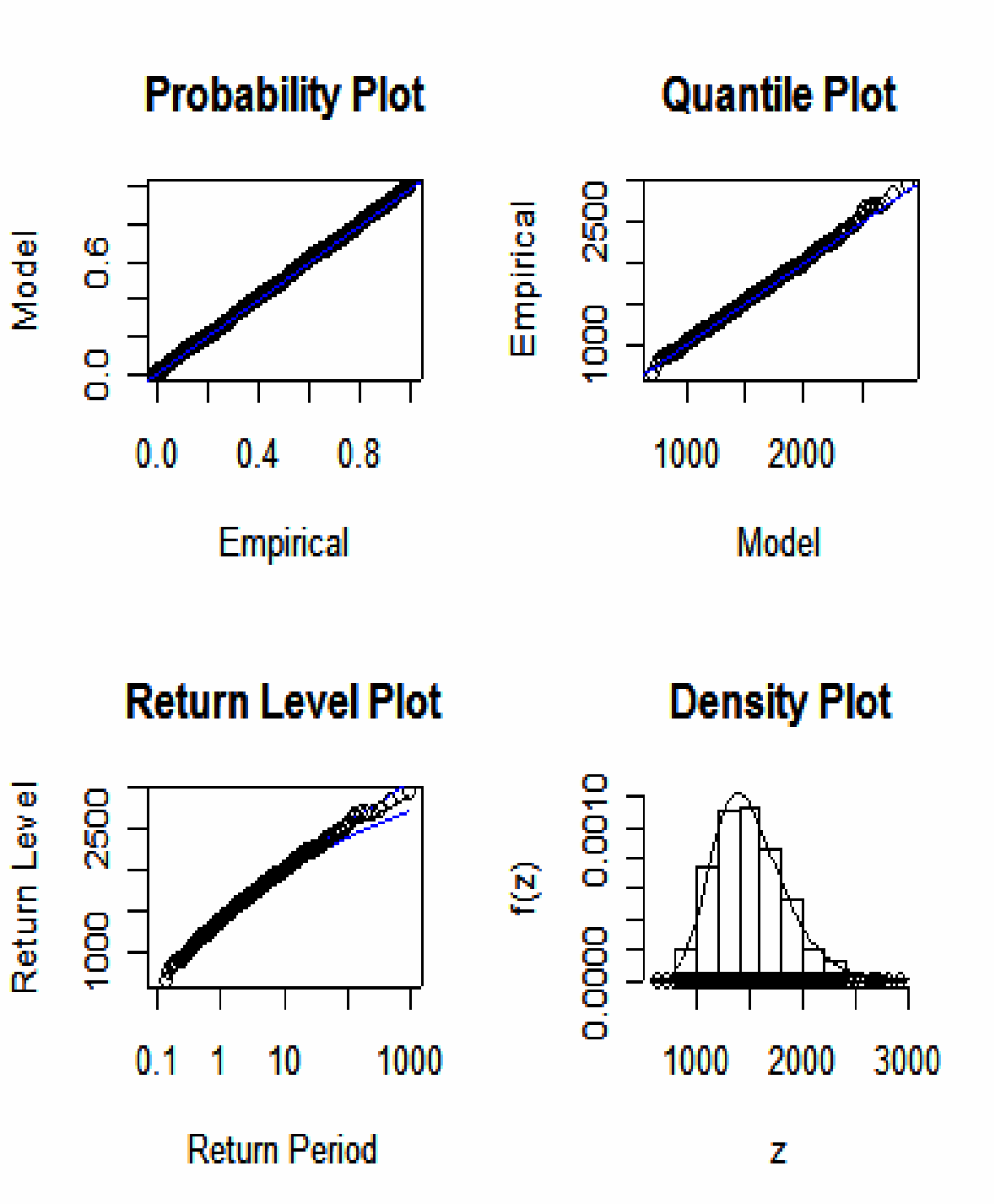


Figure 4.1 Diagnostic Plots for GEV $(n = 1000)$

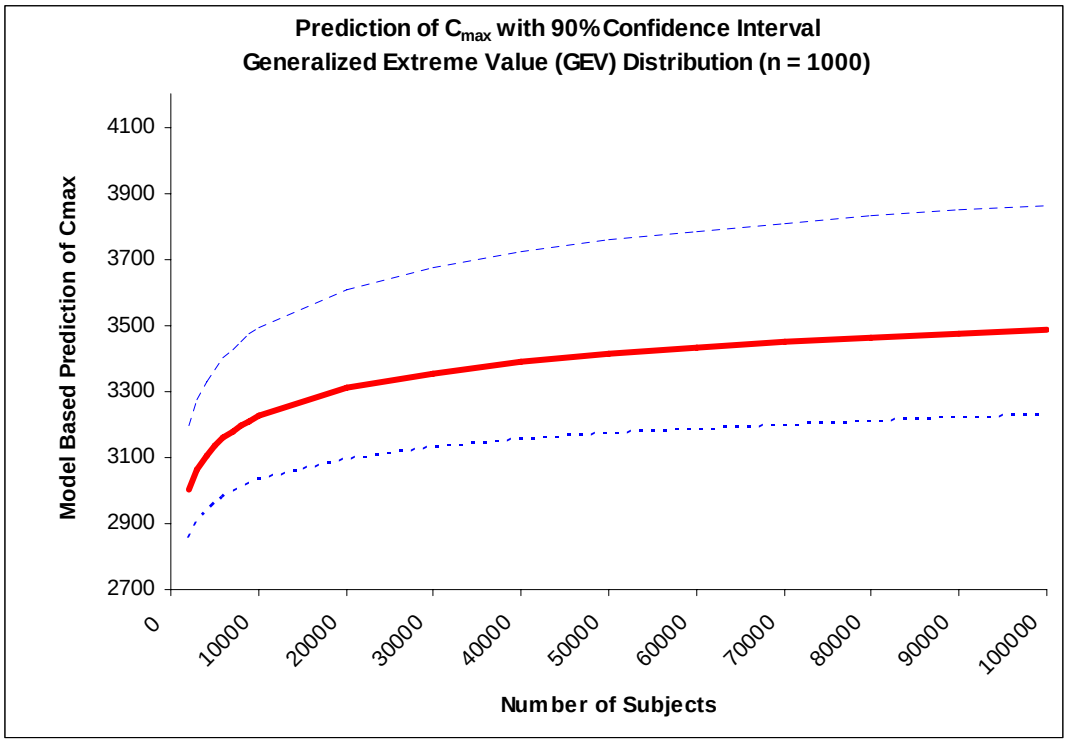


Figure 4.2 Prediction of C_{max} versus Number of Subjects with 90% Confidence Interval

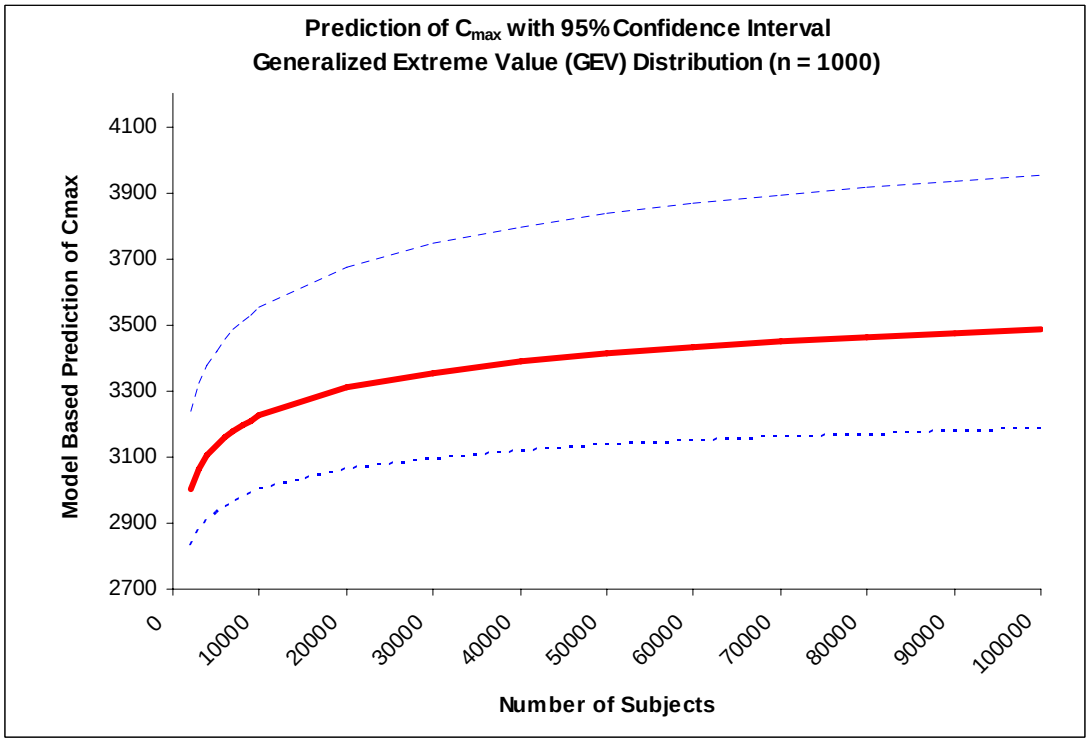


Figure 4.3 Prediction of C_{max} versus Number of Subjects with 95% Confidence Interval

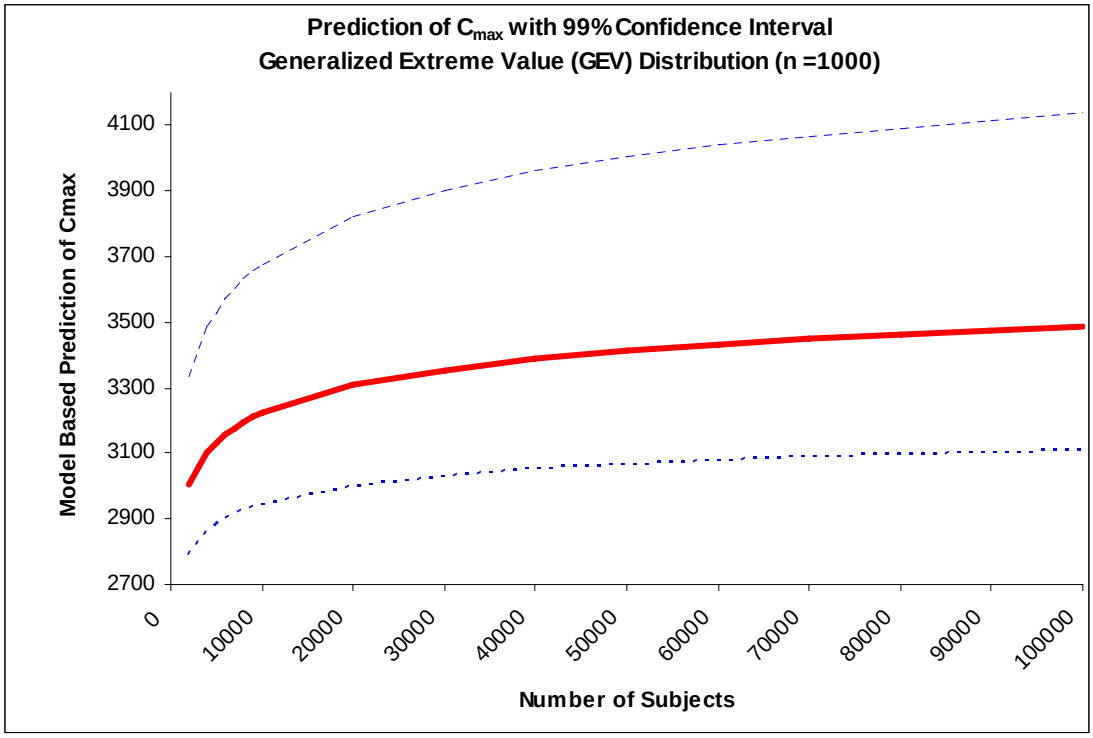


Figure 4.4 Prediction of C_{max} versus Number of Subjects with 99% Confidence Interval

4.6.2 Assessment of Gumble Model

Next, the suitability of replacing the GEV family with Gumbel family ($\xi = 0$) was assessed.

Table 4.2 Parameter Estimates (SE) for the Gumbel Model

n	μ (SE)	σ (SE)	DEVIANCE	DEVIANCE (GEV)	DF	$^{*}\chi^2$
1000	1344.58 (10.1)	296.25 (7.2)	14487	14469	1	18
500	1349.6 (14.28)	299.56 (11.2)	7257	7250	1	7
400	1334.66 (15.2)	285.95 (11.1)	5770	5766	1	4
300	1336.09 (19.5)	320.46 (17.3)	4395	4389	1	6

$^{*}p$-values < 0.05

The p-value corresponding to each of the four χ^2 values presented in Table 4.2 is less than 0.05, leading to the rejection of the null hypothesis that the shape parameter is 0.

The degree of significance increases as the sample size increases. This leads to the conclusion that the Gumbel model is not a suitable model to model the maximum drug concentration data when the sample size is large.

4.6.3 Assessment of Weibull Model

Since the value of the shape parameter is negative, the suitability of replacing the GEV model with the Weibull model was assessed. Although the Weibull model was suitable for small samples with its shortcomings to adequately explain the bimodal nature of the data, its suitability to model the large sample data was inconsistent. Table 4.3 displays the parameter estimates for the Weibull fit.

Table 4.3 Parameter Estimates (SE) for the Three-Parameter Weibull Model

n	μ (SE)	σ (SE)	ξ (SE)	$p-value^*$
1000	624.54 (9.55)	993.54 (16.24)	2.68 (0.07)	0.004
500	775.20 (19.9)	835.42 (29.38)	2.18 (0.10)	0.014
400	704.72 (17.1)	889.85 (27.3)	2.43 (0.11)	0.064
300	774.5 (36.4)	833.9 (49.6)	2.06 (0.16)	0.47

*p-value corresponding to χ^2 goodness-of-fit test

The inconsistency of the suitability of the Weibull model stems from the fact that the p-value is statistically more significant as the sample size becomes large, leading to the rejection of the null hypothesis that the Weibull model fits the maximum drug concentration data when the sample size becomes large.

4.6.4 Assessment of Generalized Pareto Model

Also examined was the suitability of the generalized Pareto distribution to model the maximum drug concentration data. The results from this approach indicated that the Pareto distribution was an inadequate distribution to model the data. Table 4.4 displays the parameter estimates for the generalized Pareto fit.

Table 4.4 Parameter Estimates (SE) for the generalized Pareto Model

n	σ (SE)	ξ (SE)	$p-value^{*}$
1000	636.7 (16.9)	1.196 (0.14)	0.0
500	800.4 (21.9)	1.634 (0.19)	0.0
400	724.5 (25.3)	1.432 (0.22)	0.0
300	835.8 (27.2)	1.774 (0.23)	0.0

*p-value corresponding to χ^2 goodness-of-fit test

The goodness-of-fit test used was Chi-square, and the p-values were statistically significant, leading to the rejection of the null hypothesis that the Pareto distribution is an appropriate distribution for any of the large sample sizes 300, 400, 500, or 1000.

4.7 Return Level Profiles Based on the Quartiles

The complete data set (n =1000) was split into four groups based on the quartiles. The first quartile value is 1249.32. The second and third quartile values are 1463.41 and 1710.56 respectively. The first group, named as Q1, consists of all values of maximum drug concentrations that are less than or equal to 1249.32. The second group, named as Q2, consists of all values of maximum drug concentrations that are less than or equal to 1463.41. The third group, named as Q3, consists of all values of maximum drug concentrations that are less than or equal to 1710.56. The last group, named as ALL, consists of all 1000 original maximum drug concentration values. The generalized extreme value (GEV) distribution was fit into each of these groups and maximum likelihood parameters were estimated. The following table displays the parameter estimates along with the standard errors for each of the four groups.

Table 4.5 Parameter Estimates (SE) for the GEV fit on four groups of data.

n	μ (SE)	σ (SE)	ξ (SE)
Q1	1072.02 (13.22)	166.72 (12.44)	-0.9404 (0.00)
Q2	1163.85 (11.80)	208.79 (13.21)	-0.6847 (0.04)
Q3	1269.82 (9.67)	223.78 (7.09)	0.4497 (0.03)
ALL	1360.81 (10.63)	303.45 (7.51)	0.0950 (0.02)

The graphs of the quantiles derived from the maximum likelihood estimators for each of the four groups are displayed in a single graph in Figure 4.5. If the maximum value of maximum drug concentrations that can be attained is limited by these quartile values that define the groups Q1, Q2, and Q3, then the quantile graphs for the groups Q1,

Q2, and Q3 are almost linear as n becomes a very large number. The graph is displayed in Figure 4.5.

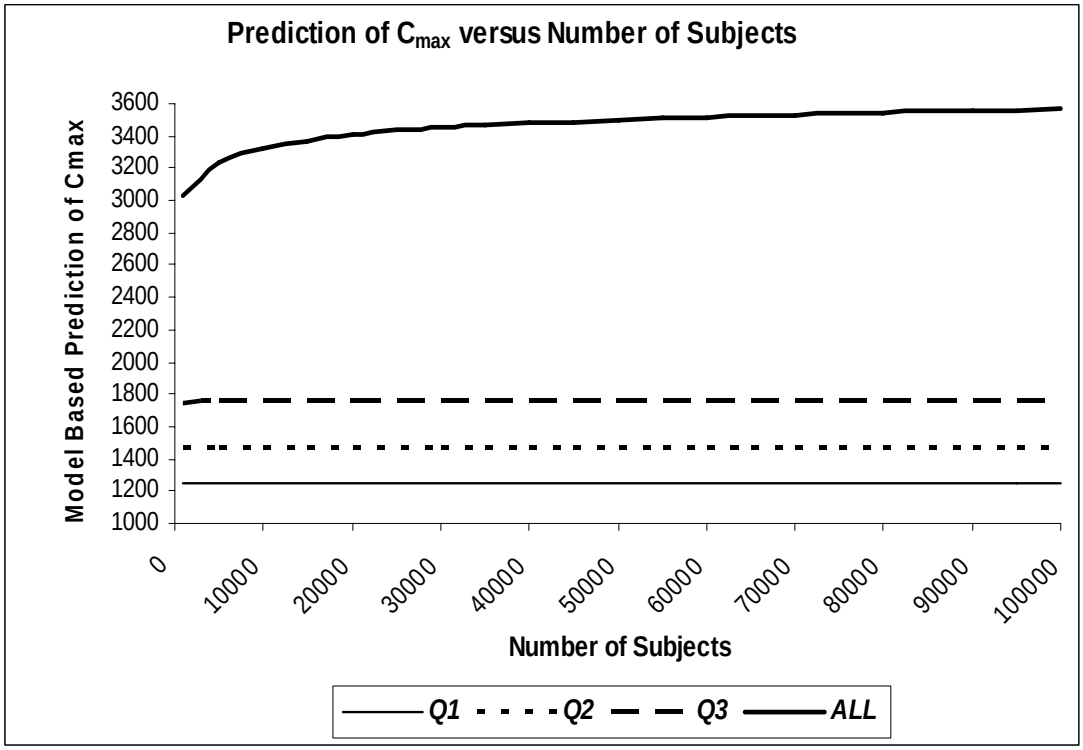


Figure 4.5 Prediction of C_{max} versus Number of Subjects from the Maximum Likelihood Estimates of the GEV Model

4.8 Summary and Conclusions

The results from the application of extreme value theory to model the maximum drug concentration data for large samples are summarized in the form of tables in section 4.6. The Gumbel distribution (testing the $H_0 : \xi = 0$) is an inadequate model based on χ^2 goodness-of-test's *p*-value shown in Table 4.2. A similar conclusion can be drawn for the 3-parameter Weibull because of its inconsistency in fitting to the data as sample size increases. This is evident from the *p*-values shown in Table 4.3 for different values of *n*. Table 4.4 tabulates the *p*-values corresponding to the Pareto fit to the data for

different sample sizes. From this, it is easy see that the Pareto distribution is an inadequate model to fit to the maximum drug concentration data.

From the model diagnostic plots for the generalized extreme value distribution (GEV) shown in Figure 4.1, and the deviance statistic shown in Table 4.1, it is evident that GEV is the only suitable model to model the maximum drug concentration data for large samples.

Figures 4.2, 4.3, and 4.4 show the quantile graphs obtained using the maximum likelihood estimators from the GEV model with 90%, 95%, and 99% confidence intervals, respectively. The usefulness of these graphs is that one can predict the maximum drug concentration value in a given patient after the administration of the drug with 90%, 95%, and 99% confidence limits.

Using the extreme value theory to model and predict $C_{\max}$ has not appeared in the literature prior to this study. The behavior and the effect of any drug in a human body are influenced by gender, age, weight, interactions with other drugs, severity of the disease, and other factors which are unique to a particular drug. The methodology developed in this study can easily be applied to construct profiles of $C_{\max}$ of a drug for each covariate or combination of covariates.

Chapter Five

Statistical Modeling of a Pharmacokinetic System

5.0 Introduction

In the pharmaceutical drug discovery and development process, mathematical modeling, computer-intensive simulations and statistical validations have become an integral part of the process. For any chemical compound to be useful as a drug it must exhibit a sufficient binding to the target receptor, that is, *it must have efficacy at the site of action*. In addition, sufficient quantities of the compound must reach the site of action and remain there long enough for the desired therapeutic effect, that is, *it must have desirable pharmacokinetics*. In order to accomplish these two goals it is necessary that the drug be absorbed into the body, be distributed to the site of action, and remain in the body long enough for its benefits to emerge. Each of these areas is subject to research in pharmaceutical drug development and requires extensive mathematical modeling.

Mathematical modeling and the resulting computer simulations have become part of a successful method of doing scientific research by complementing and blending experiment and theory [13, 14, 28]. Mathematical/numerical modeling not only offers many advantages but also aids and accelerates research and development by predicting the response of the system under different, often critical, conditions, giving insight into processes over a time scale, and eliminating the need of doing expensive tests on unsuccessful designs.

In multi-compartment pharmacokinetic models the two rate constants that are of importance are the *absorption rate* constant and the *elimination rate* constant. The absorption rate constant expresses the speed of absorption and affects such pharmacokinetic parameters as the maximum drug concentration, time at which the maximum concentration occurs, and the area under the concentration-time curve. The average drug concentration during a long-term therapy depends on the extent of absorption, and the degree of fluctuation is related to the absorption rate constant. The elimination rate constant is a function of how a drug is cleared from the blood by the eliminating organs and how the drug distributes throughout the body [93]. As a result of increasing interest in the kinetics of drug absorption, distribution, and elimination, many analytical and mathematical techniques have been developed to perform highly sophisticated pharmacokinetic analyses [37].

The purpose of mathematical models in biology and medicine is to obtain a comprehensible representation of complex biological activities. Models are based on the experiments, and experiments do contain errors of random nature. In addition, biological systems have intrinsic variability. These factors indicate the importance of stochastic approach to formulations of mathematical models in medicine and biology [63, 64].

The use of compartment models in pharmacokinetics was first proposed by Teorell [92] in 1937, and ever since, these models have been used extensively for the study of drug concentration problems in pharmacokinetics [39, 40]. An *n*-compartment model is described by the following system of differential equations:

$$\dot{x}(t) = -k_{ii}x_i(t) + \sum_{\substack{j=1 \\ i \neq j}}^{n} k_{ji}x_j(t), \quad j,i = 1,2,3,.....,n \tag{5.0.1}$$

with initial conditions $x_i(0) = x_{i0} \geq 0$, where $x_i(t)$ is the drug concentration in the i^{th} compartment; $k_{ij}\,(i,\ j = 1,2,3,.....,n)$ are the rate constants which represent the proportional constants associated with the rate of absorption or elimination from the compartments. The rate constants are non-negative and satisfy the relationship

$$k_{ii} = \sum_{i \neq j} k_{ij}\,,\ i,\ j = 1,2,3,.....,n\,.$$

Assuming that the rate constants are known, the system of differential equations, (5.0.1), can be solved to obtain a theoretical description of the drug concentration as a function of time in each compartment. But the complicating factor in this problem is the uncertainty in the rate constants due to factors such as varied experimental conditions like sampling and measurement errors, variations in patient parameters as functions of time and environmental effects. These uncertainties naturally lead to the assumption that the rate constants behave as stochastic variables, and hence, the use of stochastic procedures in the analysis is appropriate [86]. A well constructed model in a stochastic setting will converge in the mean to the deterministic solution.

By considering the initial concentration or rate constants or both to be stochastic, the system of differential equations, (5.0.1), can be made stochastic. One can then seek to develop the probability density function that will characterize the random behavior of drug concentration. Having the probability distribution one can obtain the mean behavior of the drug concentration and compare it to its deterministic counterpart. Tsokos *et al*. [97] have discussed in detail the procedure for stochastizing the pharmacokinetic model that describes the profile of an antibiotic drug coumermycin A_1, and have given a

numerical comparison between the deterministic trajectory of the above drug with the mean value of the random solution as a function of time.

5.1 Focus of the Chapter Five

The focus of the present study is to develop numerical solutions for a system of random differential equations that represents the three-compartment open system which describes the disposition of coumermycin A_1 that was considered by Tsokos *et al.* [97]. Schematically, the open three-compartment pharmacokinetic model that we are studying is illustrated in Figure 5.2.1 and the system of differential equations describing the model is given by the equation 5.2.1. The rate constants used in the system are assumed to be random and have been characterized by a trivariate probability distribution. The trivariate probability distributions that will be used to simulate these rate constants are the truncated trivariate normal and exponential probability distributions.

More precisely, the numerical solutions for the system of random differential equations will be obtained using the rate constants that are simulated from the trivariate probability distribution under different combinations of covariance structures and initial conditions, where the initial conditions are non-random.

The effects of different values of covariance structures and initial conditions on the deterministic behavior of the drug concentration and the mean behavior of the random solutions as a function of time will be discussed in addition to comparing these two behaviors. Also discussed is the suitability of the use of the two probability distributions to simulate the rate constants.

5.2 A Pharmacokinetic Model for the Antibiotic Drug, Coumermycin A_1

In the model for the disposition of coumermycin A_1 the first compartment reflects solely the plasma water, while the second compartment comprises the remaining body distribution space and the third compartment reflects the biotransformation and excretion of the drug [57]. In designing a model for the disposition of coumermycin A_1, the additional factors to be considered are the non-elimination of the drug from the central or plasma compartment. Therefore biotransformation is occurring elsewhere in the body. The appropriate model for the disposition of coumermycin A_1, introduced by Kaplan [57], is as follows:

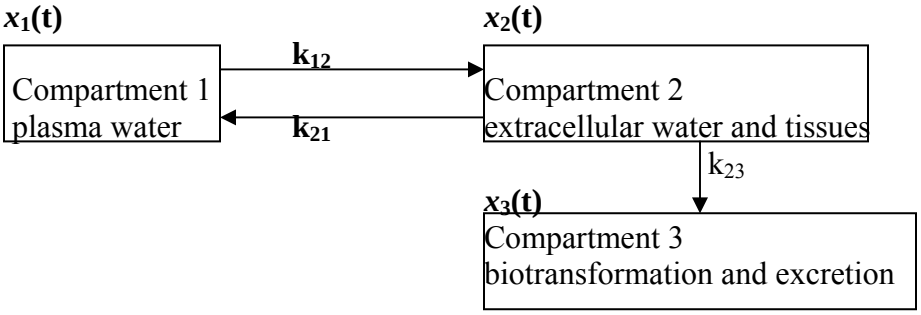


Figure 5.2.1: Model for the disposition of Coumermycin A_1

The pharmacokinetic profile of coumermycin A_1 was determined in four human subjects based on the serum level data from the report of a clinical study [85]. The associated rate constants were calculated [57] and are presented in Table 5.2.1 below.

Table 5.2.1 Rate constants for the disposition of coumermycin A_1

SUBJECT	SUBJECT A	SUBJECT B	SUBJECT C	SUBJECT D
k_{12}, hr^{-1}	0.338	0.297	0.273	0.222
k_{21}, hr^{-1}	0.060	0.097	0.030	0.033
k_{23}, hr^{-1}	0.041	0.065	0.027	0.017

The above compartmental model is described by the following system of differential equations:

$$\begin{aligned} \dot{x}_1(t) &= -k_{12}x_1(t) + k_{21}x_2(t) \\ \dot{x}_2(t) &= k_{12}x_1(t) - (k_{21}+k_{23})x_2(t) \\ \dot{x}_3(t) &= k_{23}x_2(t) \end{aligned} \qquad (5.2.1)$$

with the initial conditions $x_1(0) = c_1$, $x_2(0) = c_2$, $x_3(0) = c_3$; where $x_1(t)$, $x_2(t)$, and $x_3(t)$ are the amounts of the drug concentration in compartments 1, 2, and 3, respectively, at time $t \geq 0$; k_{12}, k_{21}, and $k_{23} > 0$ are the rate constants of the system.

In the matrix form, the system (3.1) can be written as follows:

$$\dot{x}(t) = \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} = \begin{bmatrix} -k_{12} & k_{21} & 0 \\ k_{12} & -(k_{21}+k_{23}) & 0 \\ 0 & k_{23} & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}, \text{ with initial conditions } C = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \qquad (5.2.2)$$

The solution to the differential system (5.2.1) with its corresponding initial conditions is given by Tsokos *et al*. [97] as follows,

$$x(t)=\begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix} \left(\frac{c_1\lambda_3}{k_{23}} - \frac{k_{21}\lambda_3 c_2}{k_{23}(\lambda_3+k_{12})}\right) a e^{\lambda_2 t} + \left(\frac{k_{21}\lambda_2 c_2}{k_{23}(\lambda_2+k_{12})} - \frac{\lambda_2 c_1}{k_{23}}\right) b e^{\lambda_3 t} \\ \left(\frac{c_1\lambda_3}{k_{23}} - \frac{k_{21}\lambda_3 c_2}{k_{23}(\lambda_3+k_{12})}\right) c e^{\lambda_2 t} + \left(\frac{k_{21}\lambda_2 c_2}{k_{23}(\lambda_2+k_{12})} - \frac{\lambda_2 c_1}{k_{23}}\right) d e^{\lambda_3 t} \\ c_1\left(\frac{\lambda_2-\lambda_3}{k_{23}}\right) - \frac{k_{21}\lambda_2(k_{23}c_2-\lambda_3 c_3)}{k_{23}^2(\lambda_2+k_{12})} + \frac{k_{21}\lambda_3(k_{23}c_2-\lambda_2 c_3)}{k_{23}^2(\lambda_3+k_{12})} \\ + \left(\frac{c_1\lambda_3}{k_{23}} - \frac{k_{21}\lambda_3 c_3}{k_{23}(\lambda_3+k_{12})}\right) e^{\lambda_2 t} + \\ \left(\frac{k_{21}\lambda_2 c_2}{k_{23}(\lambda_2+k_{12})} - \frac{\lambda_2 c_1}{k_{23}}\right) e^{\lambda_3 t} \end{bmatrix} \tag{5.2.3}$$

where $\lambda_1 = 0,$

$$\lambda_2 = 0.5 * \left(-(k_{21}+k_{23}+k_{12}) + \sqrt{\left((k_{21}+k_{23}+k_{12})^2 - 4k_{12}k_{23}\right)} \right),$$

$$\lambda_3 = 0.5 * \left(-(k_{21}+k_{23}+k_{12}) - \sqrt{\left((k_{21}+k_{23}+k_{12})^2 - 4k_{12}k_{23}\right)} \right),$$

$$a = \frac{k_{21}\lambda_2}{k_{23}(\lambda_2+k_{12})}, \qquad b = \frac{k_{21}\lambda_3}{k_{23}(\lambda_3+k_{12})}, \; c = \frac{\lambda_2}{k_{23}}, \qquad d = \frac{\lambda_3}{k_{23}},$$

and $\Delta = \frac{k_{21}\lambda_2\lambda_3}{k_{23}^2}\left\{\frac{\lambda_3-\lambda_2}{(\lambda_2+k_{12})(\lambda_3+k_{12})}\right\}.$

In the matrix form, the general solution can be written as:

$$x(t) = \frac{1}{\Delta} AC \text{ where } \quad C = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \text{ and}$$

A is a $3x3$ matrix whose three columns *A.1*, *A.2*, and *A.3* are given as follows:

$$A.1 = \begin{bmatrix} \frac{1}{k_{23}}\left(\lambda_3 a e^{\lambda_2 t} - \lambda_2 b e^{\lambda_3 t}\right) \\ \frac{1}{k_{23}}\left(\lambda_3 c e^{\lambda_2 t} - \lambda_2 d e^{\lambda_3 t}\right) \\ \frac{1}{k_{23}}\left(\lambda_2 - \lambda_3 + \lambda_3 e^{\lambda_2 t} - \lambda_2 e^{\lambda_3 t}\right) \end{bmatrix} \qquad A.2 = \begin{bmatrix} -\frac{k_{21}}{k_{23}}\left(\frac{\lambda_3 a e^{\lambda_2 t}}{\lambda_3 + k_{12}} - \frac{\lambda_2 b e^{\lambda_3 t}}{\lambda_2 + k_{12}}\right) \\ -\frac{k_{21}}{k_{23}}\left(\frac{\lambda_3 c e^{\lambda_2 t}}{\lambda_3 + k_{12}} - \frac{\lambda_2 d e^{\lambda_3 t}}{\lambda_2 + k_{12}}\right) \\ \left[\frac{k_{21}\lambda_2}{k_{23}\left(\lambda_2 + k_{12}\right)}\left(-1 + e^{\lambda_3 t}\right) + \frac{k_{21}\lambda_3}{k_{23}\left(\lambda_3 + k_{12}\right)}\right] \end{bmatrix}$$

$$A.3 = \begin{bmatrix} 0 \\ 0 \\ \frac{k_{21}\lambda_2\lambda_3}{k_{23}^2\left(\lambda_2 + k_{12}\right)} - \frac{k_{21}\lambda_3}{k_{23}\left(\lambda_3 + k_{12}\right)}\left(\frac{\lambda_2}{k_{23}} + e^{\lambda_2 t}\right) \end{bmatrix}$$

Tsokos *et. al.* [97] compared the deterministic behavior of the concentration of coumermycin A_1 and the mean behavior of the random solution as a function of time with the assumption that initial concentrations being $c_1 = 1, c_2 = 0$ and $c_3 = 0,$ and the joint probability density function of random vector k being probabilistically characterized by

$$f(k;t) = \begin{Bmatrix} D_0\left[2\pi\left|\Sigma\right|^{\frac{1}{2}}\right]^{-1} \exp\left[-\frac{1}{2}(k-\mu)^T \Sigma^{-1}\left((k-\mu)\right], & k \geq 0 \\ 0 & \text{elsewhere} \end{Bmatrix},$$

where $k = \begin{bmatrix} k_{12} \\ k_{21} \\ k_{23} \end{bmatrix}, \quad \mu = \begin{bmatrix} E(k_{12}) \\ E(k_{21}) \\ E(k_{23}) \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 & \rho\sigma_1\sigma_3 \\ \rho\sigma_2\sigma_1 & \sigma_2^2 & \rho\sigma_2\sigma_3 \\ \rho\sigma_3\sigma_1 & \rho\sigma_3\sigma_2 & \sigma_3^2 \end{bmatrix}$

and D_0 is a normalization factor.

5.3 Simulation of Rate Constants Using Truncated Trivariate Normal Distribution

The importance of rate constants is quite evident as they affect such pharmacokinetic parameters as maximum drug concentration, time at which this maximum concentration is attained and the area under the concentration-time curve as well as how a drug is cleared from the blood and distributed through out the body.

Chaung and Lloyd [6] specifically mention that the rate constants are correlated, but the presumption of their independence or of having some specific correlation is necessary in order that an analysis of the problem to be relevant. Thus, we study the drug concentration behavior in a stochastic setting with respect to small (0.25), medium (0.50) and high (0.75) correlation structure.

The standard deviation in any type of pharmacokinetic system plays a major role in the final interpretation of the drug concentration behavior in each of the compartments. Thus, we study its effect on the drug concentration behavior with respect to small (5% of the mean), medium (10% of the mean), and large (20% of the mean) increase in the standard deviation.

The two necessary components, the mean vector and the covariance matrix, to generate rate constants that are truncated trivariate normally distributed are obtained as follows:

Mean Vector: The first entry in the mean column vector was computed as the average of k_{12} values measured on four patients, the second entry was the average of k_{21} values measured on the same four patients, and the third entry was the average of k_{23} values measured on the same four patients.

Covariance Matrix: The values used for the correlation coefficient and standard deviations to obtain the covariance matrix structure are described in Figure 5.3.1 in the form of a flow-chart. It also describes the set of initial conditions used in solving the system of random differential equations.

The numerical solutions for the system of random differential equations are obtained for the following three sets of initial conditions:

1. vector $(c_1, c_2, c_3) = (1.0, 0.00, 0.00)$
2. vector $(c_1, c_2, c_3) = (1.0, 0.05, 0.05)$
3. vector $(c_1, c_2, c_3) = (1.0, 0.10, 0.10)$.

For each set of initial conditions, three correlation coefficients (0.25, 0.50 and 0.75) and three values for the standard deviation (5%, 10% and 20%), represented as a percentage of the mean value, will be considered. Thus, numerical solutions will be obtained for a total of 3•3•3 = 27 models. All the numerical solutions summarized in this study are based on 1000 simulations.

It is helpful to have a convenient notation to refer to individual models. For example, the notation for the model consisting of the third set of initial conditions, correlation = 0.50 and standard deviation = 10% of the mean is M (3, 0.50, 10%). Similarly, M(3, 0.25, 5%) refers to the model consisting of the third set of initial conditions, correlation = 0.25 and standard deviation = 5% of the mean. Thus, a total of twenty seven model configurations that we studied are presented in Figure 5.3.1

In section 5.4 we present the representative graphs along with the discussion of the numerical results pertaining to the behavior of the deterministic characterization in

each of the three compartments as the correlation structure and standard deviation values change.

In section 5.5 we present the representative graphs along with the discussion of the numerical results pertaining to the behavior of the stochastic characterization of individual compartments as the correlation structure and standard deviation values change.

In section 5.6 we discuss the simulation of rate constants using the trivariate exponential probability distribution and the results from solving the system of random differential equations using these rate constants.

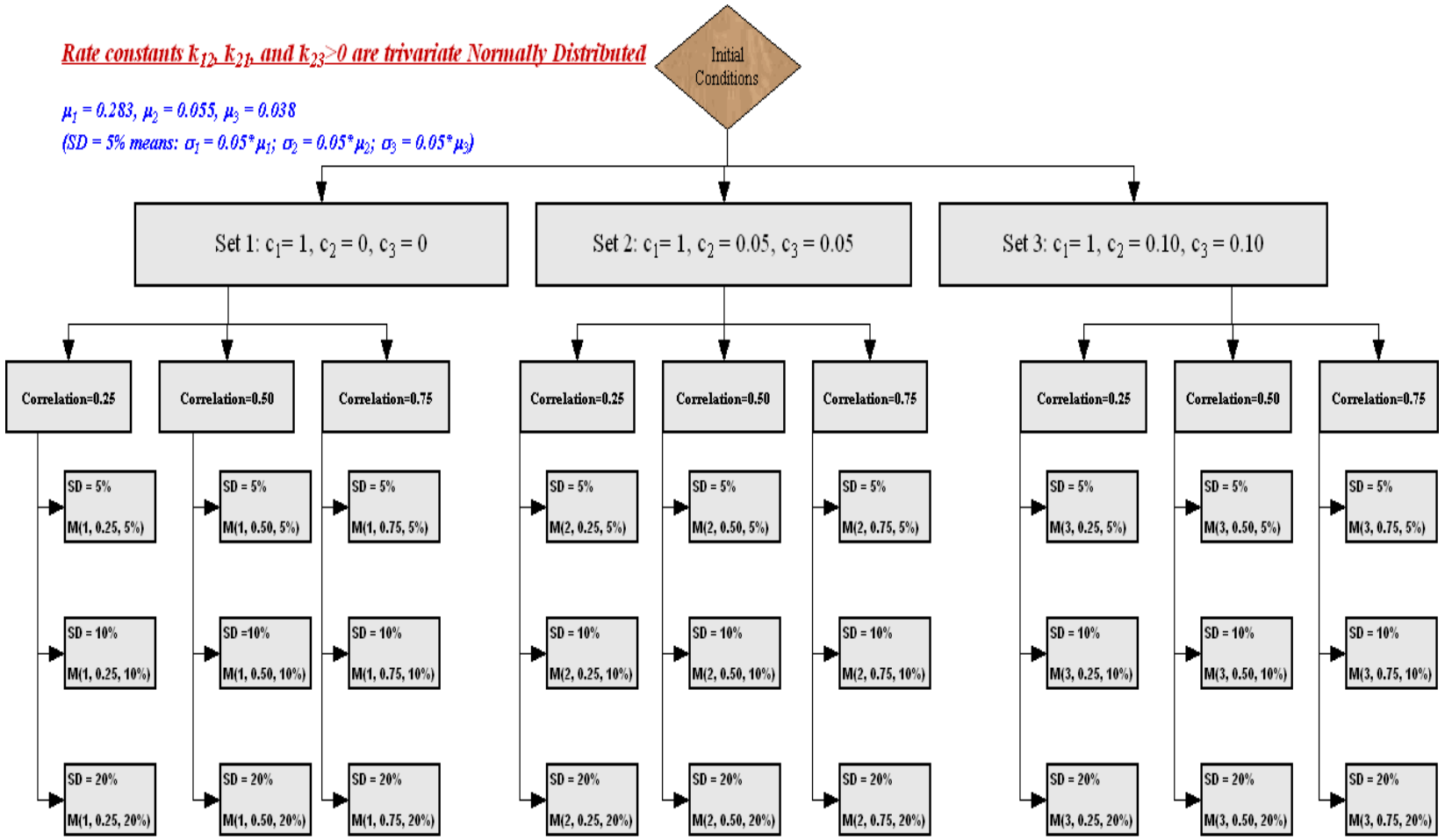


Table 5.3.1 Flow-chart describing all possible configurations in our numerical study of the pharmacokinetic system.

5.4 Discussion of the Results

The numerical results of the different models, as outlined in Figure 5.3.1, that we study are presented in Figures 5.4A through 5.4D.

Each of the three representative graphs in Figure 5.4A shows the deterministic behavior of coumermycin A_1 concentration in all three compartments for the first 50 hours. This behavior of the drug concentration in all three compartments is as expected. The drug concentration in the first compartment (*plasma water*) decays rapidly as indicated by the graph of $x_1(t)$, while the drug concentration in the second compartment (*extracellular water and tissues*) increases during the *first six to eight hours* as indicated by the graph of $x_2(t)$ before it starts decaying. The concentration behavior in the third compartment (*biotransformation and excretion*) increases exponentially as indicated by the graph of $x_3(t)$.

Discussion of the results displayed in the graphs in Figures 5.4B through 5.4D is given in Table 5.4.1. More precisely, the behavior of the deterministic characterization in each of the three compartments as is affected by the change in the correlation structure and increase in the standard deviation is displayed.

5.4.1 Effects of Varying the Values of the Standard Deviation and the Correlation Structure for Studying the Behaviors of $x_1(t)$, $x_2(t)$ and $x_3(t)$

Table 5.4.1 Effects of varying the value of the standard deviation and correlation structure for studying the behavior of $x_1(t)$, $x_2(t)$ and $x_3(t)$.

Change in Standard Deviation	**Change in Correlation**	
	0.25 to 0.50	0.25 to 0.75
5% to 10%		
$x_1(t)$:	Rate of decay is almost identical for the first 6 hours when r = 0.50, and after that, significantly slower rate of decay.	Rate of decay is significantly faster when r = 0.75 for the first 10 hours and after that, rates are almost identical.
$x_2(t)$:	Amount of absorption is significantly higher when r = 0.50 after the first 4 hours, and this difference increases with time.	Amount of absorption is higher for r = 0.75 for the first 7 hours, and after that, a significantly lower absorption.
$x_3(t)$:	Amount of biotransformation is significantly lower when r = 0.50 and this differences increases with time.	Amount of biotransformation is significantly higher when r = 0.75.
5% to 20%		
$x_1(t)$:	Rate of decay is significantly slower when r = 0.50 after the first 2 hours.	Rate of decay is significantly slower when r = 0.75 between 2 and 24 hours and after that, the two rates are almost identical.
$x_2(t)$:	Amount of absorption is slightly lower when r = 0.50 between 2 and 10 hours, and after that, higher amount of absorption.	Amount of absorption is significantly lower when r = 0.75 all through.
$x_3(t)$:	Amount of biotransformation is significantly lower when r = 0.50 and this differences increases with time.	Amount of biotransformation is almost identical for the first 20 hours and after that, it is slightly higher for r = 0.75.

5.4.2 Summary of the Effects of the Standard Deviation and the Correlation Structure on the Behaviors of $x_1(t)$, $x_2(t)$ and $x_3(t)$

Drug Concentration, $x_1(t)$

For a combination of changes in the standard deviation from 5% to 10% and the correlation from 0.25 to 0.50, the rate of decay is almost identical for the first six hours (delay effect) and then a significantly slower rate of decay. For correlation change from 0.25 to 0.75, rate of decay is faster for the first ten hours, and after that, almost identical.

For a combination of changes in the standard deviation from 5% to 20% and the correlation from 0.25 to 0.50, the rate of decay is slower after the first 2 hours, while it is significantly slower between 2 and 24 hours when the correlation changes from 0.25 to 0.75.

From these results we can conclude that, for the combination of higher values of the correlation and standard deviation, the rate of decay of drug concentration is significantly affected as shown in the sample graphs of $x_1(t)$ in Figure 5.4B.

Drug Concentration, $x_2(t)$

As standard deviation changes from 5% to 10% and correlation changes from 0.25 to 0.50, the amount of absorption of the drug concentration is significantly higher after the first four hours, while for the change in correlation from 0.25 to 0.75 it is higher for the first 7 hours (delay effect) and after that the amount of absorption is significantly lower.

For a combination of changes in the standard deviation from 5% to 20% and correlation from 0.25 to 0.50, the amount of absorption is lower between 2 and 10 hours but higher after that time point, whereas for the change from 0.25 to 0.75, the absorption is lower all through.

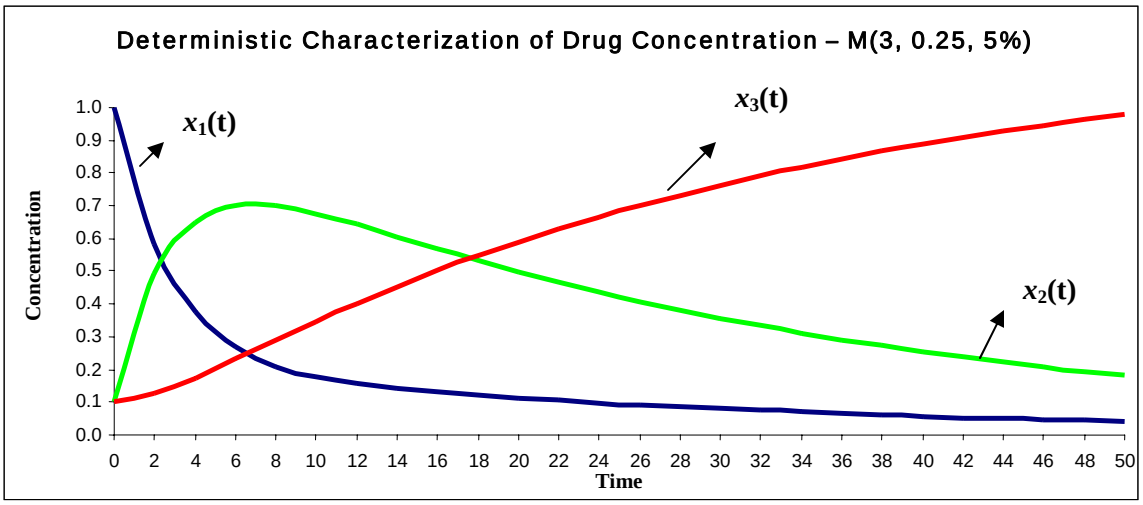
Hence, there is a significant effect of both standard deviation and correlation on the absorption rate as displayed in the sample graphs of $x_2(t)$ in Figure 5.4C. However, it is important to note this effect on the absorption rate as it affects such pharmacokinetic parameters as the maximum drug concentration, time at which the maximum drug concentration occurs, and the area under the concentration-time profile.

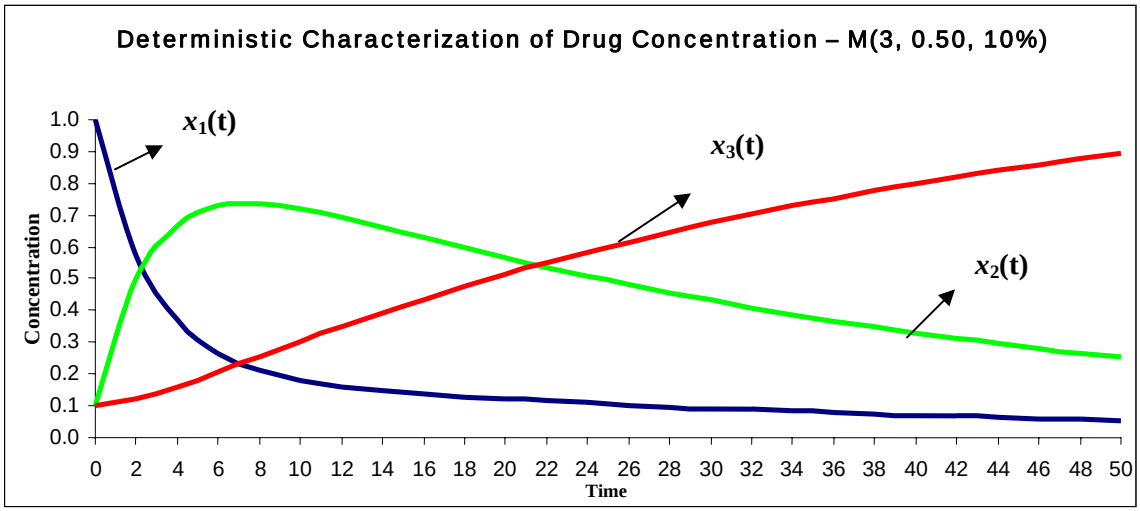
Drug Concentration, $x_3(t)$

For a combination of changes in the standard deviation from either 5% to 10% or 5% to 20%, and the correlation from 0.25 to 0.50, there is a significantly lower amount of biotransformation, and this difference increases with time.

When the correlation changes from 0.25 to 0.75, and standard deviation from 5% to 10%, the amount of biotransformation is significantly higher but for the change in standard deviation from 5% to 20%, the amount of biotransformation is higher after the first 20 hours after being almost identical during the first 20 hours (delay effect).

Thus, higher standard deviation and correlation values do have a significant effect on the biotransformation of the drug concentration as depicted by the sample graphs of $x_3(t)$ shown in Figure 5.4D.





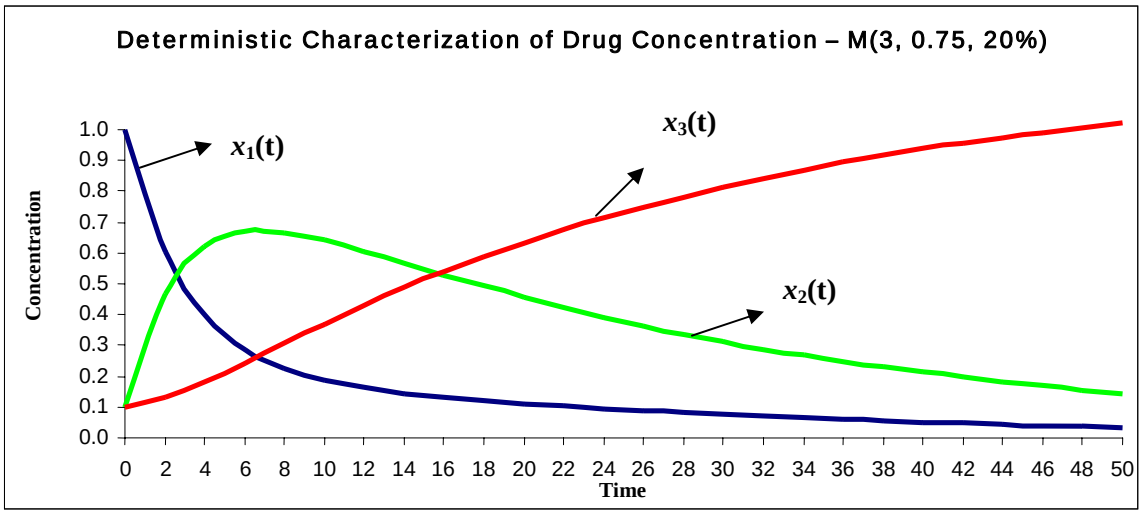


Figure 5.4A Deterministic characterization of drug concentration in all compartments

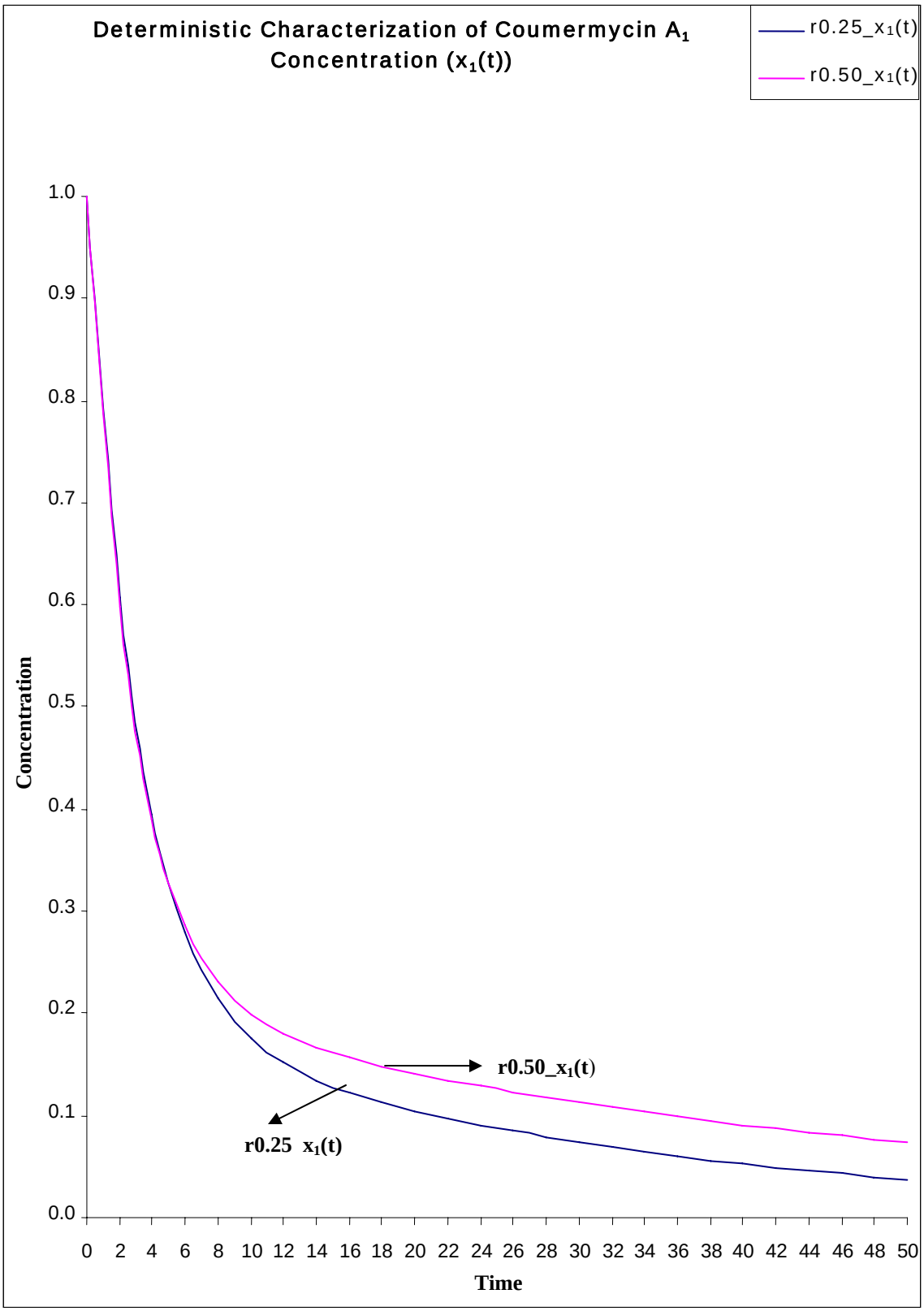


Figure 5.4B Deterministic characterization of coumermycin A_1 concentration ($x_1(t)$)

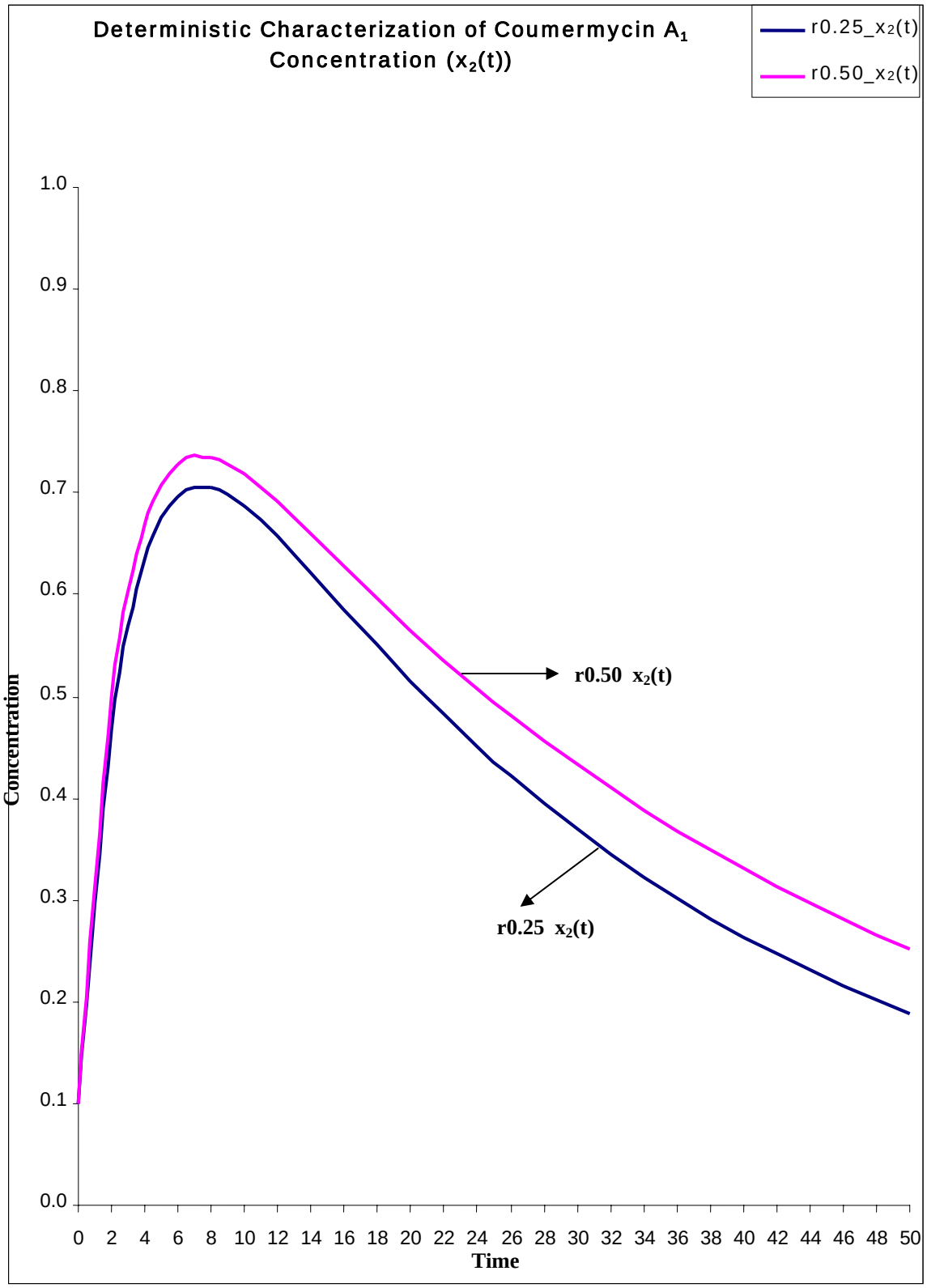


Figure 5.4C Deterministic characterization of coumermycin A_1 concentration ($x_2(t)$)

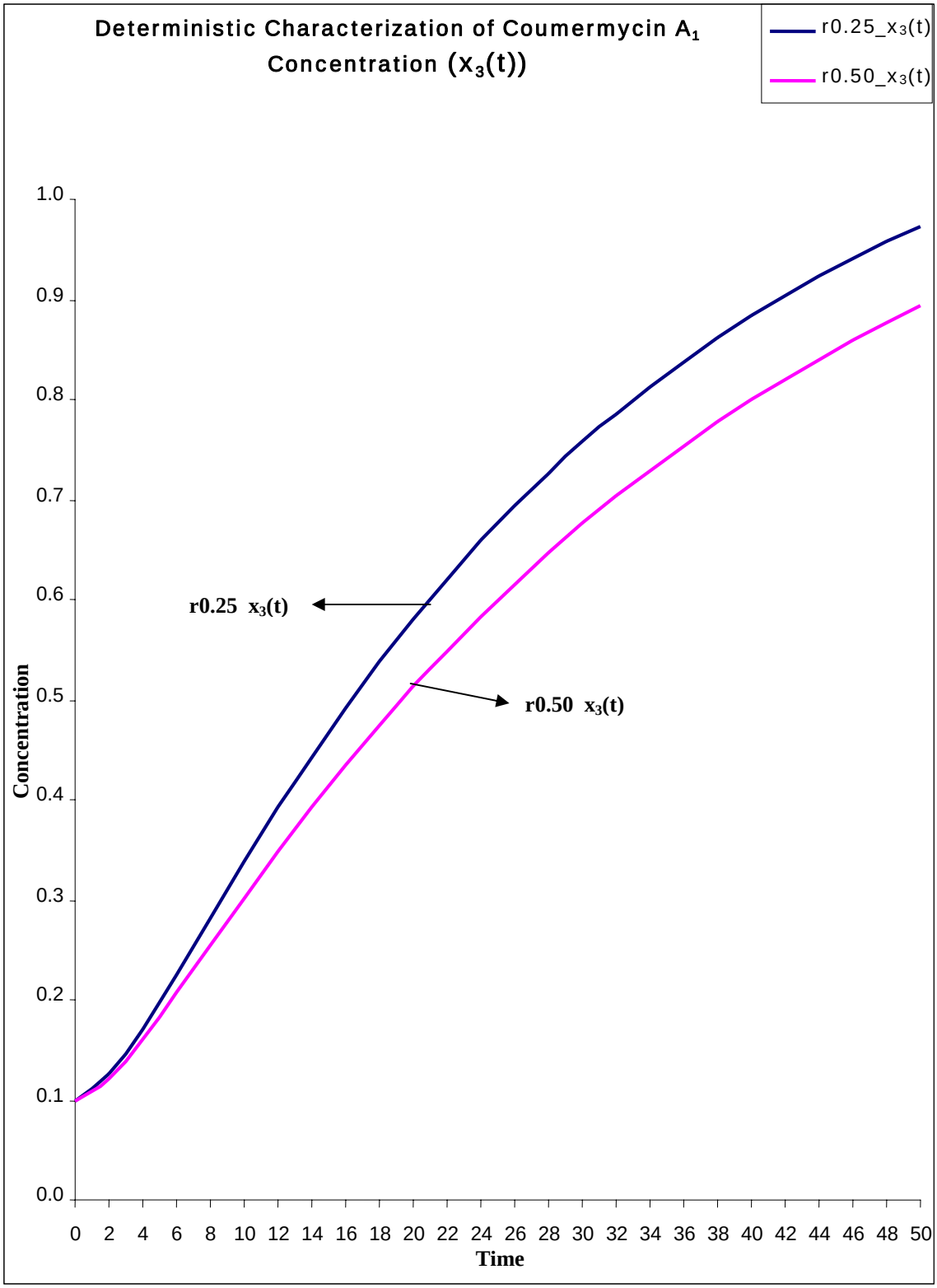


Figure 5.4D Deterministic characterization of coumermycin A_1 concentration ($x_3(t)$)

5.5 Stochastic Behavior of the Drug Concentration in Each Compartment

5.5.1 Effects of Varying the Values of the Standard Deviation and Correlation Structure on the Deterministic ($\boldsymbol{x_i(t)}$) and Stochastic ($\mathbf{E[x_i(t)]}$) Behaviors of the Individual Compartments

The stochastic results of the drug concentration behavior in each of the three compartments are summarized in Table 5.5.1 for varying values of correlation structure and varying increase in the standard deviation.

Table 5.5.1 Effects of varying the values of the standard deviation and correlation structure on the deterministic and stochastic behaviors of the individual compartments

Change in Standard Deviation	**Change in Correlation**	
	0.25 to 0.50	0.25 to 0.75
5% to 10%		
$x_1(t)$:	Both decrease rapidly but after the first 2 hours, the stochastic is consistently higher.	Both the stochastic and deterministic behaviors decrease but the stochastic is consistently higher.
$x_2(t)$:	Both deterministic and stochastic are almost identical for the first 4 hours, and after that, stochastic is significantly lower.	Both deterministic and stochastic are almost identical although the deterministic is slightly higher for the first 6 hours, and after that, stochastic is significantly higher.
$x_3(t)$:	Stochastic characterization is higher this difference increases with time.	Deterministic trajectory is higher and this difference increases with time.
5% to 20%		
$x_1(t)$:	Stochastic characterization is significantly slower than the deterministic trajectory.	Stochastic characterization is significantly slower than the deterministic trajectory.
$x_2(t)$:	Deterministic is higher for the first 8 hours and after that, the stochastic is significantly higher and this difference increases with time.	Deterministic is higher for the first 6 hours and after that, the stochastic is significantly higher and this difference increases with time.
$x_3(t)$:	Deterministic trajectory is significantly higher than the stochastic characterization after the first 6 hours, and this difference increases with time.	Deterministic is significantly higher than the stochastic characterization after the first 4 hours, and this difference increases with time.

5.5.2 Summary of the Effects of the Standard Deviation and the Correlation Structure on the Deterministic and the Stochastic Characterizations

Deterministic and Stochastic Characterizations of $x_1(t)$

For a combination of changes in standard deviation from 5% to 10% and the correlation from 0.25 to 0.50 (or 0.25 to 0.75), the stochastic characterization has a rate of decrease that is consistently slower than that of the deterministic trajectory. This slower rate of decrease of the stochastic characterization is even more pronounced as the standard deviation changes from 5% to 20% and correlation changes from 0.25 to 0.50 (or 0.25 to 0.75). This can be seen from the representative graph presented in Figure 5.5A.

More precisely, in Figure 5.5A, we note that after approximately 3.5 units of time, there is a difference between the deterministic and stochastic behaviors of the drug concentration in compartment one.

Deterministic and Stochastic Characterizations of $x_2(t)$

As the standard deviation changes from 5% to 10% and correlation changes from 0.25 to 0.50, from both characterizations being almost identical for the first 4 hours (delay effect), the stochastic is significantly lower than the deterministic after that point of time. But when the correlation changes from 0.25 to 0.75, after being almost identical for the first 6 hours, the stochastic behavior of $x_2(t)$ is significantly higher than the deterministic.

For a combination of changes in standard deviation from 5% to 20% and the correlation from 0.25 to 0.50, the deterministic behavior is higher for the first 8 hours (delay effect), but after that period the stochastic characterization is significantly higher with the difference increasing with time. However, when the correlation changes from 0.25 to 0.75, the deterministic is higher for the first 6 hours (delay effect), and after that,

the stochastic behavior of $x_2(t)$ is significantly higher with the difference increasing with time. This identifies the time delay of initial drug concentration to reach compartment two. This can be seen from the representative graph presented in Figure 5.5B.

Deterministic and Stochastic Characterizations of $x_3(t)$

For a combination of changes in standard deviation from 5% to 10% and the correlation from 0.25 to 0.50, the stochastic characterization is significantly higher with the difference increasing with time while, for the change in correlation from 0.25 to 0.75, the deterministic is significantly higher with the difference increasing with time.

As the standard deviation changes from 5% to 20% and correlation changes from 0.25 to 0.50, the deterministic is significantly higher after about approximately the first 6 hours (delay effect), and this difference increases with time. The same behavior repeats but after about approximately the first 3 hours when the correlation changes from 0.25 to 0.75. This can be seen in the representative graph presented in Figure 5.5C. This identifies the time delay of the drug concentration to reach compartment three from two.

Thus, the deterministic and the stochastic characterizations of the drug concentrations $x_1(t)$, $x_2(t)$, and $x_3(t)$ are significantly affected by the variations in the values of the standard deviation and the correlation structure. This significance in the effect is evident, especially, for the higher values of the standard deviation and the correlation structure.

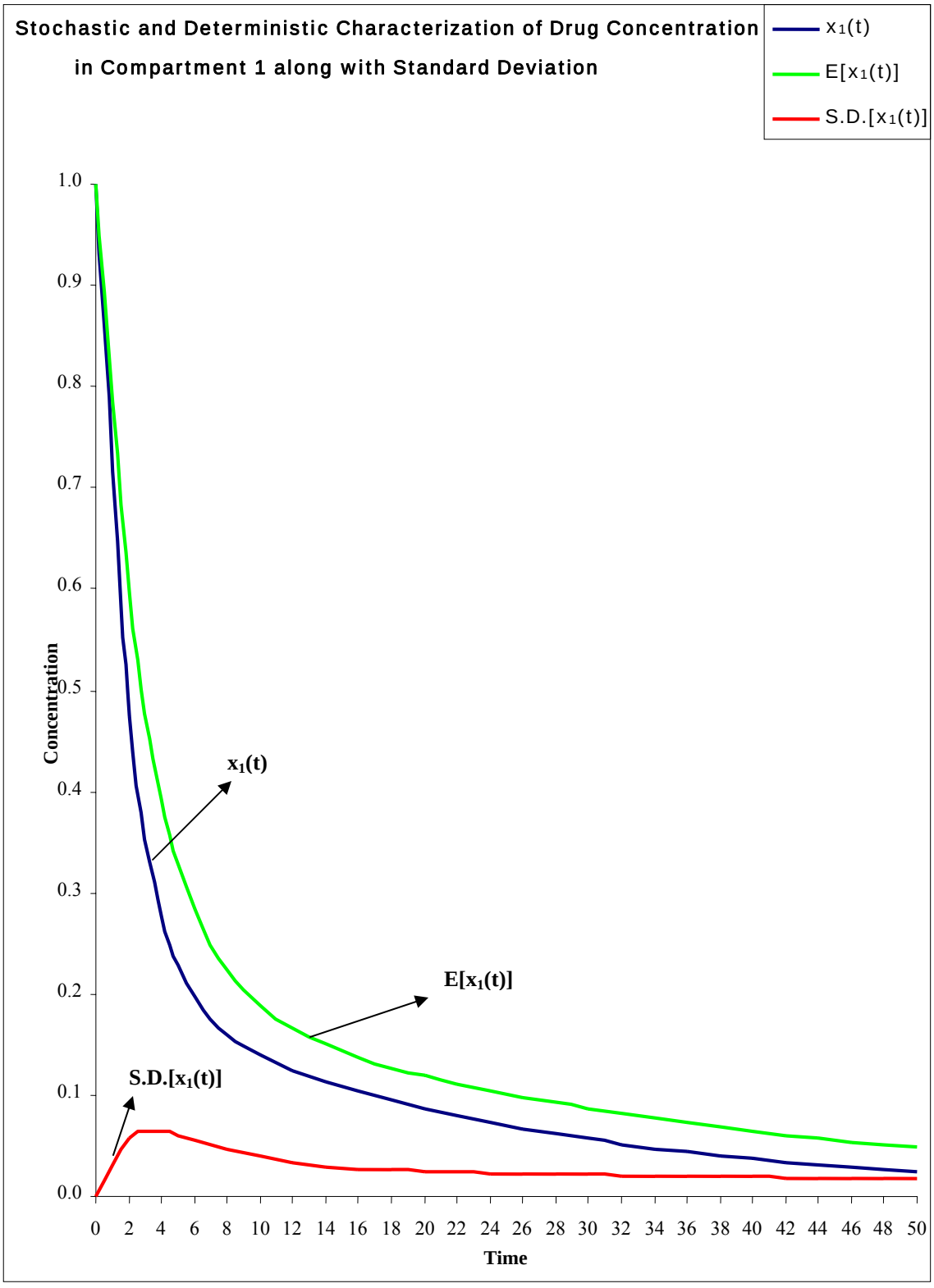


Figure 5.5A Deterministic and stochastic characterizations of $x_1(t)$ as correlation changes from 0.25 to 0.75, and standard deviation, from 5% to 20%.

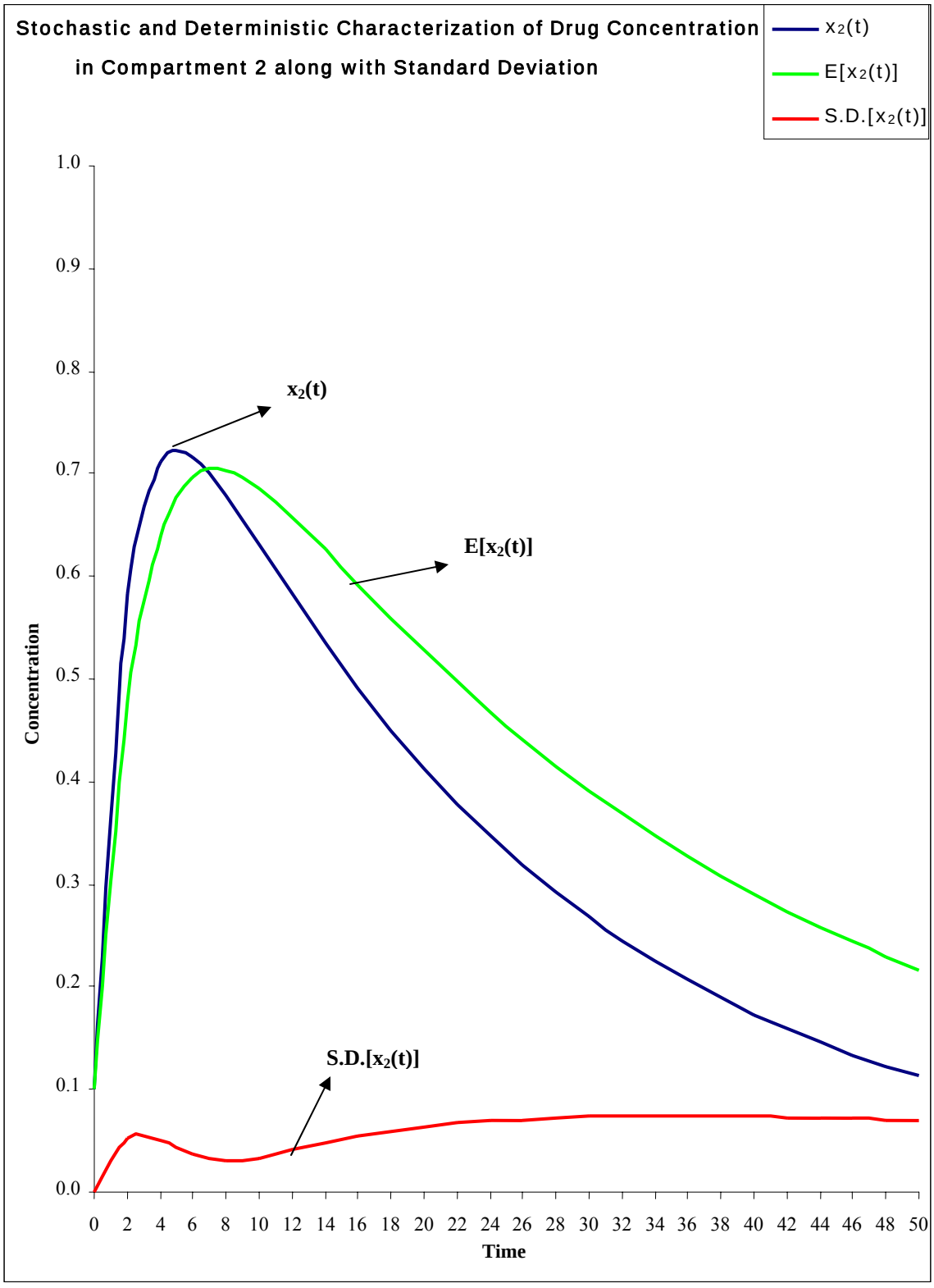


Figure 5.5B Deterministic and stochastic characterizations of $x_2(t)$ as correlation changes from 0.25 to 0.75, and standard deviation, from 5% to 20%.

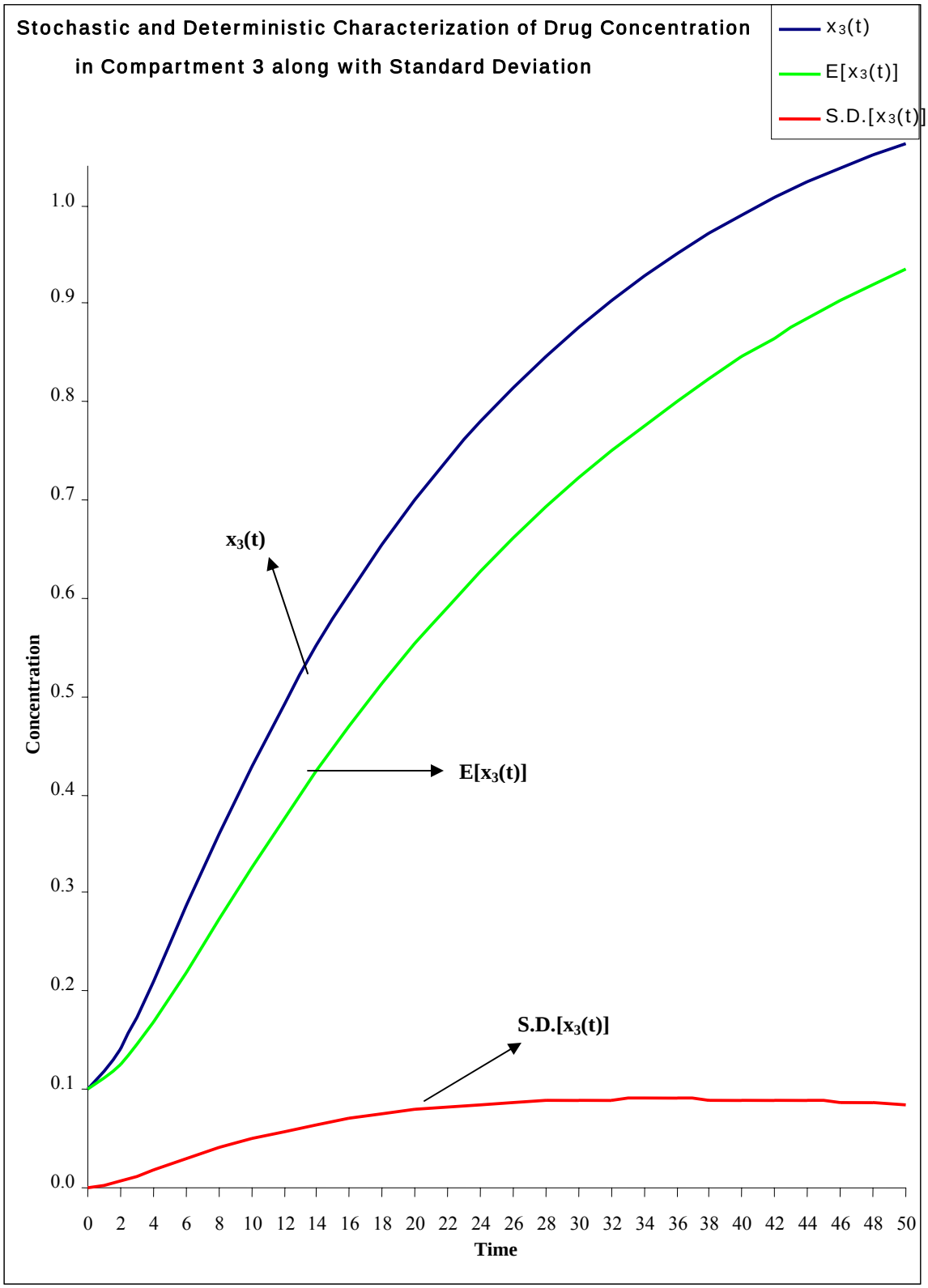


Figure 5.5C Deterministic and stochastic characterizations of $x_3(t)$ as correlation changes from 0.25 to 0.75, and standard deviation, from 5% to 20%.

5.6 Simulation of Rate Constants Using Trivariate Exponential Distribution

5.6.1 Introduction

Except for the normal distribution, multivariate extensions of univariate distributions are not obvious. The normal is simple because the first and the second moments extend in an obvious way, and they completely determine the multivariate normal distributions.

There are various ways to define a multivariate exponential distribution. Marshall and Olkin defined a multivariate exponential distribution that arises naturally in modeling observable processes, in terms of Poisson shocks. The idea is to consider a system of **d** components that experience shocks that follow independent Poisson processes. There are **d** Poisson processes that affect only one component; there are $\binom{d}{2}$ that affect two components simultaneously; and so on; and one that affects all **d** components simultaneously.

Each component is therefore affected by 2^{d-1} Poisson shock processes. The i^{th} element of the multivariate exponential random variable is the interval between two successive shocks experienced by the i^{th} component; that is, it is the minimum of 2^{d-1} independent exponential random variables. The exponential rate parameter for the i^{th} element is the sum of the rates of the Poisson shock processes that affect the i^{th} component. The magnitudes of the rates of the processes that affect both the i^{th} and j^{th} components simultaneously determine the covariance between the i^{th} and j^{th} elements of the multivariate exponential random variable. Marshall and Olkin discuss a number of

interesting properties of this distribution, including the fact that it, like the univariate exponential distribution, is memoryless.

For a **d**-variate exponential random variable, choose 2^d-1 exponential random variables with rate parameters

$$\lambda_1,\ldots,\lambda_d,\ \lambda_{12},\ldots,\lambda_{1d},\ \lambda_{23},\ldots,\lambda_{2d},\ldots,\lambda_{123},\ldots,\lambda_{12d},\ldots,\lambda_{12\cdots d}.$$

We generate 2^d-1 exponential random variates

$$t_1,\ldots,t_d,t_{12},\ldots,t_{1d},t_{23},\ldots,t_{2d},\ldots,t_{123},\ldots,t_{12d},\ldots,t_{12\cdots d}$$

independently from these distributions. Then, for the i^{th} element, we take x_i as the minimum of all of the t^s with a subscript containing i. For example,

$$x_1 = \min\{t_1,\ldots,t_d,t_{12},\ldots,t_{1d},t_{123},\ldots,t_{12d},\ldots,t_{12\cdots d}\}.$$

The Marshall-Olkin **d**-variate exponential distribution can also be expressed in terms of conditional distributions. A complication, however, is the fact that this multivariate distribution has regions of positive probability for which the Lebesgue measure is 0. (These correspond to events that affect two or more components simultaneously.) Dagpunar (1988) describes a method based on conditional distributions. The discontinuities require special handling, and so the method only works well for small values of **d**.

Based on the above, for the trivariate case, that is, for a 3-variate exponential random variable, choose $2^3-1=7$ exponential random variables with rate parameters $\lambda_1,\lambda_2,\lambda_3,\lambda_{12},\lambda_{13},\lambda_{23}$, and λ_{123}. The value of λ_1 is taken as the average of k_{12} values

measured on four patients. Similarly, the values for λ_2 and λ_3 are the average of k_{21} values and the average of k_{23} values, respectively measured on the same four patients.

The value for λ_{12} is computed as the average of the values of λ_1 and λ_2; the value of λ_{13} is computed as the average of λ_1 and λ_3; the value of λ_{123} is computed as the average of λ_1, λ_2 and λ_3.

Generate $2^3 - 1 = 7$ exponential random variates $t_1, t_2, t_3, t_{12}, t_{13}, t_{23}$, and t_{123}. For the i^{th} element, take x_i as the minimum of all the t^s with subscript containing i. That is,

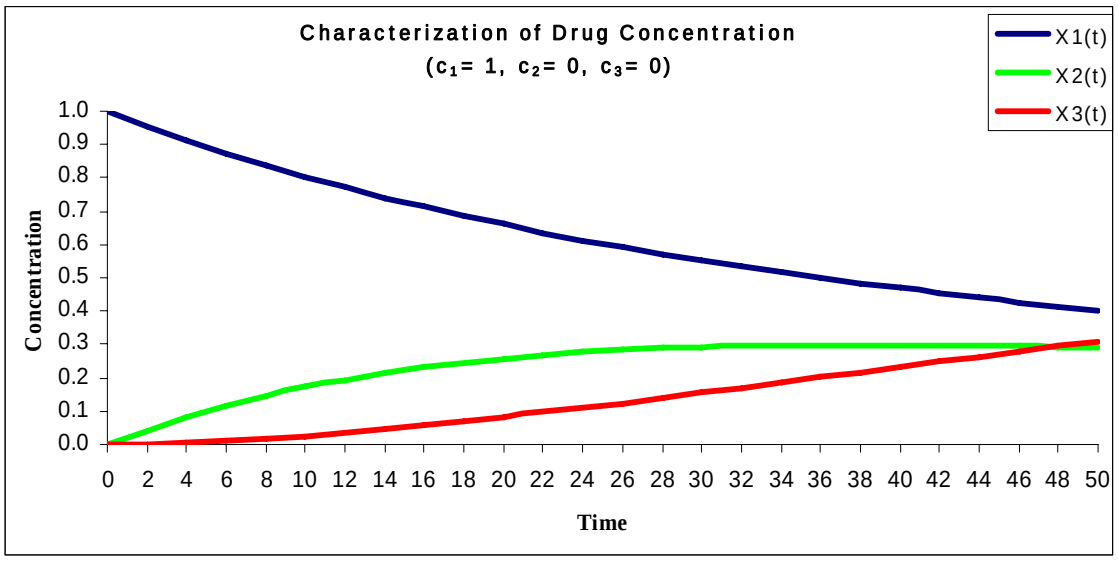
$$x_1 = \min\{t_1, t_{12}, t_{13}, t_{123}\} = k_{12}$$
$$x_2 = \min\{t_2, t_{12}, t_{23}, t_{123}\} = k_{21}$$
$$x_3 = \min\{t_3, t_{13}, t_{23}, t_{123}\} = k_{23}$$

The rate constants that follow trivariate exponential distribution are generated using the programming codes.

5.6.2 Discussion of Results

The values of the rate constants generated by the exponential probability distributions are such small values for the majority of the simulations that the overall behaviors of the drug concentration in the three compartments *do not* reflect the expected behavior of the system. The sample graphs presenting the overall behavior of the system are shown in Figure 5.6A when the initial conditions for the system of random differential equations are $c_1 = 1, c_2 = 0, c_3 = 0$. Figure 5.6B and Figure 5.6C show similar sample graphs displaying the overall behavior of the system when the initial conditions are $c_1 = 1, c_2 = 0.05, c_3 = 0.05$ and $c_1 = 1, c_2 = 0.10,$ and $c_3 = 0.10,$ respectively.

Thus, this approach that uses the rate constants generated from the exponential probability distribution in solving the system of random differential equations resulted in a poor and inconsistent characterization of the coumermycin A_1 drug concentrations as shown in the graphs 5.6A through 5.6C. This is a strong indication of the unsuitability of the use of exponential probability distribution to simulate the rate constants.



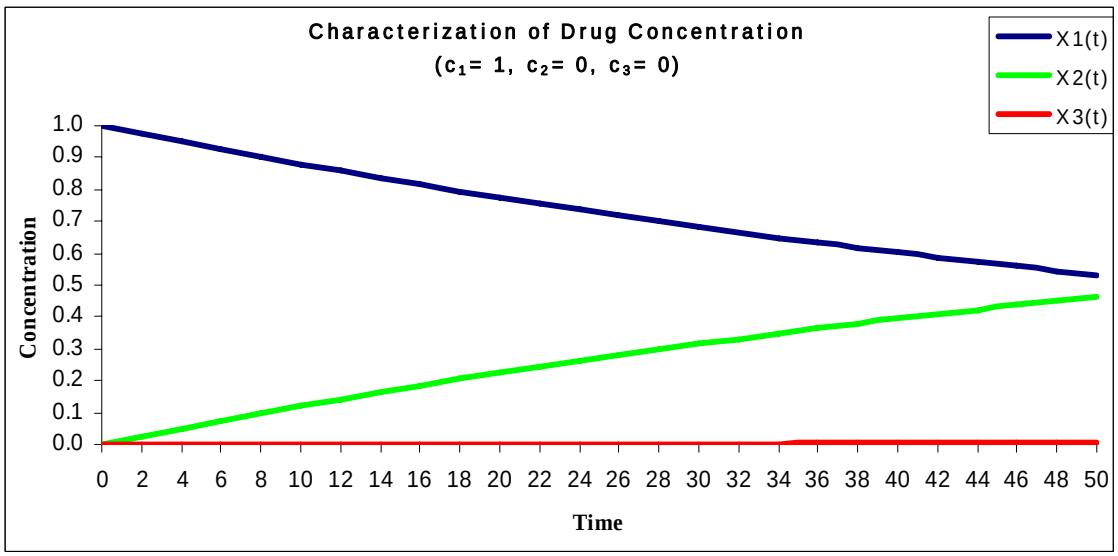
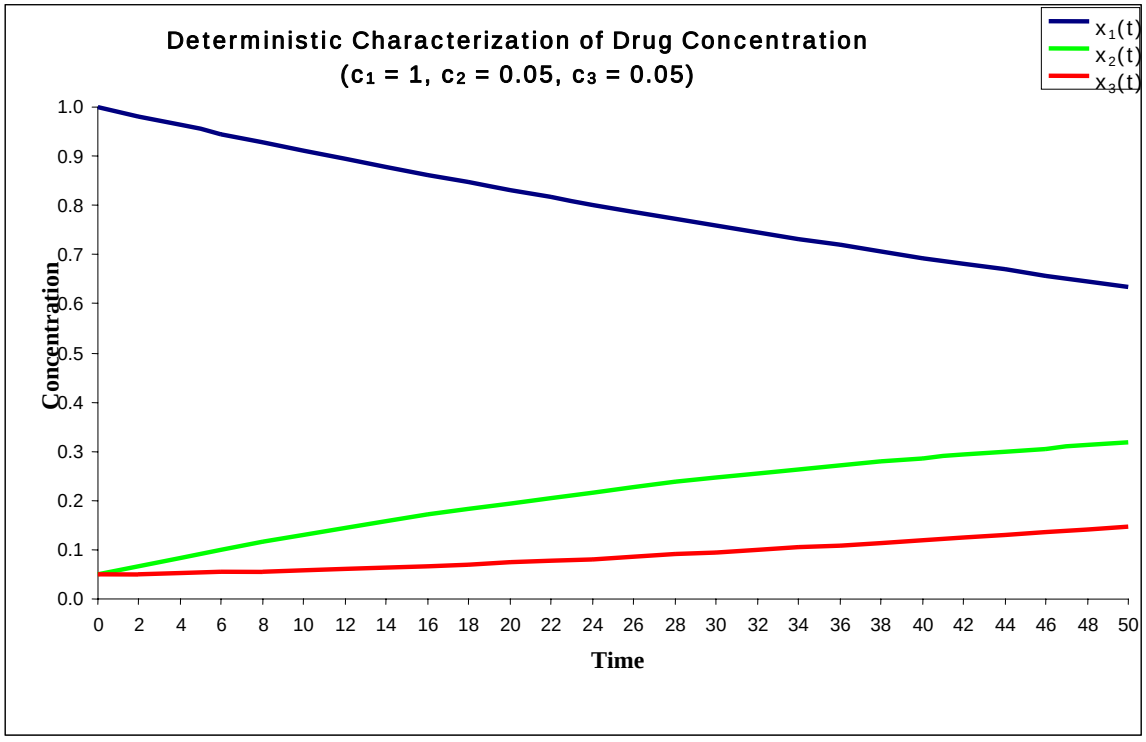


Figure 5.6A Graphs of the drug concentration behaviors in the three compartments of the two simulations when the initial conditions are $c_1 = 1$, $c_2 = 0$, and $c_3 = 0$.



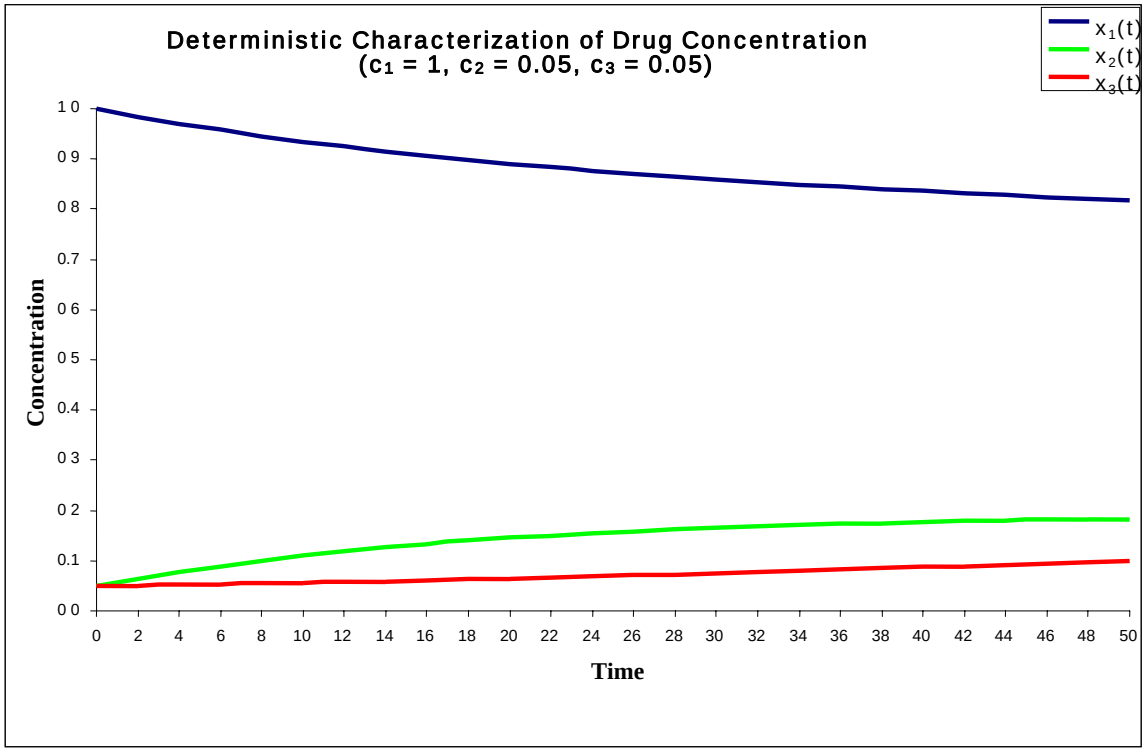
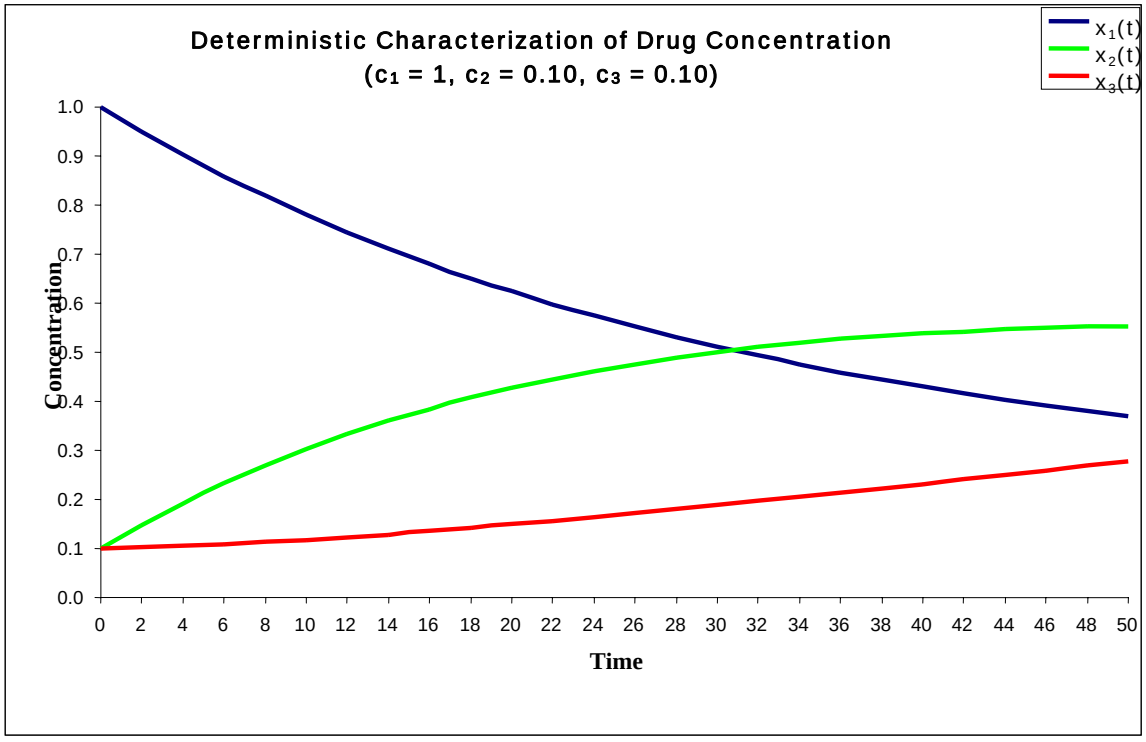


Figure 5.6B Graphs of the drug concentration behaviors in the three compartments of the two simulations when the initial conditions are $c_1 = 1$, $c_2 = 0.05$, and $c_3 = 0.05$.



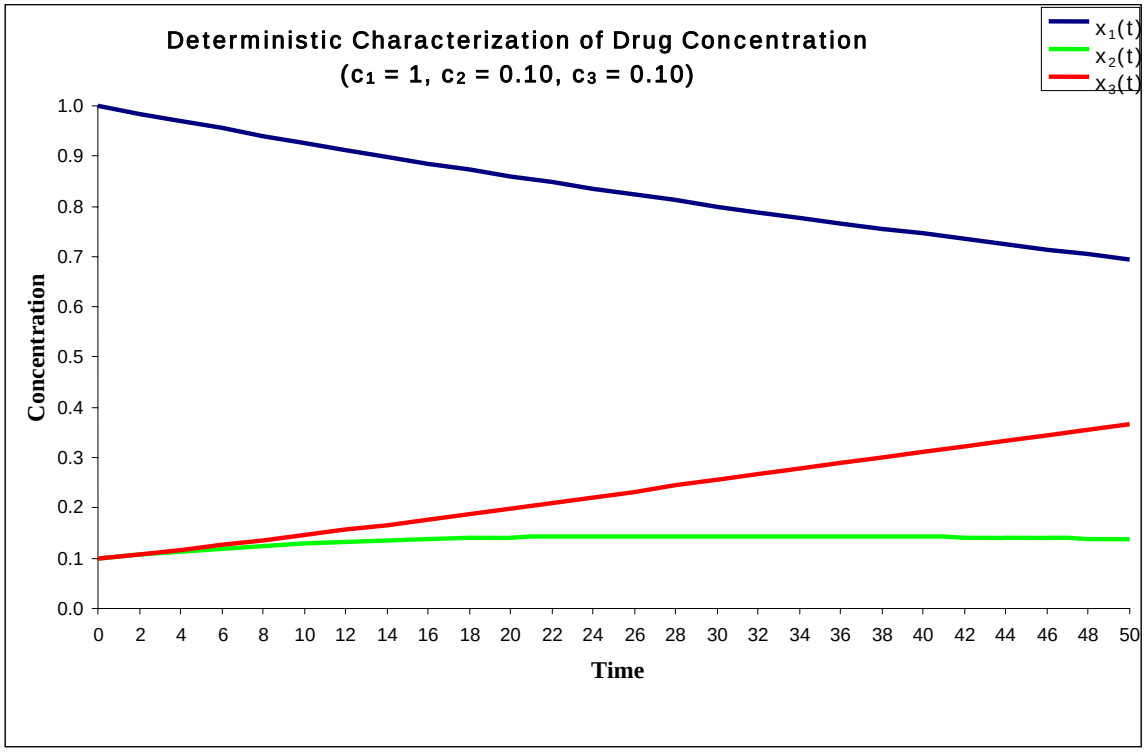


Figure 5.6C Graphs of the drug concentration behaviors in the three compartments of the two simulations when the initial conditions are $c_1 = 1$, $c_2 = 0.10$, and $c_3 = 0.10$.

5.7 Conclusion

We have presented a deterministic and stochastic analytical formulation of a three-compartment open pharmacokinetic system that describes the behavior of the coumermycin A_1 concentration in each compartment. A stochastic comparison with the deterministic characterization of the drug concentration behavior in each compartment is given. The results clearly identify the dependence effect that exists among the three compartmental structure of the antibiotic.

Using as the basic structure the actual data that was collected from administering coumermycin A_1 to four different patients, we studied numerically the deterministic and stochastic behaviors. As expected, stochastic version uniformly is different than that of its deterministic counterpart. Our numerical study included the behavior of the drug concentration as we vary the correlation structure from small, medium to high and also varying the variance from small, medium to high. Thus, we recommend that the stochastic formulation and analysis to be more appropriate for decision making for the antibiotic drug, coumermycin A_1, commonly used for the Lyme disease.

Chapter Six

Statistical Modeling of a Pharmacokinetic System Using a System of Delay Random Differential Equations

6.0 Introduction

Pharmacokinetics is a discipline that uses mathematical models to describe and predict the time-course of drug concentrations in body fluids. Our ability to apply pharmacokinetic principles has vastly improved in the past few decades as a result of advances in analytical chemistry techniques, the development of highly sensitive methods of quantitation of drug concentrations in plasma and tissues, and the availability of iterative computer methods.

Pharmacodynamics is a discipline that studies the time course and intensity of drug action or response. The ability to understand and predict individual differences in drug response is of critical importance. Advances in the clinical applications of pharmacodynamics are due to improved precision and objectivity in methods for measuring human drug response. The kinetic-dynamic modeling which uses mathematical methods to link drug concentrations directly to clinical effect has contributed to this advancement (Pharmacokinetics and Pharmacodynamics, David J. Greenblatt, Jerold S. Harmatz, Lisa L. von Moltke, and Richard I. Shader).

In any pharmacokinetics or pharmacodynamics study when a process that develops in the body over a course of time, such as drug absorption, drug dissolution, drug bioavailability, etc. is delayed, then this delay, referred to as time delay, is an

important parameter of the process. In such cases, it is necessary to build a structured model of a pharmacokinetic system with time delays. The time delay is independent of the method used to study a particular process, and signifies the delay between the time of dosing and time of appearance of concentration in the sampling compartment. The delay in drug response may complicate drug monitoring. For example, digoxin's effect on the heart is delayed because the drug requires time to be distributed to the active site, and hence, digoxin concentrations should not be measured within six hours of a dose, even after intra-venous administration. Before that time, the plasma concentration does not reflect the concentration at the active site [93].

In the previous chapter, the focus was to investigate the drug concentration behavior in a three-compartment open pharmacokinetic model which describes the disposition of an antibiotic drug coumermycin A_1. We studied a system of non-delay random differential equations representing this model. The three rate constants that were used in the system of differential equations were simulated using the trivariate truncated normal probability distribution. The initial values of the rate constants that were used in the simulation were calculated from the pharmacokinetic profile of coumermycin A_1 determined in four human subjects based on the serum level data obtained from the report of a clinical study. The extensive numerical solutions for the system of non-delay random differential equations under different combinations of the covariance structure and the initial conditions were developed.

Numerical comparisons of the deterministic characterizations of the drug concentration as a function of time of the individual compartments to study the effect of various combinations of the covariance structure and the initial conditions on these

characterizations were presented. A similar comparison between the deterministic and the stochastic characterizations was also presented.

6.1 Focus of Chapter Six

The focus of the present chapter is to incorporate the time delay into the system of random differential equations that was considered in the previous chapter, and develop numerical solutions for the system which describes the disposition of coumermycin A_1. Schematically, the open three-compartment pharmacokinetic model with the time delay components that we are studying is illustrated in Figure 6.2.1.

More precisely, the numerical solutions for the system of delay random differential equations will be obtained for different sets of constant time delay values of τ_1, τ_2 and τ_3 but under the same combinations of the covariance structure and the initial conditions that were described in the previous chapter. The rate constants used in the system are same as those simulated using the trivariate truncated normal distribution as described in the previous chapter.

From the previous chapter we observed a time delay of 3.5 hours for the drug to reach compartment two from compartment one, a time delay of 6 hours for the drug to reach compartment one from compartment two, and a time delay of 3 hours for the drug to reach compartment three from compartment two.

In the present chapter we use these time delays $(\tau_1 = 3.5h, \tau_2 = 6h, \tau_3 = 3h)$ to obtain numerical solutions for the system of delay differential equations and study its impact on the rate of decay, absorption and biotransformation of the drug concentration. In addition we consider other sets of constant time delay values to study the behavior of the drug concentration under different time delays.

For all 27 configurations shown in Figure 6.1.1 the numerical solutions for the system of delay random differential equations are obtained based on 1000 simulations and using each of the following four sets (referred to as DT1, DT2, DT3, and DT4) of constant time delay values in hours:

DT1. vector(τ_1, τ_2, τ_3) = (0.5, 1, 1.5)

DT2. vector(τ_1, τ_2, τ_3) = (1, 2, 3)

DT3. vector(τ_1, τ_2, τ_3) = (3.5, 6, 3)

DT4. vector(τ_1, τ_2, τ_3) = (4, 6, 8)

At first we present the discussion of the effect of each of the four sets of constant time delay values on the overall behavior of the drug concentration in all three compartments.

Secondly, the comparison of the deterministic behavior of the drug concentration with and without time delay for each compartment will be discussed.

Finally, the effects of different sets of constant time delay values on the deterministic behavior of the drug concentration and the mean behavior of the random solutions as a function of time will be discussed.

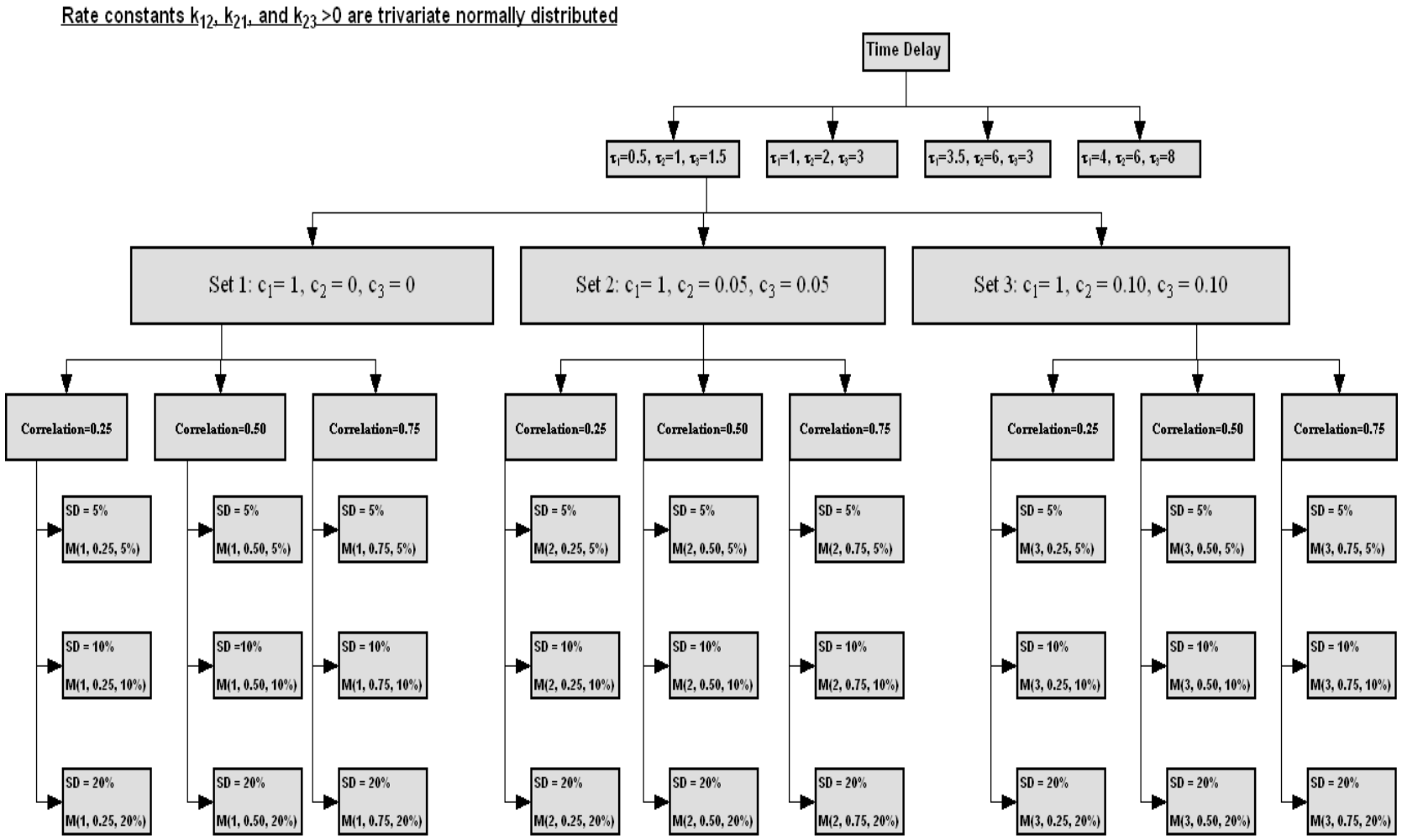


Figure 6.1.1 Flow-chart describing the configurations for all 27 models for each of the four sets of constant time delay values

6.2 Pharmacokinetic Model

The pharmacokinetic model with the delay components is as follows:

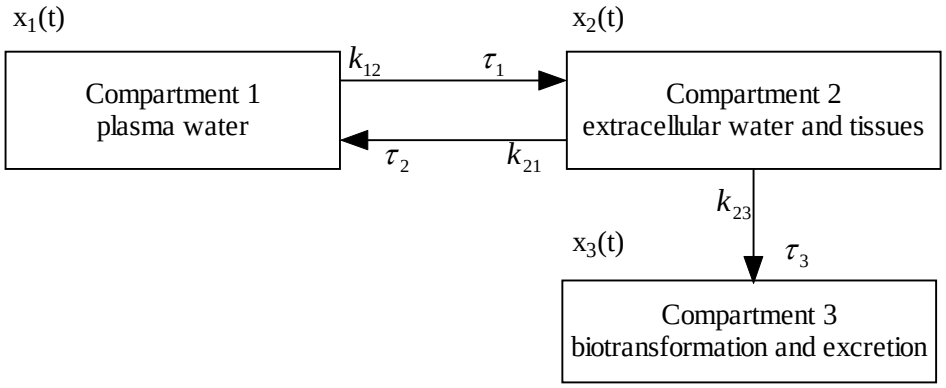


Figure 6.2.1 Model for the disposition of coumermycin A_1

The above model is described by the following system of random differential equations:

$$\dot{x}_1(t) = -k_{12}x_1(t-\tau_1) + k_{21}x_2(t-\tau_2)$$
$$\dot{x}_2(t) = k_{12}x_1(t-\tau_1) - k_{21}x_2(t-\tau_2) - k_{23}x_2(t-\tau_3)$$
$$\dot{x}_3(t) = k_{23}x_2(t-\tau_3)$$

with the initial conditions $x_1(0) = c_1$, $x_2(0) = c_2$, $x_3(0) = c_3$; $x_1(t)$, $x_2(t)$, and $x_3(t)$ are the amounts of the drug concentration in compartments 1, 2, and 3, respectively, at time $t \geq 0$; τ_1, τ_2, and τ_3 are the time delays before the drug concentration reaches compartment 2 from compartment 1, compartment 1 from compartment 2, and compartment 3 from compartment 2, respectively; k_{12}, k_{21}, and $k_{23} > 0$ are the rate constants of the system.

In the matrix form, the above system of differential equations can be written as follows:

$$\dot{x}(t)=\begin{bmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\\ \dot{x}_3(t)\end{bmatrix}=\begin{bmatrix}-k_{12} & k_{21} & 0\\ k_{12} & -k_{21} & -k_{23}\\ 0 & 0 & k_{23}\end{bmatrix}\begin{bmatrix}x_1(t-\tau_1)\\ x_2(t-\tau_2)\\ x_2(t-\tau_3)\end{bmatrix}, \text{ with initial conditions } C=\begin{bmatrix}c_1\\ c_2\\ c_3\end{bmatrix}$$

In general, solutions of this type of equations are difficult to find in closed forms. Using an integral approach, integrating each side of the equations, we have,

$$x_1(t)=\int -k_{12}x_1(t-\tau_1)dt+\int k_{21}x_2(t-\tau_2)dt+m_1$$

$$x_2(t)=\int k_{12}x_1(t-\tau_1)dt-\int k_{21}x_2(t-\tau_2)dt-\int k_{23}x_2(t-\tau_3)dt+m_2$$

$$x_3(t)=\int k_{23}x_2(t-\tau_3)dt+m_3$$

where $m_1, m_{2,}$ and m_3 are constants of integration. By knowing the values of each of the functions prior to $t-\tau_i=0$, this system can be solved using the method of steps.

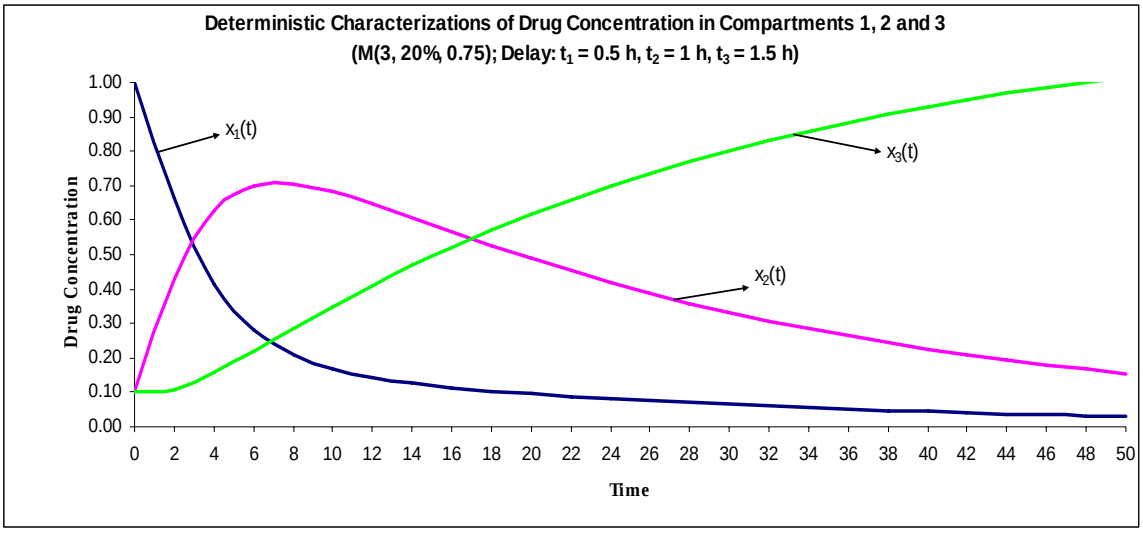
However, for the present study the numerical solutions to the above system of delay differential equations were obtained through a computer program written in *Mathematica 5.1* and *SAS 9.1* computer languages. Both programs produced identical results.

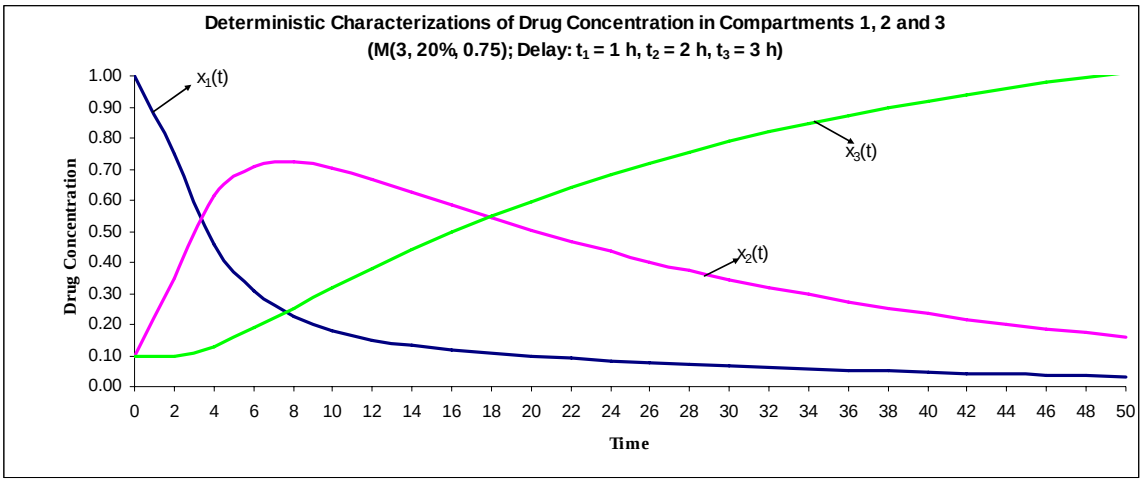
In section 6.3, we present the discussion of the numerical results along with the representative graphs, followed by conclusion in section 6.4.

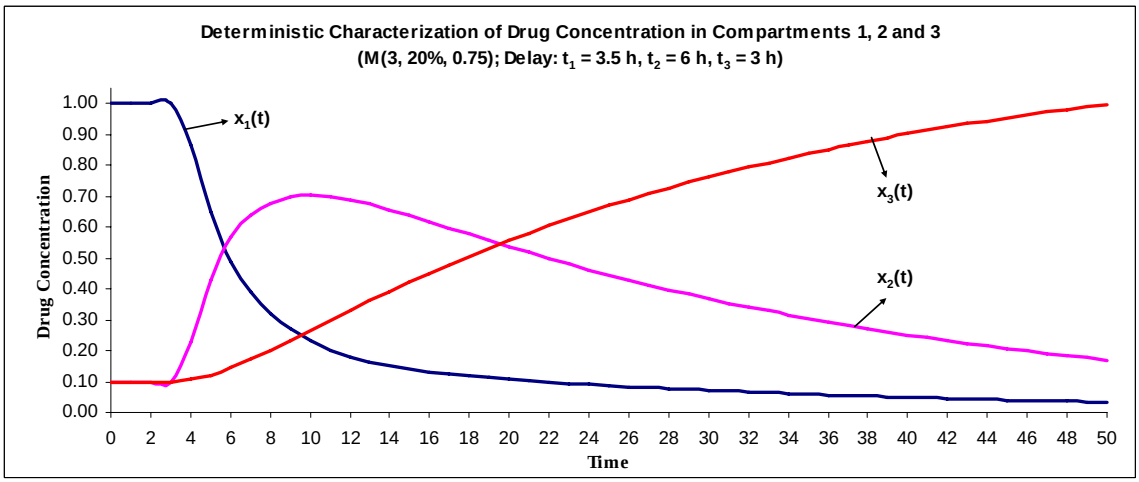
6.3 Discussion of the Results

6.3.1 Overall Behavior of the Drug Concentration in all Three Compartments

The graphs presented in Figure 6.3.1 show the deterministic behavior of coumermycin A_1 concentration in all three compartments for the first 50 hours under four different sets of constant time delay values that are described in section 6.1. The behavior of the drug concentration in all three compartments is as expected. The drug concentration in the first compartment (*plasma water*) decays rapidly as indicated by the graph of $x_1(t)$ after the initial delay, while the drug concentration in the second compartment (*extracellular water and tissues*) increases during the *first eight to ten hours* as indicated by the graph of $x_2(t)$ before it starts decaying. The peak values are slightly higher than those for the non-delay models of the previous chapter. The concentration behavior in the third compartment (*biotransformation and excretion*) increases exponentially as indicated by the graph of $x_3(t)$. Across all models this overall behavior of the drug concentration in all three compartments is similar.







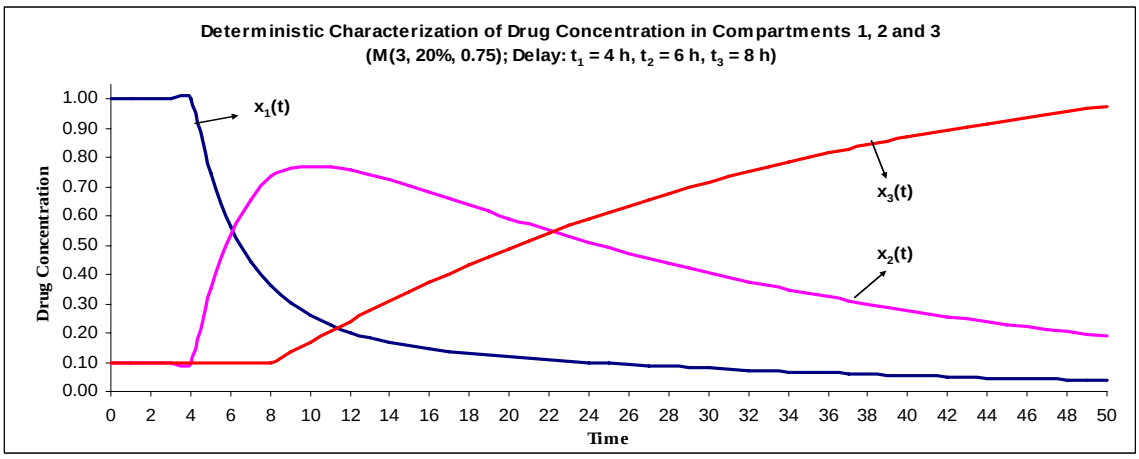


Figure 6.3.1 Deterministic characterizations of drug concentration in all 3 compartments

6.3.2 Comparison of Deterministic Behavior Between Delay and Non-Delay Models

Deterministic Behavior of Drug Concentration, $x_1(t)$

The difference in the rate of decay in the drug concentration in compartment one due to a small delay of 0.5 hour is almost negligible during the first 12 hours, while this difference increases as the time delay increases from 0.5 to 1 hour. The deterministic behaviors of the drug concentration in compartment one for both delay and non-delay cases are almost identical after about the first 12 hours when the time delay in compartment one is small (τ_1 = 0.5 or 1 h) compared to the length of time (50 hours) the drug concentrations are being monitored. These can be observed from the graphs shown in the top two panels of Figure 6.3.2A.

As the time delay increases from 1 hour to 3.5 or 4 hours, not only the difference in the rate of decay of drug concentration in compartment one is highly significant compared to the non-delay profile but also takes longer period of time, beyond 24 hours, before the difference between the two profiles become negligibly small. For the case of $\tau_1 = 3.5h$, the graph is shown in the bottom panel of Figure 6.3.2A.

Deterministic Behavior of Drug Concentration, $x_2(t)$

The rate of change in the drug concentration in compartment two is subject to two time delay constants. It takes a time delay of τ_1 for the drug to reach compartment two from compartment one, and a time delay of τ_2 for the drug to reach compartment one from compartment two. The difference in the rate of absorption of the drug concentration in compartment two due to a small delay of 1 hour is almost negligible during the first 8 hours, while this difference increases as the time delay increases from 1 to 2 hour. The peak value in both cases of time delay is almost the same. The deterministic behaviors of

the drug concentration in compartment two for both delay and non-delay cases are almost identical after about the first 8 hours when the time delay in compartment one is small (τ_1=1) compared to the length of time (50 hours) the drug concentrations are being monitored. This can be observed from the graph shown in the top panel of Figure 6.3.2B.

As the time delay increases from 1 hour to 2 or 6 hours, not only the difference in the rate of absorption of drug concentration in compartment two is significant compared to the non-delay profile but also attains a higher peak value, especially in case of 6-hour delay. All through the observation period of 50 hours the amount of absorption is slightly higher than those observed in its corresponding non-delay case. These are displayed in the graphs shown in the middle and bottom panels of Figure 6.3.2B.

Deterministic Behavior of Drug Concentration, $x_3(t)$

The rate of change in the drug concentration in compartment three is subject to one time delay constant, that is, the time τ_3 taken by the drug to reach compartment three from compartment two. Although the amount of biotransformation and excretion in case of a small time delay of 1.5 hours is almost identical to that of the non-delay case, there is a uniformly significant difference in the deterministic behaviors of for both delay and non-delay as the time delay increases from 1.5 to 3 hours or 3 to 8 hours. These behaviors are shown in Figures 6.3.2C.

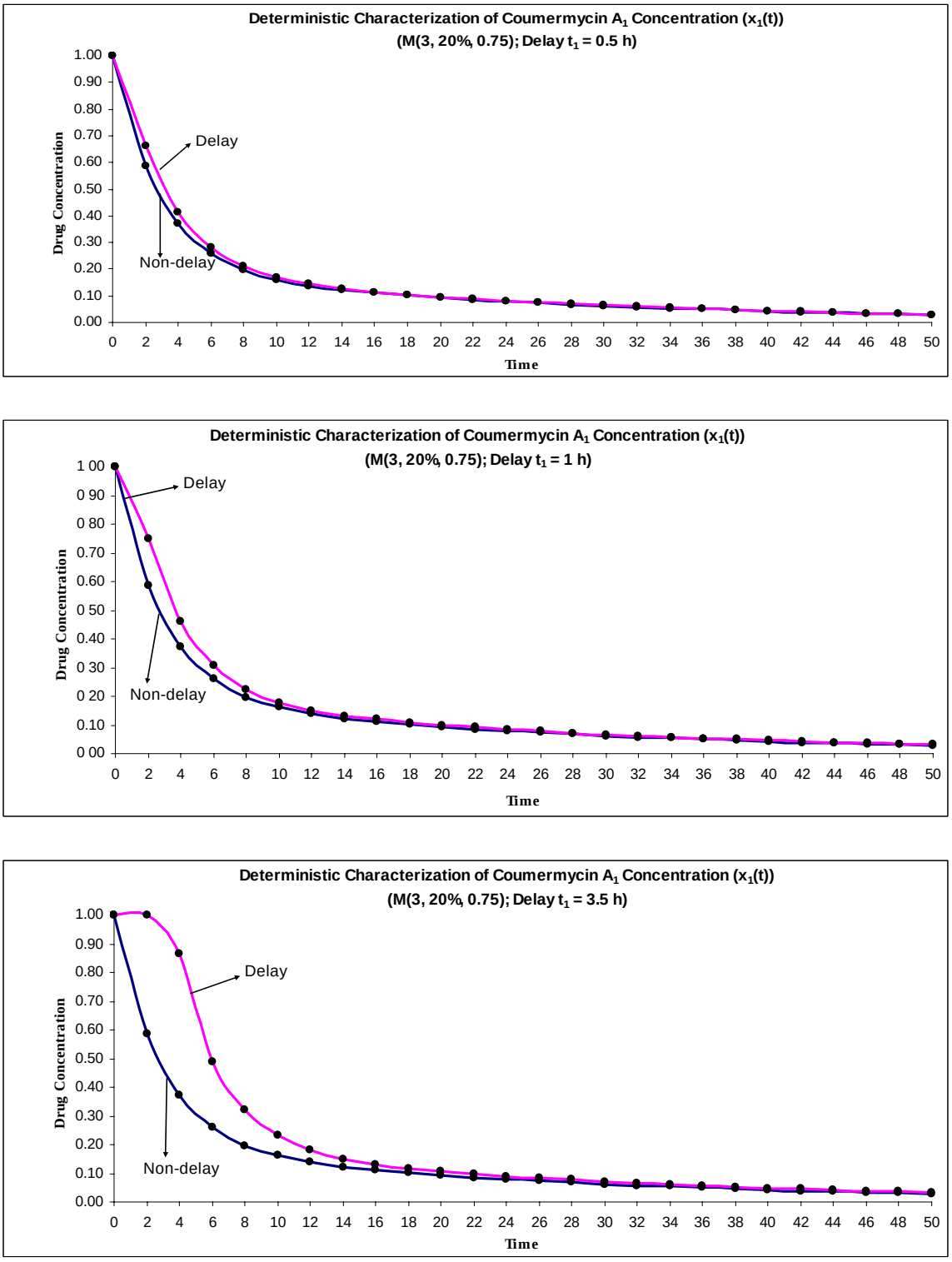
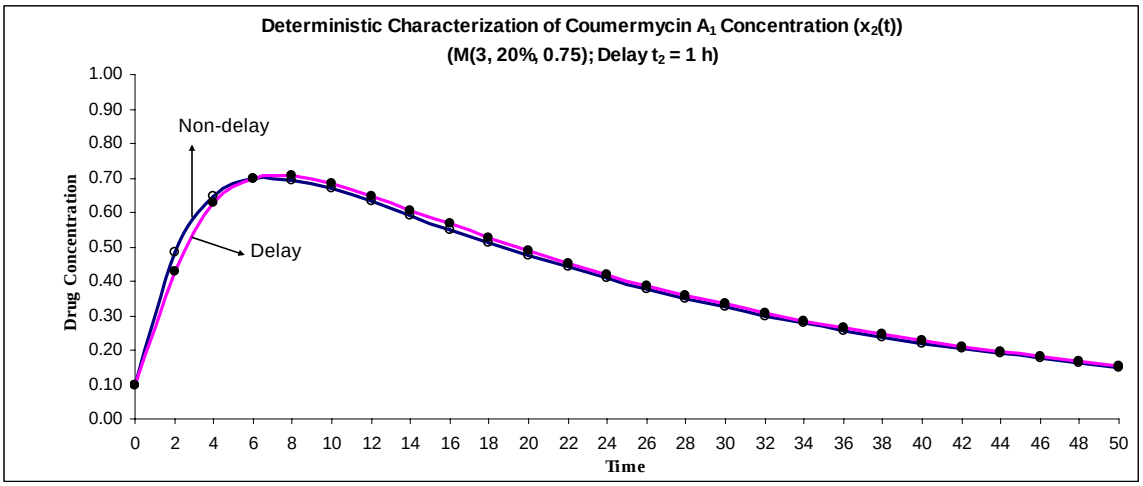
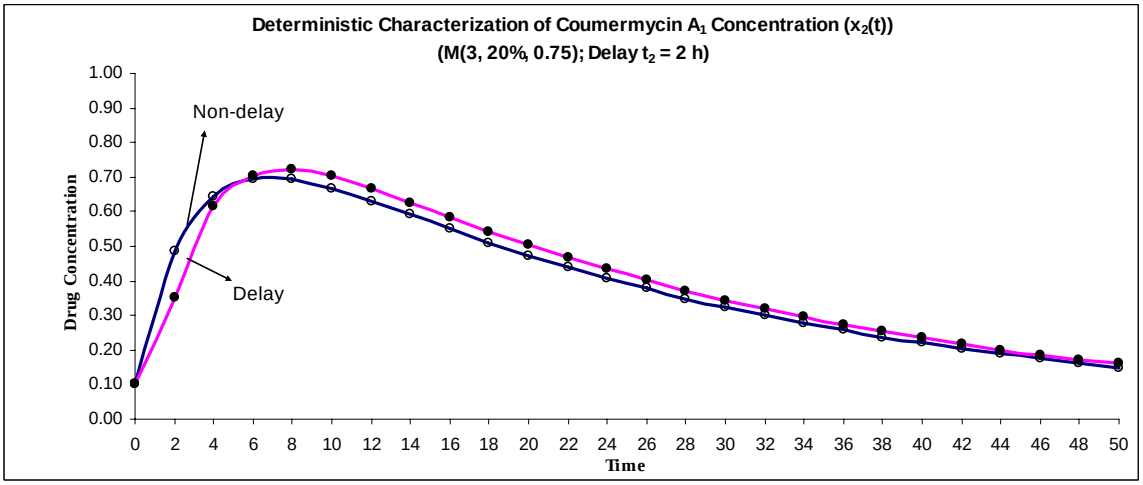


Figure 6.3.2A Deterministic characterization of coumermycin A_1 concentration ($x_1(t)$) for time delay values 0.5h, 1h, and 3.5h





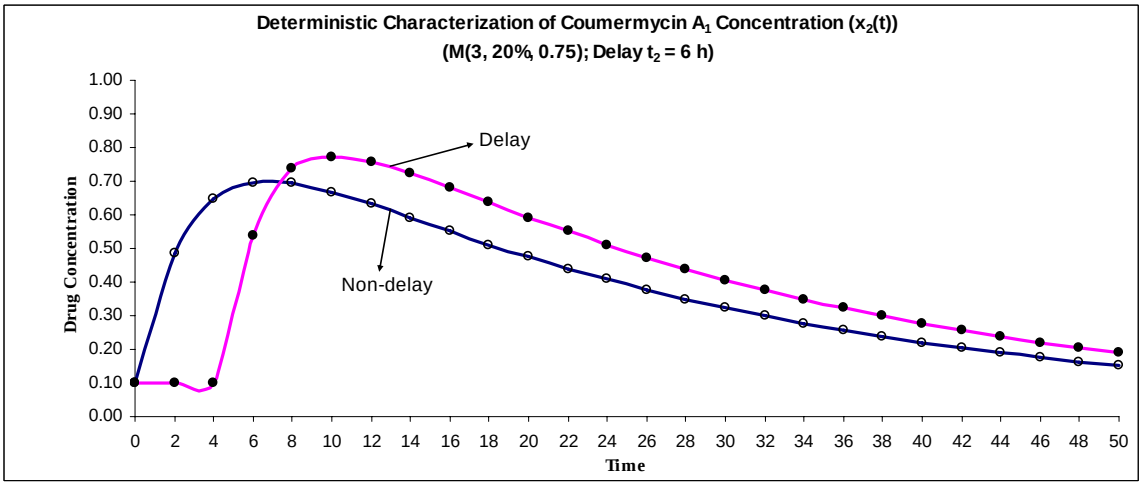


Figure 6.3.2B Deterministic characterization of coumermycin A_1 concentration ($x_2(t)$) for time delay values 1h, 2h, and 6h

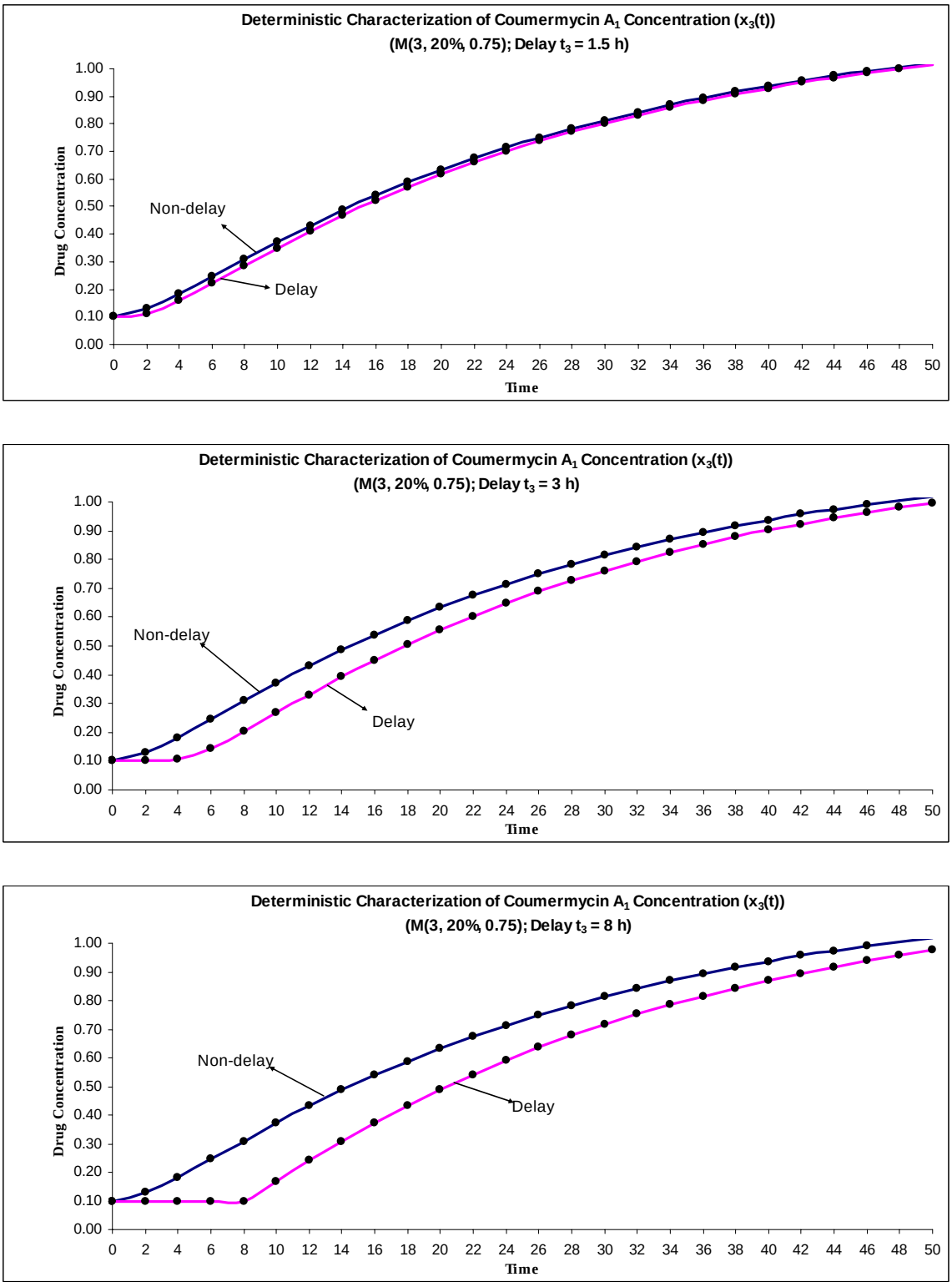


Figure 6.3.2C Deterministic characterization of coumermycin A_1 concentration ($x_3(t)$) for time delay values 1.5h, 3h, and 8h

6.3.3 Comparison of Deterministic and Stochastic Behaviors for the Delay Models

Deterministic and Stochastic Characterizations of x_1(t)

The deterministic and stochastic characterizations are almost identical during the first 6 hours for the first two sets of time delay values whereas for the last two sets of delay constants they are almost identical during the first 7 hours. But after these time points, the stochastic characterization is uniformly higher than the deterministic for all four sets of time delay values. Despite the varying time delay values the two characterizations are almost identical during the first 6 to 7 hours.

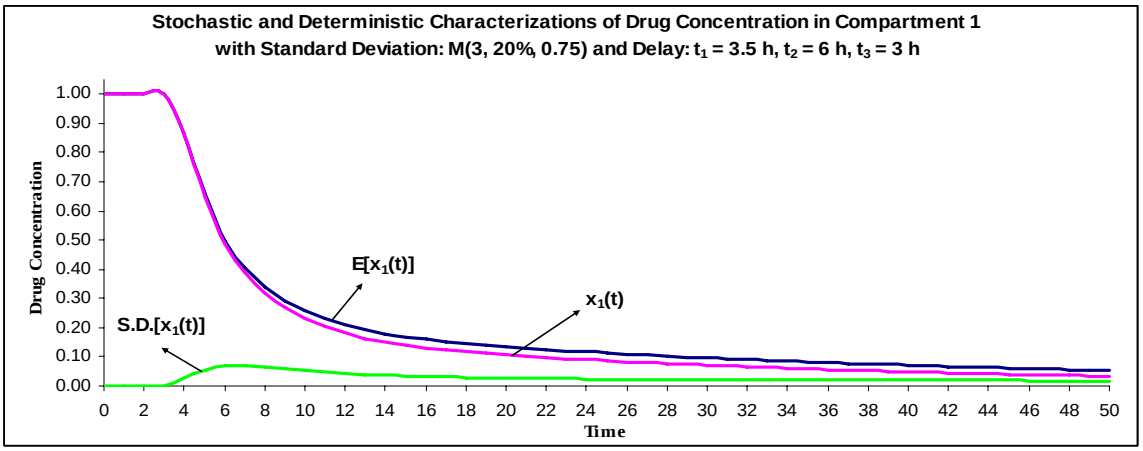
Deterministic and Stochastic Characterizations of x_2(t)

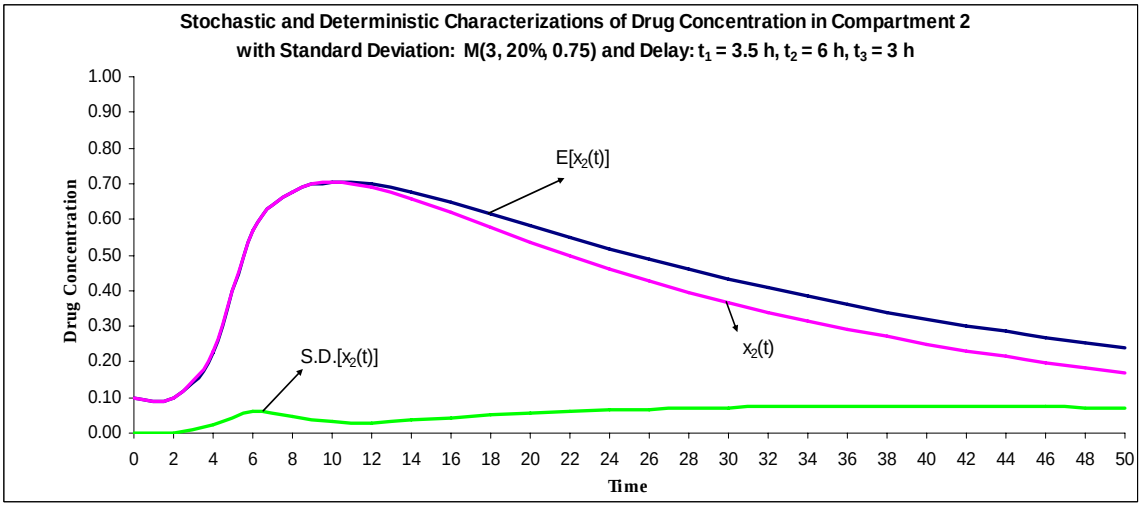
The deterministic and stochastic characterizations are almost identical during the first 10 hours for the first two sets of time delay values whereas for the last two sets of delay constants they are almost identical during the first 12 to 14 hours. But after these time points, the stochastic characterization is significantly higher than the deterministic for all four sets of time delay values.

Deterministic and Stochastic Characterizations of x_3(t)

The deterministic and stochastic characterizations are identical during the initial time delay values for all four sets of time delay values. But after these time points, the stochastic characterization is significantly higher than the deterministic in all four cases, and the difference increases with time.

Three panels that are displayed in Figure 6.3.3A show the representative graphs of the deterministic and stochastic characterizations of the drug concentration for the case of time delay constants, $\tau_1 = 3.5h$, $\tau_2 = 6h$, and $\tau_3 = 3h$. Figures 6.3.3B – 6.3.3D show both characterizations of all three compartments and for all four sets of time delay values.





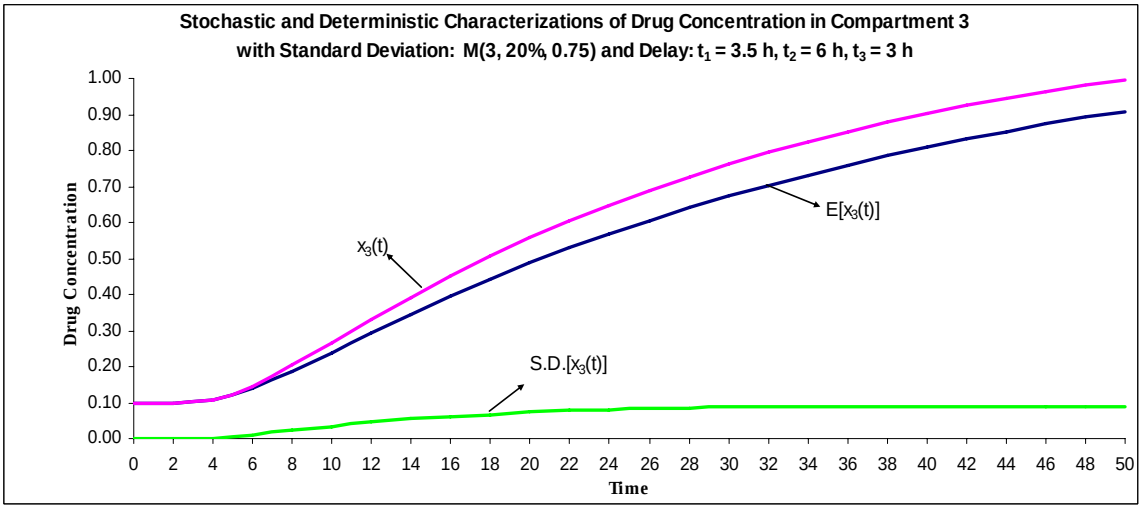


Figure 6.3.3A Stochastic and deterministic characterizations of drug concentration in all three compartments with standard deviation for time delay values: t_1=3.5h, t_2=6h, t_3=3h

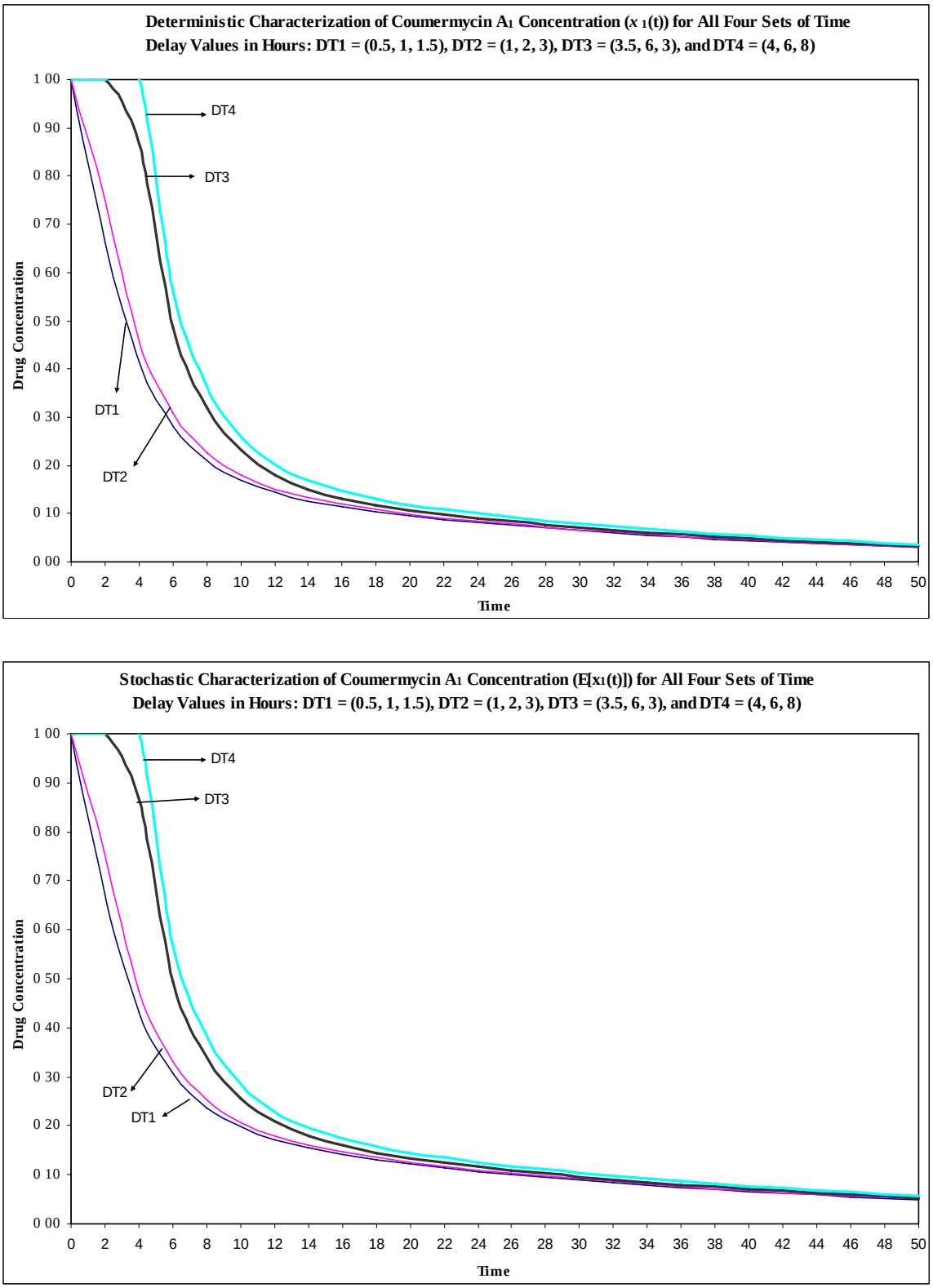
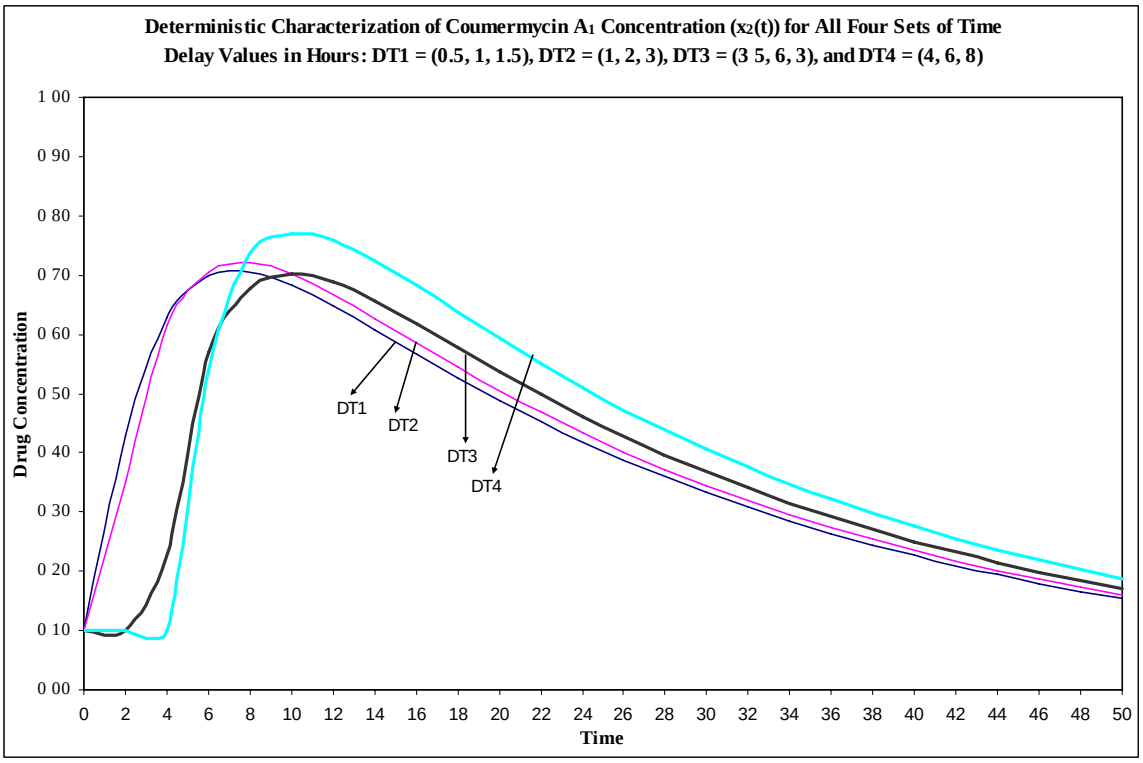


Figure 6.3.3B Deterministic ($x_1(t)$) and stochastic ($E(x_1(t))$ characterizations of coumermycin A_1 drug concentration for all four sets of time delay values in hours



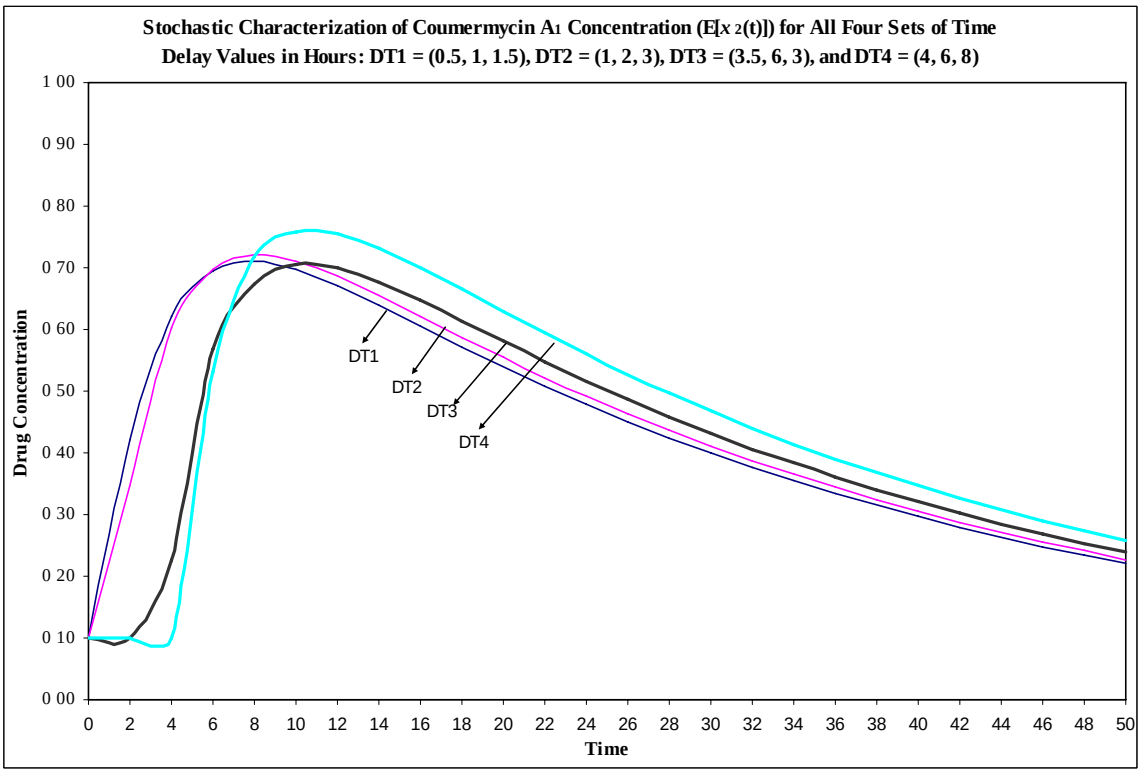
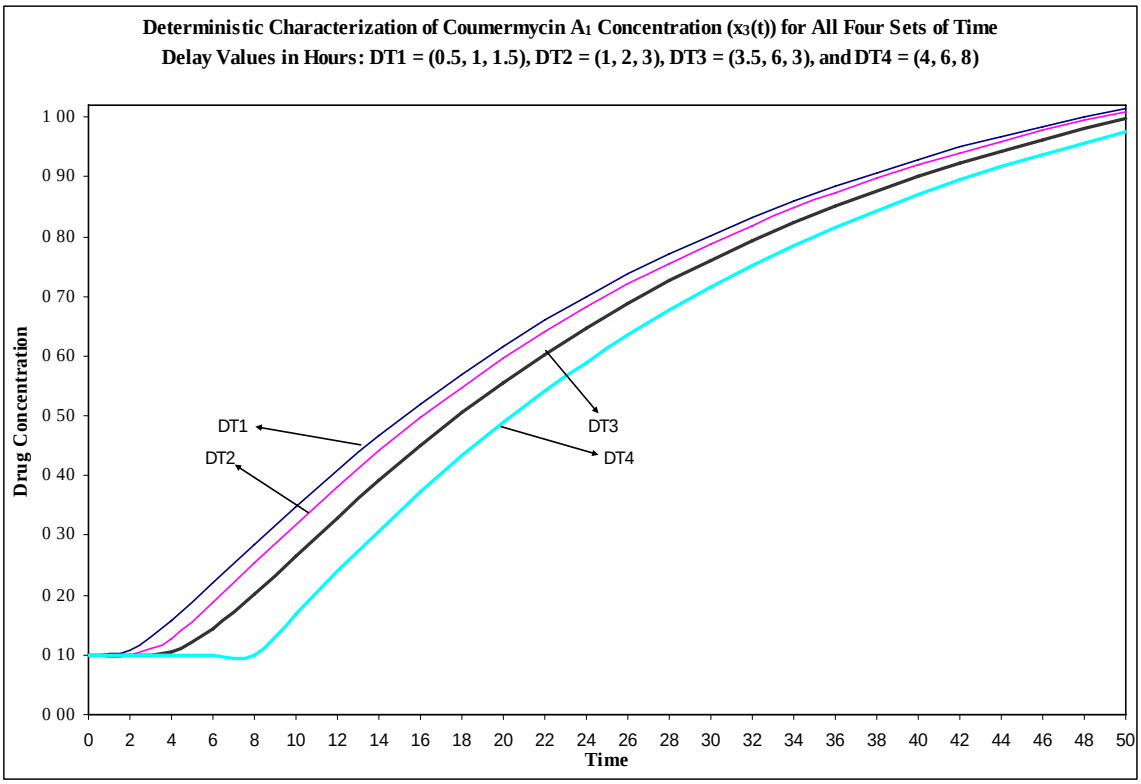


Figure 6.3.3C Deterministic ($x_2(t)$) and stochastic ($E(x_2(t))$ characterizations of coumermycin A_1 drug concentration for all four sets of time delay values in hours



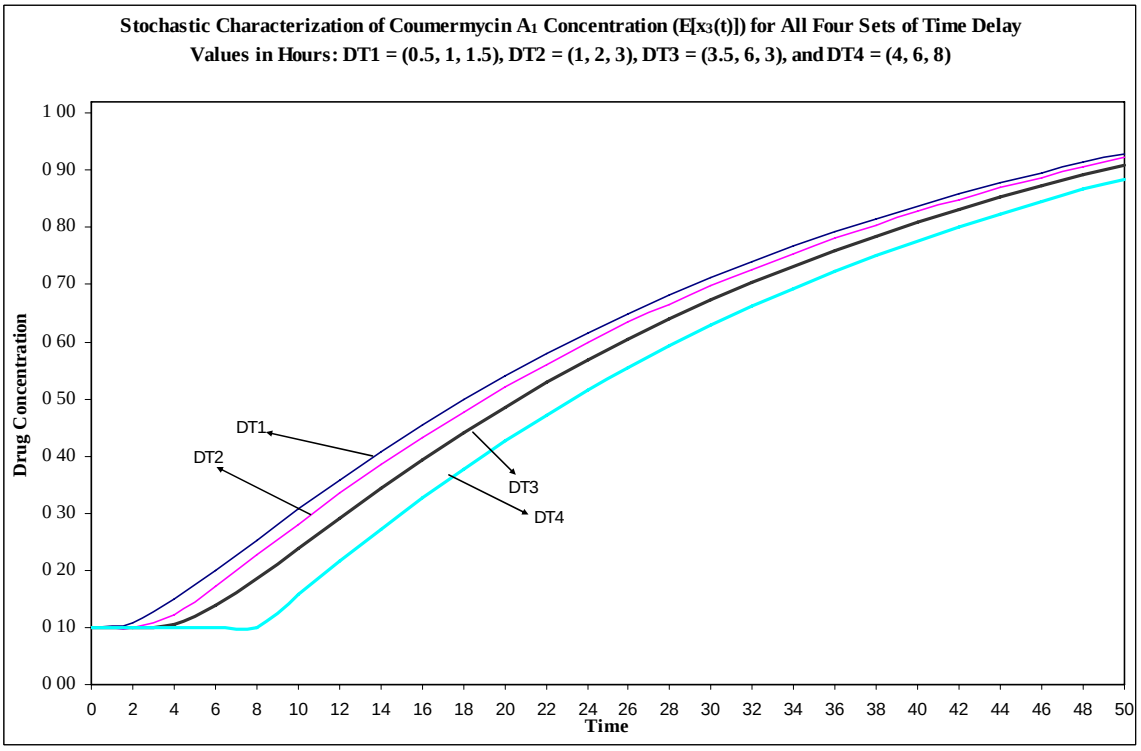


Figure 6.3.3D Deterministic ($x_3(t)$) and stochastic ($E(x_3(t))$ characterizations of coumermycin A_1 drug concentration for all four sets of time delay values in hours

6.4 Conclusion

In this study we have presented a deterministic and stochastic formulation of a three-compartment open pharmacokinetic system with pre-determined constant time delays incorporated in the system that describes the behavior of the coumermycin A_1 in each compartment. We have presented the effects of constant time delays on the drug concentration behavior in each compartment. We studied numerically the deterministic behavior of delay and non-delay systems, and performed the comparison of the drug concentration behaviors. The results clearly indicate the dependence effect that exists among the compartments when the time delays are larger.

We studied numerically the deterministic and stochastic behaviors of the delay system. As expected the stochastic version uniformly is different than that of its deterministic counterpart, especially at larger values of the time delays. Our numerical study included the behavior of the drug concentration as we vary time delay values from small, medium to high.

Understanding the time delay that exists between the time of administration of the drug and the first time point at which drug concentration is observed in the system is crucial in drug monitoring process.

Chapter Seven

Future Research Studies

7.0 Possible Extensions of the Present Research Studies

In this chapter we shall pose some possible extensions of the present research.

In chapter two, we have used the Generalized Extreme Value (GEV) distribution for statistical modeling of the annual monthly maximum rainfall from 44 locations in the State of Florida. Using the model parameters we obtain estimates and confidence intervals for return levels of various return periods. We further classify all forty four locations into clusters based on the similarity profiles.

It would be of interest to develop a statistical model that identifies the contributory independent variables that affect the rainfall characterizations in the State of Florida. For example, to determine if such independent variables as humidity, air temperature, altitude, longitude, wind velocity, air pressure, among others are specifically statistically significantly contributing to the amount of rainfall in Florida.

Having established such models which may be non-linear in nature with a high degree of accuracy, namely, high (greater than 0.85) r^2 and adjusted-r^2, we want to look at similar models for regional segments in the United States and furthermore, how such models relate to non-tropical regions such as desert regions in the United States.

In chapters three and four, we have utilized the family of extreme value theory to statistically model the maximum drug concentration $(C_{\max})$, a critical pharmacokinetic

parameter in the drug development process in a clinical trial, for predictive purposes. This is the first application of extreme value theory to model $C_{\max}$.

Extreme maximum drug concentration levels of any drug in a human body can cause toxicity, inefficaciousness, or even serious health problems. Hence, it is of paramount importance that a model, incorporating such covariates as age, gender, presence of other diseases and use of other concurrent medications, among others, be developed. Using such models, one can predict the maximum drug concentration levels in a patient whose profile is known. This can revolutionize the concept of individualized, optimal patient care.

For the problem of statistical modeling of a three-compartment pharmacokinetic system using a system of random differential equations, from chapters five and six, we propose to investigate the identification of interactions among the compartments. Furthermore, we want to determine if a markovian approach to modeling the three-compartment model is meaningful to get a better understanding of the drug.

In addition, we like to determine if connecting the third and first compartment will provide more information about the drug disposition. This will require the modification of the stochastic system process to a system of four differential equations. Also, the question of interest is to investigate the stochastic system from a non-linear perspective and determine and justify in deviating from a linear to non-linear statistical characterization of the behavior of the drug, coumermycin A_1.

List of References

1 Aarssen, K and De Haan, L. (1994): On the maximal life span of humans, Mathematical Population Studies, 4, 259-281.

2 Abdallah, H.Y. and Ludden, T.M. (1995): A spreadsheet program for simulation of bioequivalence and bioavailability studies. Comput Biol Med, 25: 349-354,

3 Achcar, J.A., Bolfarine, H. and Pericchi, L.R. (1987). Transformtaion of survival data to an extreme value distribution, Statistician 36, 229-234.

4 Aithal, G, Day, P and Daly, K.A.: Association of polymorphisms in the cytochrome P450 CYP2C9 with Warfarin does requirement and risk of bleeding complications. The Lancet, Volume 353, Issue 9154, Pages 717-719.

5 Artzner P., F. Delbaen, J.M. Eber & D. Heath (1997). Thinking coherently. Risk 10, 68-71.

6 Artzner P., F. Delbaen, J.M. Eber & D. Heath (1999).Coherent measures of risk. Mathematical Finance 9, 203-228.

7 Balkema, A., and L. de Haan (1974). *Residual life time at great age*, Annals of Probability, **2**, 792-804.

8 Behr, R.A., Karson, M.J. and Minor, J.E. Reliability-analysis of window glass failure pressure data, Structural Safety, 1991, 11, 43-58.

9 Benet L.Z. and Goyan, J.E. Bioequivalence and narrow therapeutic index drugs. Pharmacotherapy 1995, 15:433-440

10 Beran, M., Hosking, J.R.M. and Arnell, N. (1986). Comment on “Two-component extreme value distribution for flood frequency analysis” by Fabio Rossi, Mauro Fiorentino, Pasquale Versace, *Water Resources Res*. 22, 263-266.

11 BestFit, Version 4.5.4. Probability Distribution Fitting software for Microsoft Windows, Palisade Corporation.

12 Bortkiewicz, L., von (1922). Variationsbreite und mittlerer Fehler, *Sitzungsber. Berli. Math. Ges.* 21, 3-11.

13 Bourne, David W.A.: Mathematical Modeling of Pharmacokinetic Data. Technomic Publishing Company, 1995.

14 Bourne, David W.A.: Re: Pharmacokinetic Software [Online] 1996 July 23 <URL: http://www.cpb.uokhsc.edu/pkin/soft.html>.

15 Broussard, J.P. and Booth, G.G. The behavior of extreme values in Germany's stock index futures: An application to intra-daily margin setting, European Journal of Operational Research 1998, 104, 393-402.

16 Brown, B.G. and Katz, R.W. (1995). Regional-analysis of temperature extremes: Spatial analog for climate-change, Journal of Climate 1995, 8, 108-119.

17 Canfield, R.V. (1975). The type-I extreme-value distribution in reliability, *IEEE Trans. Reliab*. 24, 229-236.

18 Canfield, R.V. and Borgman, L.E. (1975). Some distributions of time to failure for reliability applications, *Technometrics* 17, 263-268.

19 Chavez-Demoulin, V. and Roehrl, A. (2004): Extreme Value Theory can save your neck.

20 Chow, Shein-Chung and Liu, Jen-Pei (1992). Design and Analysis of Bioavailability and Bioequivalence Studies, Marcel Dekker, Inc.

21 Chuang, S.N. and Lloyd, H.L.:Analysis and identification of stochastic compartment models in pharmacokinetics: Implication for Cancer Chemotherapy, Math. Biosci., 22, 1974, pp. 57-74.

22 Coldwell, R. (2002): Extreme value theory and applications to flood probability calculations.

23 Coles, S. (2001). An Introduction to Statistical Modeling of Extreme Values. Springer series in statistics.

24 Cubasch, U., Meehl, G.A. (2001): Projections of future climate change. In: Houghton JT et al., editors. Climate change 2001: the scientific basis. Cambridge: Cambridge University Press; 2001. p.525-82.

25 De Haan, L. (1970). *On Regular Variation and its Application to the Weak Convergence of Sample Extremes*, Mathematical Centre Tracts 32, Mathematisch Centrum, Amsterdam.

26 Deshotels, B. and Fitzgerald M. (2001). Designing for extreme temperatures, Chance, vol. 14, no. 3, Summer issue.

27 Dodd, E.L. (1923). The greatest and least variateunder general laws of error, Trans. Amer. Math. Soc. 25, 525-539.

28 Erdman, D.: A study of kinetics: The estimation and simulation of systems of first-order differential equations. SAS Institute Inc., Cary, NC.

29 Farago, T. and Katz, R. (1990): Extremes and design values in climatology, *World Meteorological Organization*, WCAP-14, WMO/TD-No. 386.

30 FDA Guidance for Industry – Bioavailability and Bioequivalence Studies for Orally Administered Drug Products – General Considerations.

31 Fisher, R.A. and Tippett, L.H.C. (1928). On the estimation of the frequency distributions of the largest or smallest member of a sample. Procedding of the Cambridge Philosophical Society 24, 180-190.

32 Folland, C.K., Karl, T.R. (2001): Observed climate variability and change. In: Houghton JT et al., editors. Climate change 2001: the scientific basis. Cambridge: Cambridge University Press; 2001. p.99-181.

33 Frechet, M. (1927). Sur la loi de probabilite de l'ecart maximum, Ann. Soc. Polon. Math. Cracovie 6, 93-116.

34 Friesen, M.H. and Walker, S.E. Are the current bioequivalence standards sufficient for the acceptance of narrow therapeutic index drugs? Utilization of a computer simulated Warfarin bioequivalence model. J. Pharm Pharmaceut Sci, 2(1): 15-22, 1999.

35 Fuller, W.E. (1914). Flood flows, *Trans Amer. Soc. Civil Engineers* 77, 564.

36 Gabrrielsson, J. and Weiner, D. Pharmacokinetic and Pharmacodynamic Data Analysis: Concepts and Applications. Swedish Pharmaceutical Press, Stockholm, 1994.

37 Garrett, E.R.: Antibiot. Chemotherapia, 12, 227 (1964)

38 Gibaldi, M. and Perrier, D. Pharmacokinetics, Second Edition, Marcel Dekker, Inc. New York, 1982.

39 Gibaldi, M., Levy, G. and Weintraub, H.: Drug distribution and pharmacological effects, Clin. Pharmacol. Therap., 12, 1971, pp. 734-742.

40 Gibaldi, M. and Levy, G.: Dose-dependent decline of pharmacologic effects of drugs with linear pharmacokinetic characteristics, Journal of Pharmaceutical Sciences, Vol. 61, No. 4, April 1972.

41 Gilleland, E., Katz, R. and Young, G. Extremes Toolkit: Weather and Climate Applications of Extreme Value Statistics.

42 Gnedenko B.V. (1941). Limit theorems for the maximal term of a variational series. Comptes Rendus de l'Academie des Sciences del'URSS 32, 7-9.

43 Gnedenko, B.V. (1943), *Sur la distribution limite du terme maximum d'une serie aleatoire*, Annals of Mathematics, **44**, 423-453. Translated and reprinted in: *Breakthroughs in Statistics*, Vol I, 1992, eds. S. Kotz and N.L. Johnson, Springer-Verlag, pp. 195-225.

44 Griffith, A.A. (1920). The phenomena of rupture and flow in solids, *Philos. Trans. Roy. Soc. London* A, 221, 163-198.

45 Gumbel E.J. (1958). Statistics of Extremes. Columbia Univ. Press., New York.

46 Gumbel, E.J. (1954). *Statistical Theory of Extreme Values and Some Practical Applications*, National Bureau of Standards, Applied Mathematics Series, Vol.33.

47 Gurwitz JH, Avorn J, Ross-Degnan D, Choodnovskiy I, Ansell J. Aging and the anticoagulant response to warfarin therapy. Ann Intern Med 1992;116:901-4.

48 Haines, S.T., Reflection on generic Warfarin. Am J Health Syst Pharm, 55:729-33; 1998

49 Harlow, D.G., Smith, R.L. and Taylor, H.M. Lower tail analysis of the distribution of the strength of load-sharing systems, Journal of Applied Probability 1983, 20, 358-367.

50 Harrison L, Johnston M, Massicotte MP, Crowther M, Moffatt K, Hirsh J. Comparison of 5-mg and 10-mg loading doses in initiation of warfarin therapy. Ann Intern Med 1997;126:133-6.

51 Holman DJ. 2003.mle: a programming language for building likelihood models.

52 Horton J.D. and Bushwick, B.M. Warfarin Therapy: Evolving strategies in anticoagulation. Published by American Academy of Family Physicians, February 1999.

53 Ihaka, R. and Gentleman, R. (1996) R: A language for data analysis and graphics. Journal of Computational and Graphical Statistics, 5, 299-314.

54 Jenkinson, A.F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological events. Quarterly Journal of the Royal Meterorological Society 81, 158-172.

55 Johnson, M.E. (December 1997). Caribbean Storm Surge Return Periods: Final Report

56 Johnson, M.E. and Watson, Jr. C.C. (1999): Hurricane return period estimation, 10th Symposium on Global Change Studies, 10-15 Jan 1999. Dallas, Texas.

57 Kaplan, S.T.: Pharmacokinetic profile of Coumermycin A, Journal of Pharmaceutical Sciences, Vol. 59, No. 3, 1970, pp. 309-313.

58 Karl, T.R., Knight, R.W. (1998). Secular trends of precipitation amount, frequency, and intensity in the United States. Bull Am Meterol Soc 1998; 79:231-41.

59 Katz, R.W., Parlange, M.B. and Naveau, P. (2002). Statistics of extremes in hydrology, Advances in Water resources, 25: 1287-1304.

60 Kotz S. and Nadarajah S. (2000). Extreme Value Distributions, Theory and Applications. Imperial College Press.

61 Lawless, J.F. Statistical Models and Methods for Lifetime Data. Wiley series in Probability and Mathematical Statistics.

62 Lee, R.U. Statistical-analysis of corrosion failures of lead-sheathed cables, Materials Performance 1992, 31, 20-23.

63 Leslie, P.H.: A stochastic model for studying the properties of certain biological systems by numerical methods, Biometrika, 45, 1958, pp. 16-31.

64 Levy G. and Jusko, W.J.: Multicompartment pharmacokinetic models and pharmacological effects, Journal of Pharmaceutical Sciences, Vol. 58, No. 4, April 1969.

65 Longin M.F. (1996). The asymptotic distribution of extreme stock market returns, The Journal of Business, Volume 69, Issue 3 (July, 1996), 383-408.

66 Maple 9, Mathematical Software Package, Maplesoft.

67 Maximizing value from integrated PK-PD Prediction. Bayer Technology Service. http://www.bioinformaticsforumuk.net/oxbf/images/bioconvergence/talks/Joerg_Lippert_Bayer.ppt

68 Mejzler, D.G. (1949). On a theorem of B.V. Gnedenko. *Sb. Trudov Inst. Mat. Akad. Nauk Ukrain. SSR* 12, 31-35

69 Mises, R. von (1923). Uber die variationsbreite einer Beobachtungsreihe, *Sitzungsber. Berlin. Math. Ges*. 22, 3-8.

70 Mises, R. von (1954). La distribution de la plus grande de n valeurs. In *Selected Papers, Volume II*, pages 271-294. American Mathematical Society, Providence, RI. This paper was the reproduction of Mises, R, von's original paper published in *Rev. Math. Union Interbalk.* 1, 141-160 (1936).

71 Naess, A. Estimation of long return period design values for wind speeds, Journal of Engineering Mechanics-American Society of Civil Engineers 1998, 124, 252-259.

72 National Oceanic and Atmospheric Administration, National Data Center (http://www.ncdc.noaa.gov)

73 Pearson JD, Morrell CH, Brant LJ. 1992. Mixture models for investigating complex distributions. J Quant Anthroplo 3:325-345.

74 Pickands, J. (1975). Statistical inference using extreme order statistics. Annals of Statistics 3, 119-131.

75 R, Statistical Computing System

76 Ramachandran, G. (1982). Properties of extreme order statistics and their application to fire protection and insurance problems, Fire Safety Journal 1982, 5, 59-76

77 Rao, N.M., Rao, P.P., Kaila, K.L. The first and third asymptotic distributions of extremes as applied to the seismic source regions of India and adjacent areas, Geophysical Journal International 1997, 128, 639-646.

78 Rossi, F. , Fiorentino, M. and Versace, P. (1986). Two-component extreme value distribution for flood frequency analysis, *Water Resources Res*. 22.

79 Rossi, F. (1986). Reply to "Comment on 'Two-component extreme value distribution for flood frequency analysis'", *Water Resources Res*. 22, 267-269.

80 Smith, R.L. (1985). Maximum likelihood estimation in a class of non-regular cases. Biometrika 72, 67-90.

81 Smith, R.L. (1987). Estimating tails of probability distributions. The Annals of Statistics 15, 1174-1207.

82 Smith, R.L. and Weissman, I. (1985). Maximum likelihood estimation of the lower tail of a probability distribution. J. Roy. Statist. Soc. B, 47, 285-298.

83 Smith, R.L. and Davison, A.C. (1990). Model for exceedances over high thresholds (with discussion), J. Roy. Statist. Soc. B, 52, 393-442.

84 Smith, R.L. (2003). Statistics of extremes, with applications in environment, insurance and finance. Department of Statistics, University of North Carolina, Chapel Hill. Web reference:

85 Solomon, H.: Unpublished data on file, Hoffmann-La Roche, Nutley, N.J., 1968.

86 Soong, T.T., Pharmacokinetics with uncertainties in rate constants I, II, and III the inverse problem, Math. Biosciences, 12, 1971, pp. 235-243.

87 Soong, T.T.: Random Differential Equations in Science and Engineering, Vol. in Mathematics in Science and Engineering, Academic Press, New York, 1973.

88 Soong, T.T.: Solutions of a class of random differential equations, SIAM J. Appl. Math., Vol. 24, No. 4, June 1973.

89 Southeast Regional Climate Center (http://www.sercc.com)

90 S-PLUS functions for extreme value modeling: *An accompaniment to the book An Introduction to Statistical Modeling of Extreme Values* by Stuart Coles.

91 Statistical Analysis Software (SAS), Version 9.1. SAS Institute Inc., Cary, NC, USA.

92 Teorell, T.: Kinetics of distribution of substances administered in the body, Arch. Intern. Pharmacodyn. Therap. 57 (1937), 205-240.

93 The Merck Manual of Diagnosis and Therapy, Section 22: Clinical Pharmacology, Chapter 299: Pharmacokinetics; Chapter 303: Monitoring Drug Treatment

94 Trenberth, K.E. (1998). Atmospheric moisture residence times and cycling: implications for rainfall rates and climate change. Climatic Change 1998; 39: 667-94.

95 Trenberth, K.E. (1999). Conceptual framework for changes of extremes of the hydrological cycle with climate change. Climatic Change 1999; 42: 327-39.

96 Tsokos, C.P. and Nadarajah, S. Extreme value models for software reliability.

97 Tsokos, J.O. and Tsokos, P.T.: Statistical modeling of pharmacokinetic systems, J. Dynamic Systems, Measurement and Control, ASME, March 1976

98 Web reference: http://www.dml.co.nz/yourtest/inr.asp

99 Wettstein, Justin J. and Mearns, Linda O. (2002). The influence of the north atlantic-arctic oscillation on mean, variance and extremes of temperature in the northeastern United States and Canada. Journal of Climate, 15:3586-3600, 2002.

100 Xapson, M.A., Summers, G.P. and Barke, E.A.. Extreme value analysis of solar energetic motion peak fluxes, Solar Physics 1998, 183, 157-164.

101 Yasuda, T. and Mori, N. Occurrence properties of giant freak waves in sea area around Japan, Journal of Waterway Port Coastal and Ocean Engineering-American Society of Civil Engineers 1997, 123, 209-213.

Bibliography

1 Aarons, L.: Pharmacokinetic and pharmacodynamic modelling in drug development. Stat. Methods Med. Res. (editorial) - 1999, Vol 8, pp 181-182.

2 Bellman, R., Topics in pharmacokinetics, III: Repeated dosage and impulse control, Math. Biosci., 12, 1971, pp. 1-5.

3 Bharucha-Reid, A.T.: Random Integral Equations, Vol. in Mathematics in Science and Engineering, Academic Press, New York, 1972.

4 Lain Glen: Pharmacokinetic Analysis, Anaesthesia and intensice care medicine, Vol. 6, Issue 8, pp 280-282, 2005.

5 Lotsch J, Kobal G, Geisslinger G.: Programming of a flexible computer simulation to visualize pharmacokinetic-pharmacodynamic models. Int. J. Clin. Pharmacol. Ther. 2004 Jan: 42(1):15-22.

6 Niazi, S.: Multicompartment pharmacokinetic analysis and simulations using a programmable calculator. Int. J. Biomed. Comput. 1979 May: 10(3):245-55.

7 Russell, C.D.: A Bayesian 3-Compartment Model for ^{99m}Tc-MAG3 Clearance: Journal of Nuclear Medicine (2003), Vol. 44 No. 8 1357-1361.

8 Schoenwald, R.D.: Pharmacokinetic Principles of Dosing Adjustment, Technomic Publishing Co. Lancaster, 2001.

9 Shargel Leon, Yu Andrew B.C. Applications of Computers in PK. In: Applied Biopharmaceutics and PK. 3rd ed. New York: Prentice-Hall International Editions, 1993:571-583.

10 Thomas P. Lodise Jr., Ben Lomaestro, Keith A. Rodvold, Larry H. Danziger, and George L. Drusano: Pharmacodynamic Profiling of Piperacillin in the Presence of Tazobactam in Patients through the Use of Population Pharmacokinetic Models and Monte Carlo Simulation. Antimicrobial Agents and Chemotherapy, December 2004, p. 4718-4724, Vol. 48, No. 12.

11 Tsokos, C.P. and Padgett, W.J.: The origin and application of stochastic integral equations, Int. J. Systems Sci., 2, 1971, pp. 135-148.

12 Tsokos, C.P. and Padgett, W.J.: Random Integral Equations with Applications to Life Sciences and Engineering, Vol. in Mathematics in Science and Engineering, Academic Press, New York, 1974.

About the Author

Lakshminarayan Rajaram received a bachelor's degree in Physics and Mathematics, and a master's degree in Pure Mathematics from the University of Mysore, India. He also received a master's degree in Applied Mathematics from the New Jersey Institute of Technology, Newark, N.J. in 1986. He worked as a biostatistician for clinical trials in the Division of Endocrinology and Metabolism at the Johns Hopkins School of Medicine, and later, in the Clinical Research Division at TheraTech, Inc., Utah and Baker Norton Pharmaceuticals, Miami. He also worked as a data analyst at the Institute on Aging, University of South Florida. At present, he is a tenured faculty in mathematics and statistics at the St. Petersburg College, Tarpon Springs Campus. He is an adjunct faculty of biostatistics in the Department of Epidemiology and Biostatistics, University of South Florida. He also is an independent consultant in the database design, data management and biostatistics for health-related studies including pharmaceutical clinical trials. Additionally, he gives lectures and workshops in the pharmacokinetics, biostatistics, and SAS (Statistical Analysis Software) programming both in the United States and abroad.