Investigating the threshold of event detection with application to earthquake and operational risk theory.

by

Corné Juan de Witt

Submitted in partial fulfilment of the requirements for the degree
Master of Science (Actuarial Science)
in the Faculty of Natural & Agricultural Sciences
University of Pretoria
Pretoria

2013/12/31

Abstract

This study provides systematic analysis of points of structural change in probability distributions. In observed frequency data of earthquakes, such a threshold exists due to the non-detection of events below a certain magnitude. By examining the factors influencing the operational risk exposure of institutions, a similar threshold is hypothesized to exist in operational loss data. In both fields of study, this threshold is termed the threshold of completeness, above which 100% of events are detected. External factors have caused this level of completeness to shift over time for earthquake data. The level of complete recording influences the volume of data that can be consistently incorporated in a study of seismic activity. Such data can be used by re-insurers and direct writers of catastrophe agreements who deal in seismic hazard. Historically, a variety of methods have been proposed by authors in an attempt to gauge the location of the threshold of completeness in earthquake data. This study aims to evaluate the efficacy of some of the most prominent methods under differing assumptions regarding the incomplete portion of the data. Furthermore, a new threshold estimation scheme (MITC) is developed and tested against the prevailing methods. Additionally, earthquake data and the wealth of literature will aid in introductory analysis and assessing applicability of estimation techniques in the context of operational losses.

Keywords: Magnitude of completeness, Event detection, probability of event detection, operational risk, Moment incorporating threshold computation, event detection threshold.

Supervisor : Prof. A. Kijko
Co-Supervisor : Dr. C. Beyers
Department : Actuarial Sciences
Degree : Master of Science

Declaration

I, **Corné Juan de Witt** declare that the thesis/dissertation, which I hereby submit for the degree **Master of Science (Actuarial Science)** at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

SIGNATURE: .....

DATE: .....

Acknowledgements

I would like to thank the following contributors, knowing that without them this work would have not been possible :

- I thank the Lord, for giving me the opportunity to have started with this work and for providing me with the strength to carry out the task.
- To my darling, loving, wife, Bernice, I thank you for your continual support, encouragement.
- To my mother and father, I would like to thank you for your support and encouragement throughout this time.
- To my supervisor, Prof. A. Kijko, and co-supervisor Dr. C. Beyers, thank you for your guidance and comments.
- I thank Dr. Jochen Woessner for supplying me with the earthquake event datasets.

Contents

List of Figures

List of Tables

1 Introduction

"Everything flows, nothing stands still"
- Plato (Cratylus 402a)

This principle prevails in the work of countless authors and has shaped their views. Furthermore, as academic advancements have been made, the understanding of the influence of said philosophical cornerstone has grown immensely. Over time, systematic changes in key characteristics of a large number of processes have been observed. With the groundwork being layed more than 2 millennia ago, rigour and quantification is what might be described as the challenge of the day.

Building on this principle, one well-posed problem in modern day academia, is that in the study of time-series. This is known as the change-point problem [5], where a process can undergo considerable change at certain unknown time(s). Examples of such changes include [9]:

Genetic data analysis - Change point models can be used to quantify DNA copy number variations.

Traffic mortality rates - Traffic death rates could possibly have changed starting from 1987, due to an increase in the speed limit from 55 to 65 miles per hour on certain roads in the United States.

Quality Control - Measurable quality of products are expected to remain constant during a production process, if the process is indeed under control and stable. Points of significant change in the quality of the produced items may exist. Identification thereof can help ascertain the stability of the overall process.

Stock market analysis - Apart from stock price fluctuations, as predicted by normal economic theory, systematic developments can cause further movements [9]. Structural changes of processes, in this particular field, have received much attention and therefore a rich literature is available. Examples include methods for detecting multiple change points in multivariate time series, with specific application in financial markets [34]. A number of authors affirm the existence of structural breaks in parameters of these processes [5].

If the time domain of these processes were replaced by another quantitative measure, such as size, an adaptation of the change point problem can be studied in a plethora of fields. Two such extensions of the change point problem can be seen in the studies of seismic activity related to the realization of earthquakes and the other being the analysis of operational risk data.

In the context of earthquake data, shifts can be observed in a number of different manners: From changes in the rate of seismic event occurrence to changes in the proportion of large to smaller generated events. If one considers the stochastic process relating to detected earthquakes, rather than generated earthquakes, it can easily be hypothesized that this process might also be subject to some shifts over time, space and event size.

This study will focus on event size change-points. Typically, some characteristic of the distribution is altered at the change point. The magnitude of completeness is a specific case of a change-point in earthquake data. This point carries importance, since events equal to or larger than the change-point will be detected with a probability of 1, whereas, events with sizes below the threshold value might not be detected. This problem can therefore also be described as the estimation of a threshold.

Operational risk data, as part of the risk quantification process, carries the possibility of having various distributional thresholds. It can fairly easily be hypothesized that such data is subject to the existence of an analogous detection threshold. Academic examples of such a threshold relates to extreme events and has also received academic attention [12, p.87].

The notion of "complete reporting" can therefore be introduced. In this context, a dataset (or subset thereof) that is described as being completely reported on, is to be characterized in that all events that have realized, have indeed been reported on. When considering data where the numbers of recorded events are not indicative of the entire process, the "level of complete reporting", i.e. the event detection threshold, is the event size that differentiates between a subset that has been completely reported on, and the subset that has not. This is a generalization of the magnitude of completeness, that specifically refers to earthquake data.

2 Motivation

Two, relatively disjoint, academic fields can both benefit from a central study of threshold estimation, namely earthquake risk as well as operational risk quantification. Intuitively, smaller earthquakes as well as operational losses can be harder to detect than larger sized (valued) events. This broadly described relationship, where smaller events attract a larger probability of not being detected, is the focus of the study.

A vast number of possibilities exist relating to the structure of the probability distribution of event detection. In some instances the probability of not detecting an event is non-zero for any event size, this is termed a "soft detection threshold". This concept will be revisited shortly, together with the idea of a "sharp detection threshold". This second type of threshold can be hypothesized to exist, above which all generated events are indeed detected. In practice, both of these situations are important. This study will mainly focus on "sharp detection thresholds", but "soft detection thresholds" will form part of an auxiliary investigation into the efficacy of threshold estimation.

As will follow from more a detailed review in Section 2.2, obtaining sufficient operational risk data presents a number of difficulties. Hence, this paper will incorporate an investigation of earthquake data. The availability of such datasets is slightly better than for operational risks. By undertaking such a cross-discipline study, a general framework of threshold estimation can be developed and the efficacy of the resulting estimation methods can be gauged. From this point modifications can be made specific to the proposed distribution of operational losses. As stated, this derived processes, relating to the size and frequency of observed events (rather than generated events) has received considerable attention in the seismological community. Such a wealth of literature will further the understanding of threshold estimation techniques.

Sections 2.1 and 2.2 of this chapter will focus on the importance of accurate threshold estimation in probability distributions. The importance in these respective fields will be stressed from an actuarial perspective. This mainly relates to how subsequent actuarial calculations will be impacted based on the end results of the threshold estimation exercise.

This study also advocates the modification of probability distributions used when modelling earthquake and operational risk data. A brief outlining motivation is given in Section 2.3 for such modifications.

Section 2.4 will provide motivation for the development of the newly proposed MITC threshold estimation scheme.

2.1 Actuarial motivation for the analysis of earthquake data

As described in the Introduction (Chapter 1) the magnitude of completeness is an example of a change point, i.e. distributional threshold, that exists in earthquake event data. This point is defined as the lowest magnitude at which 100% of the events in a space-time volume are detected [46]. The reliability of a number of subsequent analyses, for example, seismicity- and hazard rate analyses, rests on the accurate determination of this threshold value [60]. Traditionally, the motivation is that the magnitude of completeness determines the specific range, from the available data, that can be used to estimate the values of other distributional parameters [60].

Affirming the importance of such distributional parameters, is their routine use by direct insurance business writers and re-insurers in the calculation of premiums and reserves for, e.g., catastrophe insurance and reinsurance agreements [7].

A variable that readily depends on the magnitude of completeness is the b-Value, found in the Gutenberg-Richter relation [22], classically expressed as

$$\log_{10} N(M) = a - bM \tag{2.1}$$

where a (seismicity rate, or productivity, in a given period) and b (ratio of large to small events) are positive constants; M represents the magnitude of a seismic event and $N(M)$ is the number of seismic events that are equal or larger than magnitude M.

This is a prime example of a cumulative frequency-magnitude distribution (Cumulative FMD). The log transformed quantity in (2.1) is a linear function of magnitude representable by a straight line, as can be seen in Figure 2.1 for illustrative values of the parameters a and b. The basis for a cumulative FMD is that the total number (frequency) of events that have magnitudes greater or equal to a magnitude M is plotted against the specific magnitude M. The cumulative frequencies displayed on the cumulative FMD are typically presented on a logarithmic scale.

The importance of accurate estimation of the magnitude of completeness has increased over time and is said to become even more important. Activities of man can induce earthquakes, leading to structural changes in the natural physical processes of certain regions. Examples of such are, reservoir triggered earthquakes in India [17] as well as shale gas extraction through hydraulic fracturing (fracking). Such induced earthquakes have already been observed in the United States and could also become a local reality due to the possibility of hydraulic fracturing practices being implemented in the Karoo region of South Africa. Over time such activities can introduce inhomogeneous factors into recorded data. This

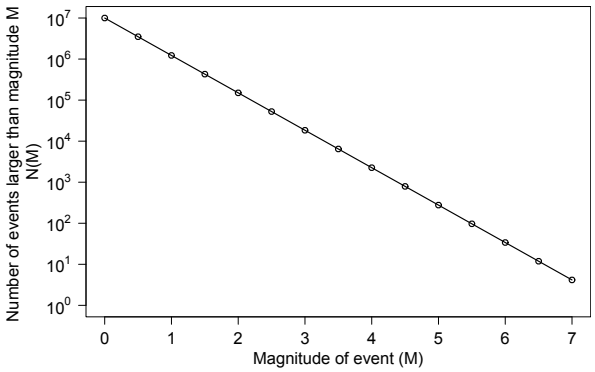


Figure 2.1: *Graphical representation of Gutenberg-Richter relation. Illustrative parameter values of* $a = 10^7$ *and* $b = 0.91$.

can ultimately alter the level of completeness by changing, for example, seismicity rates. Subsequently the importance of accurately estimating distributional thresholds in data is increased.

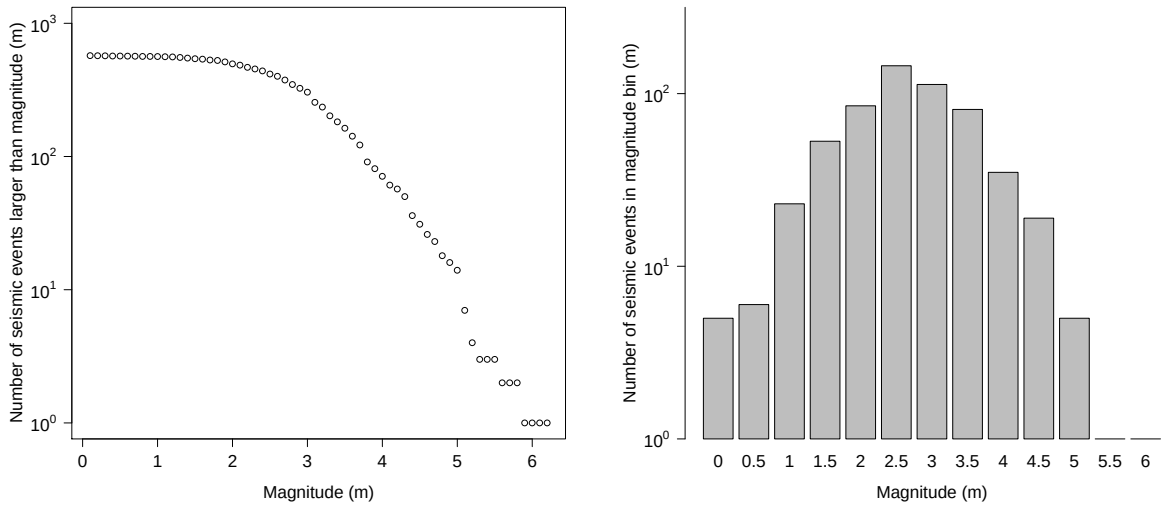
In the case of the Gutenberg-Richter frequency-magnitude distribution, such distributional change points can easily be visualized as the deviation from the straight line relation in (2.1). Figure 2.2a gives a prime example of such a visualisation of a change point. The deviation from the pattern predicted in Figure 2.1 is readily identified. Furthermore, as will be shown in (8.1), page 23, the Gutenberg-Richter relation in (2.1) can also be interpreted as a probability density. This coincides with the notion of an incremental Frequency-Magnitude Distribution (incremental FMD). For the case of the incremental version, events are grouped and counted by magnitude bins. This is done in much the same way as a histogram is constructed. What typically distinguishes an incremental FMD from a histogram is that frequencies are displayed on a logarithmic scale. In this case, the distributional threshold can be observed as the deviation from the linear pattern for bins representing progressively smaller magnitudes. This can be seen in Figure 2.2b. These illustrated anomalies in both cumulative and incremental FMDs can be linked to the level of complete reporting.

Although the distributional threshold can be identified, by visual inspection of Figure 2.2, to be near a magnitude of 2.5, this approach lacks objectivity and precision. These needs form the basis to identify accurate methods for automatically estimating distributional thresholds.

2.2 Actuarial motivation for the analysis of operational risk data

As written up in a 2007 working paper of the International Monetary Fund [31], operational risk can be broadly defined as

> "The risk of loss or some adverse outcome, such as financial loss, resulting from acts



(a) *Cumulative FMD* **(b)** *Incremental FMD*

Figure 2.2: *Frequency-magnitude distribution for earthquake events recorded in Western Freestate, South Africa (S 28°11′34′ to S 30°54′20′; E 20°24′28′ to E 29°37′35′). Data has been collected from various sources, the main source being the ISC (International Seismological Centre) for the period 1971-2005.*

> undertaken (or neglected) in carrying out business activities, such as inadequate or failed internal processes and information systems, from misconduct by people (e.g. breaches in internal controls and fraud) or from external events (e.g. unforeseen catastrophes)"

Events pertaining to operational risk and the subsequent losses form part of an integral process aimed at describing the risk exposure of a corporate entity. The approach to managing and making provisions for operational risk exposure, relating to banking activities, are further described by the Basel II framework. In contrast, South African insurers will have to comply with the regulatory framework as set out in the Solvency Assessment and Management (SAM) Roadmap. SAM is the risk-based regulatory framework of the Financial Services Board (FSB) that is modelled after its European counterpart, Solvency II.

In order to estimate this exposure, under any regime, the efficiency at which event realizations are detected must be assessed. A well founded example of events that warrant the incorporation of detection probabilities are fraud and theft. If the detection probability of such financial losses are assumed to be influenced by the monetary amount (size) of the loss, the existence of a distributional threshold of complete reporting can fairly easily be hypothesized. Analogous to the magnitude of completeness this threshold is such that all losses of monetary value greater than the threshold value are detected.

Knowledge of completeness in reporting can enable organizations to actively determine the efficiency of controls and measures implemented aimed at reducing operational risk. Furthermore, this puts organizations in a position to better understand their exposure to reported versus unreported operational losses, enabling pro-active measures to be put in place.

Publicly available operational risk data has been described as "notoriously scarce" [31] and "hard (if not impossible) to come by" [57]. Such shortages of data can be explained by the following causal factors :

1. A shortage of industry-wide initiatives for the collection and ultimate provision of operational risk data to participating parties and researchers. South Africa is a prime example of such a situation.

2. Internally and/or externally gathered data will only be adequate when sufficient time has elapsed for the realization of event data.

3. Recording of operational risk events may be subject to an inadequate detection process, e.g. unreported theft or fraud.

Such difficulties are the main motivation for considering earthquake data when undertaking this study.

2.3 Motivation for modification of probability distributions used for event severity

It is generally understood that any study needs to identify a suitable time window to be used. The aim is to eliminate any inhomogeneities that have been introduced into the data due to the passage of time. Such factors will be discussed in Sections 6.1.2 and 6.1.3. By restricting, and possibly shortening, the time window of investigation, the amount of data to be used is reduced. This especially raises concern for the inclusion of extreme events. Such unease is predominantly due to extreme events taking a large amount of time to sufficiently accumulate in data.

In order to keep the time horizon spanning the investigation as long as possible, and therefore maintain the volume of extreme events in the data, some adjustments are necessitated. The main consideration for consistency of the data is that the relationship between event realizations must be maintained. An example of such a relationship can be seen in (2.1), the Gutenberg-Richter relation, as well as subsequently derived probability density functions. For this relationship to hold, the data that is known to be inconsistent (e.g. as a result of time-dependent inhomogeneities) must be excluded. This is the traditional approach, but once again, reduces the amount of usable data and infringes on statistical reliability.

A different *modus operandi* would be to specify an alternative relationship, and subsequent probability distributions, governing the process. Through this approach, the amount of usable data will not be reduced and the investigation time window will not have to be shortened whereby extreme events are excluded. This approach is ultimately advocated in the current study and furthermore forms the motivation for undertaking an investigation into modified probability distributions used to model event severities. This will be discussed in more detail in Section 8.2.1.

2.4 Motivation for development of MITC estimation scheme

When considering the prevailing methods of threshold estimation used in the field of earthquake data analysis, only the Entire Magnitude Range (EMR) method [60] explicitly allows for the use of a detection probability. This spurred the development of a new method, that would also be able to allow for the explicit modelling of a detection probability. Furthermore, due to the restrictions on data availability, the explicit inclusion of a detection probability can increase the reliability of statistical results. This is due to the fact that not all data on the process is needed to be able to fit a distribution, but only the data on the detected events.

As will be shown in this study, the EMR method is a powerful approach to threshold estimation. However, implementation of the algorithm can be computationally intensive and lengthen the time taken to conclude a study. This has been a further motivation for the implementation of the MITC scheme, as a way to reduce computation time while still retaining much of the associated estimation power.

3 Objectives

This study is aimed at

- Establishing parallels between operational risk and earthquake magnitude events.
- Formalizing incorporation of a detection probability into a data generating distribution.
- Establishing preliminaries and formalizing some of the prevailing threshold estimation techniques. Additionally, a new threshold estimation scheme is developed. The efficacy of this new scheme, termed Moment Incorporating Threshold Computation (MITC), will also be determined relative to that of the other studied methods.
- Accurate estimation of the distributional threshold, termed the level of complete reporting, in presented data.
- The objective comparison of estimation methods must indicate the most appropriate method for varying circumstances:
  - Given a dataset with a specific form of the detection probability. This is a situation that could present itself based on prior information such as a separate investigation into the detection capabilities of a network or simply subjective beliefs.
  - Given a dataset with a specific level of completeness, i.e. a known value of the detection threshold. This situation should be very rare in practice, since this is particularly the value that is being estimated. However, from a theoretical perspective, results from such an enquiry will aid in understanding the underlying efficiency of the estimation methods.
  - Given a dataset with unknown characteristics. This situation should present itself in a typical fashion for most researchers. Here a researcher might have subjective beliefs regarding the dataset, but will most typically still have to conclude diagnostic tests on the dataset.

4 Outline of dissertation

- In **Chapter 2** the importance of taking distributional thresholds into account has briefly been discussed and a statement of the objectives of this study has been made in **Chapter 3**.
- The similarities between operational risk and earthquake data will be detailed in **Chapter 5**. It is also seen how the problem of estimating a distributional threshold can be generalized to a multi-disciplinary level.
- **Chapter 6** will provide a brief overview of the respective seismological and operational risk contexts in which the problem of threshold estimation will be discussed.
- In **Chapter 7** the assumptions underlying the study will be stated as well as the limitations that are encountered.
- In order to gauge the relative performance of the threshold estimation methods, **Chapter 8** will be devoted to deriving statistical distributions that will represent real-world data. The distributions that are derived will, however, represent a "perfect" world. Additionally, similar distributions will be derived incorporating a "soft detection threshold". The aim is to compare estimation methods in less than ideal conditions. Consequences of modelling inadequacies will subsequently be realized.
- In **Chapter 9** the most prominent methods for estimating the distributional threshold will be discussed. This will include an assessment of underlying assumptions which leads to the resultant weaknesses and strengths. A brief demonstration of each estimation method will be included. Furthermore, the newly proposed MITC estimation method is also derived in this chapter.
- Rigorous evaluation of the threshold estimation methods will be done in **Chapter 10**. This will be undertaken considering varying circumstances of the data generating process. The questions relating to which is the most appropriate estimation method for varying circumstances, as stated in the objectives of the study, will also be addressed in this chapter.
- In **Chapter 11** the threshold estimation methods will be compared by analysing real-world earthquake data in order to demonstrate practical application. **Chapter 12** will be devoted to the consideration of operational risk data on which threshold estimation techniques will be applied.
- **Chapter 13** will present a summary and conclusion on the effectiveness of the methods. In addition, objectives will be outlined for future follow up studies.
- **Appendices ...**

5 Cross discipline parallels of earthquake data and operational risk data

5.1 Motivation for assessing similarities

As mentioned, the scarcity of historical operational risk data is one of the main stumbling blocks for institutions when attempting to construct comprehensive models. Furthermore, it is stated that even the largest banks have only in the vicinity of 5 to 6 years of loss data [11]. Even though such estimates were obtained in 2006, the data can still be considered insufficient in the context of extreme events. The problem is further exacerbated due to data being subject to external factors leading to inhomogeneities.

For these reasons it is proposed to utilize earthquake data when assessing the techniques of threshold estimation. When compared to operational risk data, earthquake data can be seen to be more freely available and the time horizon of examinable events is longer. To further substantiate this act of cross discipline data analysis, a brief discussion of the data similarities is undertaken in the following section.

5.2 Similarities

A number of parallels can be drawn between data gathered for the analysis of operational and earthquake risk.

1. **Distribution of event severities**

   In general, risk can be quantified to be the effect and interaction between incident frequencies and the severity of such incidents. Specific constraints placed on the a process lead to the manner in which these two factors must be combined to obtain a quantitative measure of the risk.

2. **Incomplete data**

   Operational losses, only larger than specified monetary amounts, e.g. US $ 10,000 or €5,000, are recorded on bank's databases [10]. This truncation point is comparable to the quantity of m_{min} as seen in earthquake data.

   Distinction can be made between occurrence of events and events that are recorded, i.e. detected. Furthermore, the probability of an event to be recorded can be seen as a function of event size, with larger events having a proportionally larger probability of being detected.

3. **Alternative characterization of detection threshold**

   In earthquake literature the magnitude of completeness has, up to this point, been modelled as a sharp detection threshold. As will be shown in Section 8.2.1, a soft detection threshold can be introduced. It can however be seen that in operational risk the opposite situation has implicitly been encountered, illustrated by the following general example. In operational risk modelling, the

Log-Normal distribution has been a popular choice for severity distributions [44]. However, others argue that tail events are understated by the Log-Normal distribution and therefore authors have developed models such as the composite Pareto-Log-Normal model. This specific distribution is characterized by the body of a Log-Normal and the tail of a Pareto distribution, a combination that allows for a tail that is heavier than the traditional Log-Normal distribution. It can be seen that the density function of any Log-Normally distributed random variable X can be expressed as the combination of a Pareto severity distribution and a scaling function comparable to that of a detection probability :

$$\begin{aligned} f_X(x) &= \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(\ln x-\mu)^2} \qquad \text{for } x>0,\ \mu\in\mathbb{R} \text{ and } \sigma>0 \\ &\propto \left(x^{-(1-\frac{\mu}{\sigma^2})}\right)\left(e^{-\frac{1}{2\sigma^2}(\ln x)^2}\right) \end{aligned} \tag{5.1}$$

The second factor in (5.1), which can be interpreted as the detection probability, is a function that is increasing for $x<1$ and decreasing otherwise. The effect is that the upper tail of the distribution becomes lighter than that described by a power law and therefore resulting in reduced probability of extreme events.

Such behaviour is unwanted since the modelling of low frequency, high severity events is generally considered of utmost importance. This phenomenon is only realized when the scaling factor (detection probability) acts upon the entire support of the random variable X, as is the case when defining a soft detection threshold. This further motivates the use of a sharp detection threshold and that the use of a soft detection threshold be limited. The need for a soft detection threshold must be established through thorough investigation of the studied data. This will ensure that studies are not compromised by a reduced probability of extreme events where such events are of critical importance.

4. **Extreme events**

   For both earthquake and operational risk analysis the estimation of extreme events is a subject which attracts considerable attention. The impact of such low frequency - high severity events, can substantially add to the volatility when establishing risk exposure. Data scarcity, especially when considering longer time horizons (as can be seen in both fields), complicates the study of such events [11].

5. **Largest possible event**

   Continuing in the study of extreme events, the concept of a maximal event magnitude has attracted considerable attention in the study of earthquakes. For an example of such threshold estimation relating to maximum earthquake magnitude events as well as further references, see the 2004 study made by Kijko [32]. It can be seen that fairly a small amount of attention has been awarded to

studies of such maximal events in operational risk analysis. However, due to the finite nature of institutional exposure such a maximal event magnitude can be hypothesized to exist in operational risk.

This might however be of more use for internal understanding of the business and management purposes than fit for regulatory uses. This is due to the possibility of regulatory bodies being sceptic on the applicability of a point estimate, such as a maximal event.

6 Data gathering and existence of distributional threshold

6.1 Earthquake Data

Attention is first given to earthquake data and the existence of distributional thresholds induced by varying detection capabilities.

6.1.1 Earthquake catalogues as a source of data

Earthquake catalogues give a recorded listing of earthquakes and several of their most prominent features, such as the time of occurrence, physical location and magnitude. As such, these catalogues are one of the most important products of seismology [60]. Due to seismic event recordings, dating back as early as the 12th century B.C. [35], modern day earthquake catalogue equivalents have been made possible. One such document, published in 1956 named (The ∼) "Chronological Tables of Earthquake Data of China" [13], describes earthquake occurrences from 1189 B.C. up to A.D. 1955.

Since seismic events with large magnitudes are seen as rare occurrences, such events fall in the tail of the relevant probability distribution. Therefore, in order to accurately ascribe a monetary value to the act of covering risks under catastrophe (re-)insurance agreements, the volume of available data studied must be maximized. Unfortunately instrumental data dates back, at best, to the start of the 20th century [35]. For this reason it can be envisaged that datasets be supplemented by data from other sources, e.g. paleoseismic data. Inclusion of such pre-historic earthquake data through the study of paleoseismology might aid in the understanding of extreme earthquakes. Paleoseismic data represents earthquakes that need hundreds, or even thousands, of years worth of stress to be built up before occurring [35]. These type of events are registered when undertaking investigations of excavation trenches. However, through such augmentation, the possibility of data distortion exists. This unwanted consequence can arise due to the disproportional representation of earthquakes with respect to event size. Such resulting datasets would be contrary to the Gutenberg-Richter relation in (2.1) and could induce further distributional thresholds.

In theory, such distributional thresholds can be allowed for in similar ways as described herein. However, this remains to be investigated further and will not be directly considered in the current study.

6.1.2 Data quality of catalogues

Earthquake catalogues, although conceptually straightforward, hold data that carry a wealth of complexities. This should not be unexpected, since they are an aggregation of data that have been collected through human engineered systems. These systems include instruments, communication lines, computer programs as well as natural human involvement [23]. Alterations in any of the aforementioned proce-

dures and/or methods can potentially introduce inhomogeneities into the data gathered from the true underlying physical process. Examples of such changes in earthquake catalogues are:

- Increases or decreases in the rate of seismic events.
- Increases or decreases in the magnitude readings of individual seismic events, resulting in detected magnitudes that differ from the actual generated event magnitude.

These effects may be seen over the whole of the magnitude spectrum, or be localized to certain smaller subintervals.

In a 1987 article Habermann [23] categorized the types of deterioration in data quality due to human influences. Some of the factors that pertain to distributional thresholds are highlighted below. In addition, possible causal factors for each of the perceived changes are described.

1. Detection Changes : Changes that are related to the capability of the network to recognize and locate seismic events.
   - Increases in registered earthquake event rates:
     - Possible cause:
       * Newly installed stations that lead to increased reporting of smaller events in the region surrounding the new stations.
       * Improvements in the methods of analysis [59]
     - Possible impact:
       * The reporting of events larger than some level does not increase (such events are detected by the existing network).
   - Decreases in registered seismic rates:
     - Possible cause:
       * The closure of stations, leading to the non-detection of events, that would otherwise have been logged.
       * However counter intuitive, detection decreases in one region of a local catalogue, can be related to installation of stations in another region and the associated increase in workload.
2. Reporting changes: Changes relating to the lack of magnitude reporting for detected events.
   - Small systematic increase in the magnitudes assigned to events [24].
   - Some events can be detected and listed in catalogues, but magnitudes are not assigned.
3. Shift in magnitude: Systematic changes in the magnitudes.

- Such changes are similar to reporting changes. The difference, however, is that the assigned magnitudes shift by some small amount (e.g. 0.1 to 0.5 units) above or below the true value [25]. This can be seen as the inclusion of a magnitude error that is not centred around 0.

The current investigation will consider detection changes as the main confounding factor. This will be discussed further in the following section.

6.1.3 Further discussion : Impact of changes in detection capabilities

Improvements over the last few decades in seismic networks have led to increased detection of low magnitude earthquakes [48]. Expansion of a seismic network is a time dependent process in itself. As an example, Table 6.1 details operational time windows for the instrumental detection of seismic events in South Africa during the period 1910 to 1981.

Instrument	Location and Institute	Period of Operation
Obsolete type	Kimberley, De Beers Meteorological Station	Not known
Wiechert horizontal	Johannesburg Union Observatory	1910 to 1972
Milne-Shaw horizontal	Cape Town Royal Observatory	1920 to 1931
Milne-Shaw horizontal	University of Cape Town	1928 to 1947
Wiechert horizontal	Johannesburg, Bernard Price Institute of Geophysical Research	1938 to 1939
Benioff short-period vertical.	Grahamstown, South African Geological Survey	1949 to 1963
Willmore short period WWSSN 3-component short and long period Milne-Shaw horizontal (x2)	Hermanus Magnetic Observatory	1950 to 1983
Benioff short-period vertical Willmore short period	Pietermaritzburg, South African Geological Survey	1950 to 1972
Benioff short-period vertical. Sprengnether short-period vertical	Kimberley, South African Geological Survey	1951 to 1972
Benioff short-period vertical. Willmore short period WWSSN 3-component short and long period	Pretoria (moved to Silverton), South African Geological Survey	1963 to 1981

Table 6.1: *Seismic Instrumentation Deployed in South Africa During the Period 1910 to 1981 [47].*

Interpreting the time-dependent nature of network detection capabilities (Table 6.1) together with the discussion in the previous section (Section 6.1.2), it can be seen that complete reporting of events varies as a function of time. Therefore, when conducting a study, a suitable time window of investigation will be identified. As discussed in Section 2.3, the resulting aim of such an identification will be to eliminate any inhomogeneities in the data that has been introduced as a result of the passage of time. Unfortunately, this can reduce the amount of data relating to extreme events. For this reason, the importance of threshold estimation is reiterated, together with the subsequent modification of probability distributions

to account for changes in the underlying process.

6.2 Operational risk data

Two main sources of operational risk data exist for any organization wishing to conduct a study, namely internal and external operational risk data. Where internal data would consist of data collected by the organization itself, external data can be contributed by a number of different entities. Examples of such external data sources include [21, p. 43]:

1. The Global Operational Loss Database (GOLD) by the British Bankers Association (BBA)
2. The Operational Risk Insurance Consortium (ORIC) by the Association of British Insurers (ABI)
3. OpBase by Aon Corporation
4. The operational risk database maintained by the Operational Riskdata eXchange Association (ORX).

Due to information being obtained from a large number of sources, differences in collection processes and business management cultures, it can easily be hypothesized that available operational risk data is influenced by various factors. Such inhomogeneities might be similar to those found in earthquake event data. Ultimately, these factors have yet to be fully documented and their impact quantified. What follows is a brief discussion of some causal factors for incomplete reporting in operational risk data.

6.2.1 Causal factors relating to fluctuating operational risk exposure

A number of factors can impact the exposure and level of susceptibility of activities to operational risk. Some examples from the banking industry are set out as follows [31]:

1. Globalization and deregulation of financial markets
2. Growing complexity in banking industry.
3. Large-scale mergers and acquisitions
4. Increasing sophistication of financial products
5. Greater use of outsourcing arrangements

While some types of operational risks are measurable, such as detected fraud or system failure, others escape any measurement discipline due to their inherent characteristics and the absence of historical precedent [31]. Such incomplete reporting can relate especially to event detection.

It is also natural to suspect that the influence of such factors varies over time, possibly due to

1. Different dates of first implementation, occurrence or introduction
2. Varying speed of implementation of preventative operational risk measures.
3. Varying effectiveness between organizations regarding differing risk management procedures.

In addition, it has been stated that internal data on operational risks in the insurance industry are often limited and potentially biased. This is not only due to internal collection problems but also because large operational losses do not happen very often and take time to be fully appreciated [49].

Therefore, in much the same way as for earthquake events, data gathered in such a dynamic environment can readily be hypothesized to suffer some form of incompleteness. Such incompleteness can manifest in the form of a level of completeness, that may very well be time-varying.

6.3 Application of actuarial and statistical techniques

As stated in the introductory literature study, it is readily seen that a number of factors contribute to the creation of a rather synthetic divide : a subset of smaller valued events in earthquake and operational risk data that are represented by distorted observations, and a subset of larger valued events which are observationally unaffected.

In order to assist in the proper analysis of data and subsequent interpretation of results, the range affected by synthetically induced changes must be identified. This must be accomplished in order to detect external influences and to identify a distributional threshold that indicates either a cut-off, which eliminates the affected events from consideration, or the need to incorporate alternative models.

Table 6.2 summarizes the prevailing threshold estimation techniques and some of their prominent characteristics. The newly derived MITC scheme has also been included. These characteristics include the basis (statistical or graphical) and whether a distributional assumption is made for the data in some way (parametric or non-parametric). A method is described as having a statistical basis if the notion of probability and likelihood is utilized. A statistical basis is where the implementation of the method draws upon statistical techniques. Furthermore, a method is described as having a graphical basis when the threshold estimation does not draw as sharply on statistical techniques, but more on the characteristics of the data. Motivation for the respective methods will be outlined when each method is described in Chapter 9.

Threshold estimation method and abbreviation		Characteristics of method			
		Statistical basis	Graphical basis	Parametric	Non-Parametric
Goodness of fit	(GOF)	✓		✓	
Maximum curvature	(MAXC)		✓		✓
m_c by b-value stability	(bVS)		✓	✓	
Entire-magnitude-range	(EMR)	✓		✓	
Median based assessment of the segment slope	(MBASS)	✓			✓
Moment incorporating threshold computation	(MITC)	✓		✓	

Table 6.2: *Tabulation of threshold estimation methods and prominent characteristics.*

As will be seen when fully discussing these methods, they can fairly easily be extended to differing forms of the event magnitude distribution. A benefit of this generalization is the use in threshold estimation in operational risk data.

7 Considerations and limitations of study

7.1 Data

Historic instrumental data may most directly reflect the nature of event-generating stress. As briefly touched on in Section 6.1.1, the usable part of this earthquake data is unfortunately extremely short, and considerable caution should be taken to assess the completeness (level of $\sim$) of all datasets used [36]. For this particular reason it has been decided to first test all threshold estimation methods using synthetically generated data. Synthetically generated data will have some obvious advantages over real seismicity data:

1. The magnitude at which complete reporting starts (m_c) can be specified beforehand. Such specification aids in the objective comparison of the effectiveness of estimation methods.
2. The shape of the incomplete portion of the magnitude range can be varied to portray differing physically plausible situations, whilst retaining the position of m_c.
3. The number of events generated or, alternatively, the time period over which the synthetic dataset spans can be specified at outset. As such, statistical credibility of the analysis can be guaranteed. For the current study, the number of events will be specified.

However, it should be mentioned that some of these points have been scrutinized by authors. These parties argue that the use of synthetically created datasets is not relevant since doubts exist on the theoretical shape of frequency - magnitude distributions (FMD's) [3].

It is felt that such critique indeed motivates the use of synthetically generated FMDs, seeing as they provide an excellent opportunity to incorporate differing forms of the detection probability. Subsequent testing of threshold estimation techniques can therefore also be rated according to their robustness when dealing with these various forms of the detection probabilities. Furthermore, such incorporation can aid in the understanding of relations between possible observed patterns of earthquake magnitude and the ensuing results of statistical analyses.

7.2 Critical assumptions

A main objective of this investigation is to accurately determine the magnitude of completeness (m_c) in earthquake data. As such, assumptions must be made regarding other facets of the process, which will not form part of this investigation. A brief discussion of these assumptions follows.

1. **Events as a system with characteristics of self-organized criticality (SOC)**
   A self-organized critical process has the inherent feature that the process, without external influence, naturally moves to a critical state. This serves as a driving agent to provide complexity

in systems [4]. The notion of earthquake events as a self-organized critical process establishes a solid framework for self-similarity and magnitude scaling of said earthquake events when expressed in terms of the seismic moment (M_0) or energy. Such self-similar phenomena can consistently be modelled by a suitable power law or Pareto distribution. Pareto distributions and power law distributions furthermore possess the property of scale invariance since $f_{M_0}(\lambda m) = \lambda^{\alpha} f_{M_0}(m)$ for some α. This in turn implies self-similarity of the distribution function of M_0. Consequently the transformed random variable $M = \log_{10} M_0$ follows an exponential distribution, which is the case when magnitude is expressed in terms of the Richter-scale. Main [39, 40] discusses this topic and list further references. A similar assumption is made regarding operational risk data. This can be seen when considering the established fit of Generalized Pareto and Extreme value distributions to loss data.

2. $\boldsymbol{m_{min}}$ (Smallest event magnitude)
   In the literature on earthquakes some authors have discussed the departure from self-similarity at low magnitude ranges [2, 18, 38, 52, 55]. When considering data described by this departure from self-similarity, it might be seen that a Log-Normal distribution offers a reasonable fit to lower magnitude events, whereas a Pareto or Extreme Value fit is preferable for higher magnitude events. This can be seen to be the case in operational risk studies. In this study the assumption is made that the lower magnitude of generated events do not depart from self-similarity and therefore no distortion of lower magnitude data arises as a result of this departure from self-similarity.

   Considering the detected distribution of events, versus the generated distribution of events, it is assumed that the only reason for a departure from the law governing generated distributions is the act of not detecting and registering of events.

   Another point is that studies have generally truncated the magnitude distribution from the left due to distortion of data found in the incomplete portion. However, since the motivation for this investigation is the estimation of the magnitude of completeness, the chosen truncation point will be of a smaller magnitude. This is done so as to include the incomplete portion of data. This point of truncation will be referred to as m_{min}.

3. $\boldsymbol{m_{max}}$ (Largest event magnitude)
   In earthquake literature, a fundamental aspect of the higher magnitude range has been debated, namely whether the Gutenberg-Richter law can continue indefinitely. This has been argued due to the existence of the physical requirement of finite-energy flux, or equivalently a finite-seismic moment-release rate [39]. Due to these concerns a number of authors [14, 15, 26, 32, 43, 53, 58] have discussed an upper bound for the magnitude range, namely m_{max}. In this study the considered magnitude range will also be bounded from above by m_{max}. Furthermore, due to the

finite risk exposure of institutions, the existence assumption of m_{max} carries over very reasonably to operational loss data.

4. **Certainty of m_{min} and m_{max}**

   As this study is primarily focused on the estimation of the magnitude of completeness, the assumption is made that these points are not subject to change or uncertainty and also known in advance. This assumption can however, be relaxed through, e.g., inclusion of techniques to estimate m_{max} such as those proposed by Kijko in a 2004 article [32].

5. **Elimination of false event detections**

   Suppose G is the set of all events in a given space-time volume and D is the set of events that have been observed. The assumption is made that $D \subseteq G$, such that false detections are an impossibility, i.e. $D \cap G^C = \emptyset$. If significant contrary evidence accumulates, revision of this assumption will be required. However, this falls outside the scope of the current investigation.

6. **Observational error**

   It is assumed that reported magnitudes are the true magnitudes of earthquake events and therefore there is no need to make allowance for observational error. However, this is an assumption that can be easily relaxed through defining of a new random variable M_A. Here M_A represents the apparent magnitude which takes account of stochastic observation error, e.g. $M_A = M_G + e$, where e follows some distribution of errors.

7. **Independence of frequency and severity**

   The assumption is made that the process responsible for earthquake occurrences and the resulting size of the earthquake is independent. This simplifying assumption is generally accepted in earthquake studies. Furthermore, in studies of operational risk modelling [49] the same assumption has been made.

8 Generating severity data

In this chapter attention will be given to the complete as well as the incomplete portion of the event size distribution. The end goal is to be able to produce realistic data that includes events from the entire magnitude range. Initially, the focus will be on earthquake magnitude distributions. Hereafter the event size distribution relating to operational risk can be considered.

8.1 Complete data

In terms of earthquake events, it has been widely accepted that the Gutenberg-Richter [22] distribution can be used to model the distribution of earthquake events. In the absence of any destruction process acting on the data, data generated from this distribution will represent the complete portion of the catalogue. This translates into the scenario where any event that has been generated, has a probability of 1 to be observed, i.e. recorded.

The Gutenberg-Richter probability distribution takes the following form.

$$f_M(m) = \begin{cases} \beta e^{-\beta m} & \text{if } m > 0 \\ 0 & \text{otherwise} \end{cases} \tag{8.1}$$

where $\beta = b \ln 10 > 0$ is the original b-value in the Gutenberg-Richter frequency magnitude relation in (2.1). In deriving (8.1) from (2.1) it can be seen that the a - value has been cancelled out by the normalizing constant.

Upon restriction of the magnitude range to the closed interval $[m_{min}, m_{max}]$ and thereby modifying the support of the random variable, the probability density takes on the form

$$f_M(m) = \begin{cases} \frac{\beta \exp(-\beta m)}{\exp(-\beta m_{min}) - \exp(-\beta m_{max})} & \text{if } m_{min} \leq m \leq m_{max} \\ 0 & \text{otherwise} \end{cases} \tag{8.2}$$

By utilizing (B.30) in the appendix (p. 136) the expected magnitude can be expressed as

$$E[M^r]\Big|_{r=1} = E[M] = \frac{1}{\beta} + \frac{m_{max} e^{-\beta m_{max}} - m_{min} e^{-\beta m_{min}}}{e^{-\beta m_{max}} - e^{-\beta m_{min}}} \tag{8.3}$$

From the density in (8.2) the cumulative probability distribution, $F_M(m)$, and quantile function, $Q_M(p)$, can respectfully be stated as

$$F_M(m) = \begin{cases} 0 & \text{if } m \leq m_{min} \\ \frac{1 - \exp(-\beta(m - m_{min}))}{1 - \exp(-\beta(m_{max} - m_{min}))} & \text{if } m_{min} < m \leq m_{max} \\ 1 & \text{if } m > m_{max} \end{cases} \tag{8.4}$$

$$Q_M(p) = m_{min} - \frac{1}{\beta}\ln\left(1 - p\left(1 - e^{-\beta(m_{max}-m_{min})}\right)\right) \text{ for } 0 \leq p \leq 1 \tag{8.5}$$

In order to obtain random samples from the distribution in (8.2), the inverse probability integral transform method can be applied by using the quantile function (8.5).

8.2 Data simulation

8.2.1 The detected magnitude distribution

In this section the approach of producing earthquake catalogues and operational risk data is formalized. The synthesizing of data will be undertaken by generating random samples based on the Gutenberg-Richter probability distribution to represent earthquake risk data. However, this assumes complete reporting of all events. In order for us to generate samples that are subject to the loss of data, a probability distribution of detected events must be derived. The newly derived distribution will be described in terms of priorly known distributions. As a starting point let G denote the stochastic set of all events generated in a given space-time volume, i.e. a set where the number of elements is a random variable and the magnitude of each such event is also stochastically generated. Now, let D denote the set of all events that have been detected. As stated in Section 7.2, the assumption is made that D is a subset of G. However, the equality of the sets D and G is not disallowed. In keeping with the defined sets, M_G will denote the random variable that represents the generated magnitude of an event, whereas M_D will denote the random variable associated with the magnitude of the detected event.

From these definitions and assumptions it can be readily seen that the occurrence of detecting an event in the magnitude range $[m, m+h]$, i.e. $\{m \leq M_D \leq m+h\}$, is equivalent to having an event generated with magnitude in the range $[m, m+h]$ and subsequently having the generated event being observed, i.e. $\{(m \leq M_G \leq m+h) \cap (M_G \in D)\}$.

Using the limit definition of a probability density function, the density of the detected magnitude distribution can therefore be expressed as:

$$\begin{aligned}
f_{M_D}(m) &= \lim_{h\to 0} \frac{P[m \leq M_D < m+h]}{h} \\
&= \lim_{h\to 0} \frac{P[\{M_G \in D\} \cap \{m \leq M_G < m+h\}]}{h} \\
&= \lim_{h\to 0} \frac{P[m \leq M_G < m+h]P[M_G \in D|\{m \leq M_G < m+h\}]}{h} \\
&= \lim_{h\to 0} \frac{P[m \leq M_G < m+h]}{h} \lim_{h\to 0} P[M_G \in D|\{m \leq M_G < m+h\}] \qquad (8.6) \\
&= f_{M_G}(m)P[M_G \in D|M_G = m] \qquad (8.7)
\end{aligned}$$

As can be seen $P[M_G \in D | M_G = m]$ represents the conditional probability of an event being detected, given that the event has been generated and $f_{M_G}(m)$ represents the density function of an event being generated of magnitude m. Throughout this study $F_D(m)$ will exclusively denote the function used to model this probability, i.e. $F_D(m) = P[M_G \in D | M_G = m]$.

The relation in (8.7) states that all events, regardless of magnitude, have a non-zero probability of not being detected. This is termed the soft magnitude of completeness or more generally, an implicit or soft detection threshold. However, in the traditional sense of the discussed problem of complete reporting, only events falling in a specific magnitude have a non-zero probability of not being detected, which will be termed a sharp, or explicit, threshold of detection.

8.2.1.1 Soft detection threshold

In this study, this form of the data destruction process will only be considered for the modelling of earthquake data and not for the case of operational risk data. In this scenario, the magnitude of completeness is not defined as a specific point in the magnitude range, but the data-destruction effect is included implicitly. Artificially, this can also be achieved by equating the point m_c to m_{max}, yielding the result that all events in the magnitude range $[m_{min}, m_{max}]$ have a non-zero probability of not being detected. After gauging the core effectiveness of the threshold estimation methods, the performance and robustness of threshold estimation will be tested under this variation of the detection distribution. The sensitivities of estimation methods are checked and the number of events lost as a result of the data destruction process computed. This auxiliary study will demonstrate the decisiveness of the methods to identify a level of completeness that is regarded as complete by the specific method.

8.2.1.2 Sharp detection threshold

The magnitude of completeness is pre-specified and will be incorporated as a particular point in the magnitude range. This has the effect that only events in the magnitude range $[m_{min}, m_c)$ have a non-zero probability of not being detected. Subsequently, events larger than m_c up to m_{max}, are detected with probability 1.

Although not expressly stated, some authors of earthquake literature [42, 60] have adopted a definition of a detection probability that can be expressed in the following manner :

$$\mathrm{P}\left[\{\text{Detection of seismic event of size } M = m\}\right] = \begin{cases} 0 & \text{if } m \leq m_{min} \\ g(m) & \text{if } m_{min} < m \leq m_c \\ 1 & \text{if } m > m_c \end{cases} \tag{8.8}$$

with $g(m)$ some function dependent on magnitude M. This function is non-decreasing over the range $[m_{min}, m_{max}]$. It has also not been explicitly stated, but in a number of cases (e.g. Woessner and Wiemer

[60]) the function $g(m)$ does not assume the value of 1 at m_c. This is worrying, since the interval to which $g(m)$ is applied, as per the detection probability definition, is closed at the point m_c. Subsequently, a discontinuity can exist at this point of the detected magnitude density. To avoid such a manifestation in this study, $g(m)$ will take the form of a scaled function where the value of 1 is attained at m_c.

As shown in (A.17) in the appendix (p. 132) the detected magnitude distribution, for a sharp detection threshold, can be expressed in the following piecewise manner :

$$f_{M_D}(m) = \begin{cases} f_{M_G}(m) \cdot P[M_G \in D | \{M_G = m\} \cap \{M_G < m_c\}] & \text{if } m \in [m_{min}, m_c) \\ f_{M_G}(m) & \text{if } m \in [m_c, m_{max}] \\ 0 & \text{otherwise} \end{cases} \tag{8.9}$$

In keeping with the need for the detection probability to attain the value of 1 at the magnitude of completeness, the CDFs used to model the probability will be further normalized by rather considering the CDF of the truncated random variables. The point of truncation will be m_c, which corresponds to

$$P[M_G \in D | \{M_G = m\} \cap \{M_G < m_c\}] = \frac{F_D(m)}{F_D(m_c)} \tag{8.10}$$

$$= F_D^T(m) \tag{8.11}$$

where the superscript T is used to indicate that the distribution is truncated from the right.

8.2.2 Functional forms of detection probability

For the respective cases of a soft- and sharp detection threshold, the following probabilities must be considered : $P[M_G \in D | \{M_G = m\}]$ and $P[M_G \in D | \{M_G = m\} \cap \{M_G < m_c\}]$.

Some authors have already confirmed the Cumulative Normal distribution function as a suitable choice for a detection probability [42]. Woessner and Wiemer [60] have directly used this function to model the incomplete portion. In the same article, Woessner and Wiemer, indicate three other possibilities of directly modelling the incomplete portion of earthquake catalogues. These include the cumulative distribution functions of random variables that are either Exponentially or Log-Normally distributed as well as an Exponential decay. It is further stated that the three CDFs produce competitive likelihood scores, but the Normal CDF generally best fits the data.

Since the inclusion of a detection probability has not been investigated in a manner comparable to this study, no further direct motivation for the choice of detection probabilities can be drawn from. As an exploratory step, it is aimed to create a full and diverse spectrum of detection probability progressions. It is envisioned that this can be realized when considering the cumulative distribution functions of random variables that are distributed according to the following distributions :

1. Normal
2. Logistic
3. Log-Normal
4. Pareto type II

The three last mentioned distribution functions can be regarded as extensions of the detection probability theory and offer alternative forms that in certain circumstances may provide a better fit. This remains to be investigated and the fit of any specific detection distribution quantified.

Table 8.1 gives a summary of the cumulative distribution functions that will be used to model the detection probability, together with some of the prominent distributional characteristics.

CDF	Distribution characteristics				
	Symmetric	Non-Symmetric	Heavy-tailed	Light-tailed	Free parameters
Normal	✓			✓	2
Logistic	✓		✓		2
Log-Normal		✓	✓		2
Pareto type II		✓	✓		3

Table 8.1: *Tabulation of considered forms of the detection probability along with prominent features of respective distributions.*

Reference will be made throughout this study to the generic concept of a "*cumulative distribution function used to model the detection probability*". Due to the frequency of such reference, the aforementioned concept can be used interchangeably with the shortened reference of a "*detection distribution*".

What follows in Sections 8.2.3 up to 8.2.6 are derivations of the detected magnitude distribution ($f_{M_D}(m)$) for various forms of the detection distribution ($F_D(m)$), as described above. Thereafter, Section 8.2.7 will be dedicated to examining the sensitivities of the derived detected magnitude distributions to changes in the parameters of the detection distributions.

8.2.3 Detection probability modelled by cumulative Normal distribution

In the following two sections the detection probability takes the form of a cumulative Normal distribution function $\Phi(z_m)$, where

$$\Phi(z_m)=\int_{-\infty}^{z_m} \frac{1}{\sigma}\phi(t)\,dt \text{ where } z_m = \frac{m-\mu}{\sigma} \text{ and} \tag{8.12}$$

$$\phi(t)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2} \text{ for } t \in \mathbb{R} \tag{8.13}$$

8.2.3.1 Soft detection threshold

The probability of observing a seismic event of size m will be represented by the value from the cumulative Normal distribution, i.e. $P[M_G \in D | M_G = m] = \Phi(z_m)$. From this definition, it can be seen that the detected magnitude distribution follows to be

$$f_{M_D}(m) \propto f_{M_G}(m) \cdot \Phi(z_m) \qquad \text{for } m \in [m_{min}, m_{max}] \tag{8.14}$$

Which implies

$$f_{M_D}(m) = C_{Norm} \cdot f_{M_G}(m) \cdot \Phi(z_m) \qquad \text{for } m \in [m_{min}, m_{max}] \text{ and some } C_{Norm} \in \mathbb{R} \tag{8.15}$$

Cumulative distribution function and normalizing constant

The density of the detected magnitude must satisfy the relation

$$\int_{m_{min}}^{m_{max}} f_{M_D}(m) dm = C_{Norm} \int_{m_{min}}^{m_{max}} f_{M_G}(m) \Phi(z_m)\, dm = 1 \tag{8.16}$$

with C_{Norm} a suitable normalizing constant. From (B.35) in the appendix (p. 138) it can be seen that

$$(C_{Norm})^{-1} = \Phi(z_{m_{max}}) - c_1 \left(\Phi(z_{m_{max}}) - \Phi(z_{m_{min}}) - c_2 \left(\Phi(z^*_{m_{max}}) - \Phi(z^*_{m_{min}})\right)\right) \tag{8.17}$$

since $F_{M_G}(m_{min}) = 0$ and $F_{M_G}(m_{max}) = 1$

Where

$$c_1 = \left(1 - \exp\left(-\beta\left(m_{max} - m_{min}\right)\right)\right)^{-1} \tag{8.18}$$

$$c_2 = \exp\left(-\beta\left(\mu - m_{min} - \frac{1}{2}\sigma^2\beta\right)\right) \tag{8.19}$$

$$z^*_m = \frac{m - (\mu - \sigma^2\beta)}{\sigma} \tag{8.20}$$

Therefore the detected magnitude for $m \in [m_{min}, m_{max}]$, has the following distribution function :

$$F_{M_D}(m) = C_{Norm} \cdot \left(F_{M_G}(m)\Phi(z_m) - c_1\left(\Phi(z_m) - \Phi(z_{m_{min}}) - c_2\left(\Phi(z^*_m) - \Phi(z^*_{m_{min}})\right)\right)\right) \tag{8.21}$$

$$= C_{Norm} \cdot \left(F_{M_G}(m)\Phi(z_m) - c_1\left(\Phi(z_m) - c_2 \cdot \Phi(z^*_m)\right) + c_3\right) \tag{8.22}$$

where $c_3 = c_1\left(\Phi(z_{m_{min}}) - c_2 \cdot \Phi(z^*_{m_{min}})\right)$.

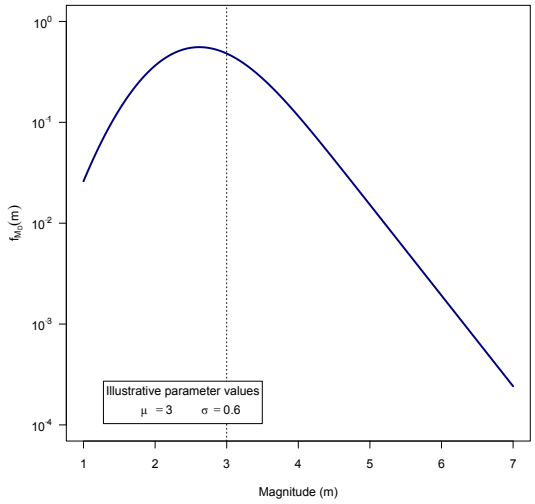
Quantile function

From (8.22), the implicit quantile function $Q_{M_D}(m) = p$ follows readily and can be numerically solved by defining :

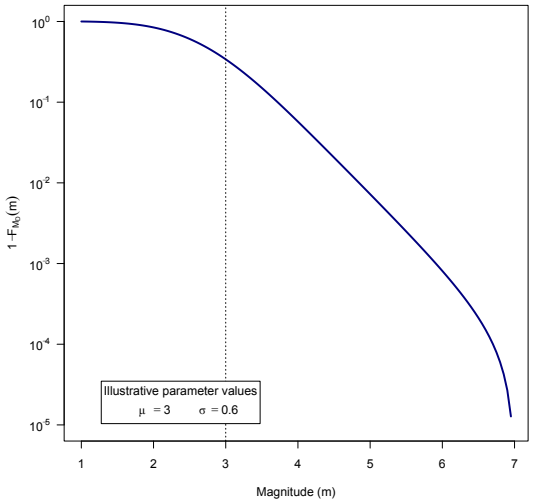
$$
\begin{aligned}
g(m) &:= \frac{p}{C_{Norm}} - F_{M_G}(m) \cdot \Phi(z_m) + c_1 \left(\Phi(z_m) - c_2 \cdot \Phi(z_m^*)\right) - c_3 \\
\Rightarrow g'(m) &= -f_{M_G}(m) \cdot \Phi(z_m) - F_{M_G}(m) \cdot \frac{1}{\sigma}\phi(z_m) + \frac{c_1}{\sigma}\left(\phi(z_m) - c_2 \cdot \phi(z_m^*)\right)
\end{aligned}
\tag{8.23}
$$

The required quantile can be obtained by utilizing the Newton-Raphson iterative scheme :

$$
m_{n+1} = m_n - \frac{g(m_n)}{g'(m_n)} \tag{8.24}
$$



(a) *Density function of detected magnitude distribution. Illustrative values used together with* $\mu = 3$ *and* $\sigma = 0.6$



(b) *Complementary distribution function of detected magnitude. Illustrative values used together with* $\mu = 3$ *and* $\sigma = 0.6$

Figure 8.1: *Illustration of density function (8.15) and complementary distribution function of detected magnitude distribution with detection probability, over entire magnitude range, modelled by the cumulative Normal distribution function. Illustrative values for parameters include :* $m_{min} = 1$, $m_{max} = 7$, *b-Value* $= 0.9$. *Reference lines are included at* μ

8.2.3.2 Sharp detection threshold

An explicit magnitude break can be produced by having the event detection probability modelled by the cumulative distribution function of a truncated Normal random variable in the following manner : $P[M_G \in D | \{M_G = m\} \cap \{M_G < m_c\}] = \dfrac{\Phi(z_m)}{\Phi(z_{m_c})} = \Phi^T(z_m)$, which leads to

$$
f_{M_D}(m) \propto \begin{cases} f_{M_G}(m) \cdot \Phi^T(z_m) & \text{if } m_{min} \leq m < m_c \\ f_{M_G}(m) & \text{if } m_c \leq m \leq m_{max} \\ 0 & \text{otherwise} \end{cases} \tag{8.25}
$$

The superscript T is used to indicate that the distribution is truncated from the right. For the specific

case of this investigation the point of truncation is the magnitude of completeness. As can be seen from Section 8.2.1.2, the motivation for this definition is to ensure that the detected magnitude distribution continues as the traditional Gutenberg-Richter distribution at the magnitude of completeness.

Normalizing constant

To ensure validity as a probability density function the normalizing constant C_{Norm} must be computed to satisfy the relation :

$$1 = C_{Norm}\left(\int_{m_{min}}^{m_c} f_{M_G}(m)\cdot\Phi^T(z_m)\,dm + \int_{m_c}^{m_{max}} f_{M_G}(m)\,dm\right) \tag{8.26}$$

Therefore,

$$\begin{aligned}(C_{Norm})^{-1} &= \frac{1}{\Phi(z_{m_c})}\left[F_{M_G}(t)\Phi(z_t) - c_1[\Phi(z_t) - c_2\Phi(z_t^*)]\right]_{m_{min}}^{m_c} + F_{M_G}(m_{max}) - F_{M_G}(m_c) \qquad \text{from (8.17)}\\ &= \frac{1}{\Phi(z_{m_c})}\left(F_{M_G}(m_c)\Phi(z_{m_c}) - c_1[\Phi(z_{m_c}) - c_2\Phi(z_{m_c}^*)] + c_1[\Phi(z_{m_{min}}) - c_2\Phi(z_{m_{min}}^*)]\right) + 1 - F_{M_G}(m_c)\\ &\qquad\qquad \text{since } F_{M_G}(m_{min}) = 0 \text{ and } F_{M_G}(m_{max}) = 1\\ &= 1 - \frac{c_1}{\Phi(z_{m_c})}\cdot\left(\Phi(z_{m_c}) - \Phi(z_{m_{min}}) - c_2\cdot(\Phi(z_{m_c}^*) - \Phi(z_{m_{min}}^*))\right)\end{aligned}$$

Cumulative distribution function

The distribution of the detected magnitude distribution follows readily :

1. for $m \in [m_{min}, m_c)$

$$\begin{aligned}F_{M_D}(m) &= \frac{C_{Norm}}{\Phi(z_{m_c})}\cdot\int_{m_{min}}^{m} f_{M_G}(m)\Phi(z_t)dt\\ &= \frac{C_{Norm}}{\Phi(z_{m_c})}\cdot(F_{M_G}(m)\cdot\Phi(z_m) - c_1\cdot(\Phi(z_m) - c_2\cdot\Phi(z_m^*)) + c_3) \qquad \text{from (8.22)}\end{aligned} \tag{8.27}$$

2. for $m \in [m_c, m_{max}]$

$$\begin{aligned}F_{M_D}(m) &= C_{Norm}\cdot\left(\frac{1}{\Phi(z_{m_c})}\int_{m_{min}}^{m_c} f_{M_G}(m)\Phi(z_t)\,dt + \int_{m_c}^{m} f_{M_G}(m)\,dt\right)\\ &= C_{Norm}\cdot\left(F_{M_G}(m) - \frac{1}{\Phi(z_{m_c})}\left(c_1\cdot(\Phi(z_{m_c}) - c_2\cdot\Phi(z_m^*)) + c_3\right)\right) \qquad \text{from (8.27)}\end{aligned} \tag{8.28}$$

Quantile function

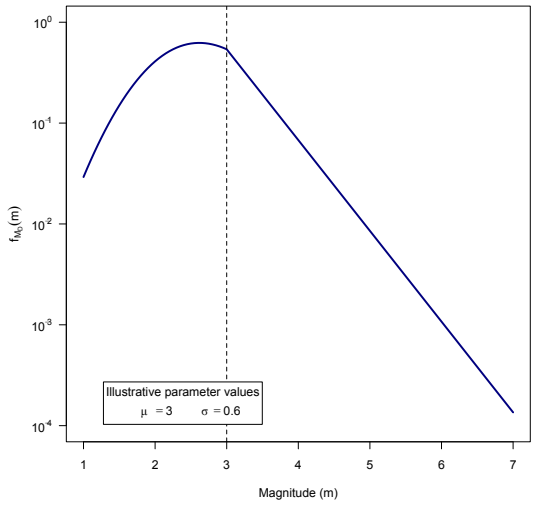
By defining

$$g(m) := p\cdot\frac{\Phi(z_{m_c})}{C_{Norm}} - F_{M_G}(m)\cdot\Phi(z_m) + c_1\left(\Phi(z_m) - c_2\cdot\Phi(z_m^*)\right) - c_3 \tag{8.29}$$

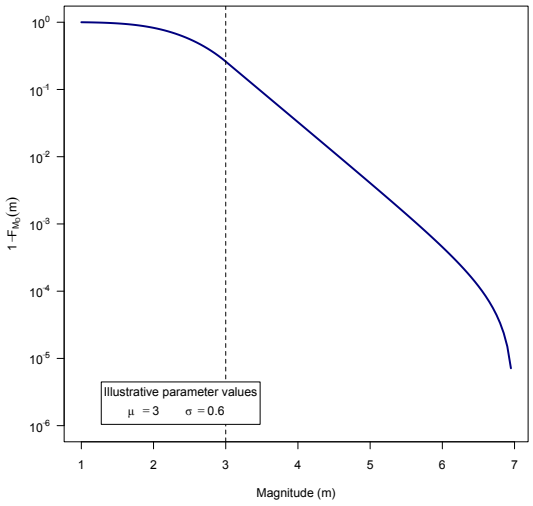
The iterative scheme presented in (8.24) can be used to solve for values of the quantile function $Q_{M_D}(p) = p$ where $p < F_{M_D}(m_c)$. Values of the quantile function where $p \geq F_{M_D}(m_c)$, i.e. $m \in [m_c, m_{max}]$, can be obtained by solving (8.28) for m :

$$\frac{p}{C_{Norm}} = F_{M_G}(m) - \frac{1}{\Phi(z_{m_c})}\left(c_1 \cdot \left(\Phi(z_{m_c}) - c_2\Phi(z^*_{m_c})\right) + c_3\right)$$

$$\Rightarrow \qquad m = Q_{M_G}\left(\frac{p}{C_{Norm}} + \frac{1}{\Phi(z_{m_c})}\left(c_1\left(\Phi(z_{m_c}) - c_2\Phi(z^*_{m_c})\right) + c_3\right)\right) \tag{8.30}$$



(a) *Density function of detected magnitude distribution. Illustrative values used together with* $\mu = 3$ *and* $\sigma = 0.6$



(b) *Complementary distribution function of detected magnitude. Illustrative values used together with* $\mu = 3$ *and* $\sigma = 0.6$

Figure 8.2: *Sensitivity illustrations of density function and complementary distribution function of detected magnitude with detection probability, below* m_c*, modelled by the cumulative Normal distribution function. Illustrative values for parameters include :* $m_{min} = 1$, $m_c = 3$, $m_{max} = 7$, *b-Value* $= 0.9$. *Reference lines are included at* m_c

8.2.4 Detection probability modelled by cumulative Logistic distribution

In the following two sections the probability of observing an event will be represented by the cumulative Logistic distribution. The cumulative Logistic distribution function can be expressed, for $m, \mu \in \mathbb{R}$ and $s > 0$ as follows:

$$F_D(m) = \frac{1}{1 + \exp(-\left(\frac{m-\mu}{s}\right))} \tag{8.31}$$

$$= \frac{1}{2}\left[1 + \tanh\left(\frac{m-\mu}{2s}\right)\right] \tag{8.32}$$

8.2.4.1 Sharp detection threshold

Breaking with the convention of the earlier section, the situation of a sharp detection threshold will first be considered. Where the event detection probability is modelled as the cumulative distribution function of a truncated Logistic random variable, the following can be written

$$P[M_G \in D | \{M_G = m\} \cap \{M_G < m_c\}] = F_D^T(m) \tag{8.33}$$

$$= \frac{F_D(m)}{F_D(m_c)} \tag{8.34}$$

Just as in (8.25), is is found that (where the superscript T indicates that the distribution is truncated to the right at m_c) :

$$f_{M_D}(m) \propto \begin{cases} f_{M_G}(m) \cdot F_D^T(m) & \text{if } m_{min} \leq m < m_c \\ f_{M_G}(m) & \text{if } m_c \leq m \leq m_{max} \\ 0 & \text{otherwise} \end{cases} \tag{8.35}$$

Normalizing constant

The normalizing constant C_{Norm} can be calculated from the following relation :

$$\begin{aligned} 1 =& C_{Norm}\left(\int_{m_{min}}^{m_c} f_{M_G}(m) \cdot F_D^T(m)\, dm + \int_{m_c}^{m_{max}} f_{M_G}(m)\, dm\right) \\ =& \frac{C_{Norm}}{c_4}\left[\frac{1}{2F_D(m_c)}\left(\int_{m_{min}}^{m_c} \beta e^{-\beta(m-m_{min})}\, dm + \int_{m_{min}}^{m_c} \beta e^{-\beta(m-m_{min})} \cdot \tanh\left(\frac{m-\mu}{2s}\right) dm\right)\right. \\ &\left. + \int_{m_c}^{m_{max}} \beta e^{-\beta(m-m_{min})}\, dm\right] \end{aligned}$$

where $c_4 = c_1^{-1} = 1 - e^{-\beta(m_{max}-m_{min})}$ is the reciprocal of the normalizing constant for $f_{M_G}(m)$.

$$\begin{aligned} =& \frac{C_{Norm}}{c_4}\left[\frac{1}{2F_D(m_c)}\left(1 - e^{-\beta(m_c-m_{min})} + \beta\int_{m_{min}}^{m_c} e^{-\beta(m-m_{min})} \cdot \tanh\left(\frac{m-\mu}{2s}\right) dm\right)\right. \\ &\left. + e^{-\beta(m_c-m_{min})} - e^{-\beta(m_{max}-m_{min})}\right] \end{aligned} \tag{8.36}$$

$$= \frac{C_{Norm}}{c_4}\left[e^{-\beta(m_c-m_{min})} - e^{-\beta(m_{max}-m_{min})} + \frac{1}{2F_D(m_c)}\left(1 - e^{-\beta(m_c-m_{min})}\right.\right.$$
$$+s\beta e^{-\beta(m_c-m_{min})}\left[\frac{\exp\left(\frac{m_c-\mu}{s}\right)}{1-\beta s}{}_2F_1\left(1-\beta s, 1; 2-\beta s; -\exp\left(\frac{m_c-\mu}{s}\right)\right)\right.$$
$$\left.+\frac{1}{\beta s}{}_2F_1\left(1, -\beta s; 1-\beta s; -\exp\left(\frac{m_c-\mu}{s}\right)\right)\right]$$
$$-s\beta\left[\frac{\exp\left(\frac{m_{min}-\mu}{s}\right)}{1-\beta s}{}_2F_1\left(1-\beta s, 1; 2-\beta s; -\exp\left(\frac{m_{min}-\mu}{s}\right)\right)\right.$$
$$\left.\left.\left.+\frac{1}{\beta s}{}_2F_1\left(1, -\beta s; 1-\beta s; -\exp\left(\frac{m_{min}-\mu}{s}\right)\right)\right]\right)\right]$$

by replacing the last remaining integral with the results of (B.40) as seen in the appendix (p.139).

Cumulative distribution function

1. For $m \in [m_{min}, m_c)$

$$F_{M_D}(m) = C_{Norm}\int_{m_{min}}^{m} f_{M_G}(t)\cdot F_D^T(t)\,dt$$
$$= \frac{C_{Norm}}{2F_D(m_c)}\left(\int_{m_{min}}^{m} f_{M_G}(t)\,dt + \int_{m_{min}}^{m} f_{M_G}(t)\cdot\tanh\left(\frac{t-\mu}{2s}\right)\,dt\right)$$

The distribution function follows readily by following the same method as that used when calculating the normalizing constant.

$$= \frac{C_{Norm}}{c_4}\frac{1}{2F_D(m_c)}\left[1 - e^{-\beta(m-m_{min})}\right.$$
$$+s\beta e^{-\beta(m-m_{min})}\left[\frac{\exp\left(\frac{m-\mu}{s}\right)}{1-\beta s}{}_2F_1\left(1-\beta s, 1; 2-\beta s; -\exp\left(\frac{m-\mu}{s}\right)\right)\right.$$
$$\left.+\frac{1}{\beta s}{}_2F_1\left(1, -\beta s; 1-\beta s; -\exp\left(\frac{m-\mu}{s}\right)\right)\right]$$
$$-s\beta\left[\frac{\exp\left(\frac{m_{min}-\mu}{s}\right)}{1-\beta s}{}_2F_1\left(1-\beta s, 1; 2-\beta s; -\exp\left(\frac{m_{min}-\mu}{s}\right)\right)\right.$$
$$\left.\left.+\frac{1}{\beta s}{}_2F_1\left(1, -\beta s; 1-\beta s; -\exp\left(\frac{m_{min}-\mu}{s}\right)\right)\right]\right] \tag{8.37}$$

2. For $m \in [m_c, m_{max}]$

$$F_{M_D}(m) = C_{Norm}\left(\int_{m_{min}}^{m_c} f_{M_G}(t)\cdot F_D^T(t)\,dt + \int_{m_c}^{m} f_{M_G}(t)\,dt\right)$$
$$= C_{Norm}\left(\int_{m_{min}}^{m_c} f_{M_G}(t)\cdot F_D^T(t)\,dt + F_{M_G}(m) - F_{M_G}(m_c)\right)$$

By utilizing the results in (8.37) the following is obtained.

$$= C_{Norm}\Bigg[F_{M_G}(m) - F_{M_G}(m_c) + \frac{1}{2\cdot c_4\cdot F_D(m_c)}\Bigg[1 - e^{-\beta(m_c-m_{min})}$$
$$+s\beta e^{-\beta(m_c-m_{min})}\Bigg[\frac{\exp\left(\frac{m_c-\mu}{s}\right)}{1-\beta s}\,{}_2F_1\left(1-\beta s, 1; 2-\beta s; -\exp\left(\frac{m_c-\mu}{s}\right)\right)$$
$$+\frac{1}{\beta s}\,{}_2F_1\left(1, -\beta s; 1-\beta s; -\exp\left(\frac{m_c-\mu}{s}\right)\right)\Bigg]$$
$$-s\beta\Bigg[\frac{\exp\left(\frac{m_{min}-\mu}{s}\right)}{1-\beta s}\,{}_2F_1\left(1-\beta s, 1; 2-\beta s; -\exp\left(\frac{m_{min}-\mu}{s}\right)\right)$$
$$+\frac{1}{\beta s}\,{}_2F_1\left(1, -\beta s; 1-\beta s; -\exp\left(\frac{m_{min}-\mu}{s}\right)\right)\Bigg]\Bigg]\Bigg] \tag{8.38}$$

Quantile function

For values of $p < F_{M_D}(m_c)$, i.e. $m < m_c$, the quantile function $Q_{M_D}(p)$ can be solved numerically by defining

$$g(m) := p\cdot\frac{2\cdot c_4\cdot F_D(m_c)}{C_{Norm}} - \Bigg[1 - e^{-\beta(m-m_{min})}$$
$$+s\beta e^{-\beta(m-m_{min})}\Bigg[\frac{\exp\left(\frac{m-\mu}{s}\right)}{1-\beta s}\,{}_2F_1\left(1-\beta s, 1; 2-\beta s; -\exp\left(\frac{m-\mu}{s}\right)\right)$$
$$+\frac{1}{\beta s}\,{}_2F_1\left(1, -\beta s; 1-\beta s; -\exp\left(\frac{m-\mu}{s}\right)\right)\Bigg]$$
$$-s\beta\Bigg[\frac{\exp\left(\frac{m_{min}-\mu}{s}\right)}{1-\beta s}\,{}_2F_1\left(1-\beta s, 1; 2-\beta s; -\exp\left(\frac{m_{min}-\mu}{s}\right)\right)$$
$$+\frac{1}{\beta s}\,{}_2F_1\left(1, -\beta s; 1-\beta s; -\exp\left(\frac{m_{min}-\mu}{s}\right)\right)\Bigg]\Bigg] \tag{8.39}$$

from where the Newton-Raphson iterative scheme can be implemented, as defined in (8.24).

For values of $p \geq F_{M_D}(m_c)$, i.e. $m \geq m_c$, the quantile function has analytical form that can be expressed as :

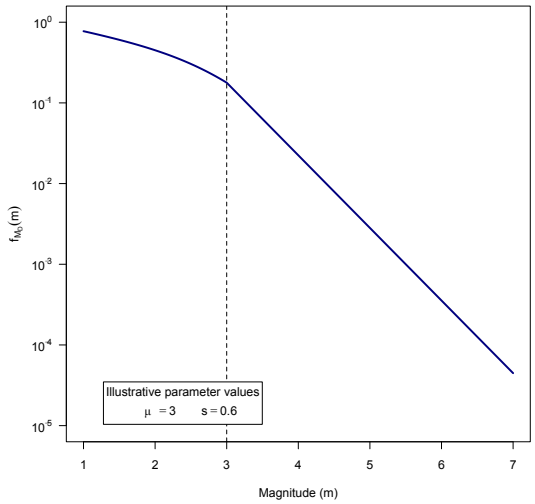
$$
\begin{aligned}
Q_{M_D}(p) = Q_{M_G}\Bigg(& \frac{p}{C_{Norm}} + F_{M_G}(m_c) - \frac{1}{2 \cdot c_4 \cdot F_D(m_c)} \Bigg[1 - e^{-\beta(m_c - m_{min})} \\
& + s\beta e^{-\beta(m_c - m_{min})} \Bigg[\frac{\exp\left(\frac{m_c - \mu}{s}\right)}{1 - \beta s} {}_2F_1\left(1 - \beta s, 1; 2 - \beta s; -\exp\left(\frac{m_c - \mu}{s}\right)\right) \\
& + \frac{1}{\beta s} {}_2F_1\left(1, -\beta s; 1 - \beta s; -\exp\left(\frac{m_c - \mu}{s}\right)\right) \Bigg] \\
& - s\beta \Bigg[\frac{\exp\left(\frac{m_{min} - \mu}{s}\right)}{1 - \beta s} {}_2F_1\left(1 - \beta s, 1; 2 - \beta s; -\exp\left(\frac{m_{min} - \mu}{s}\right)\right) \\
& + \frac{1}{\beta s} {}_2F_1\left(1, -\beta s; 1 - \beta s; -\exp\left(\frac{m_{min} - \mu}{s}\right)\right) \Bigg] \Bigg] \Bigg)
\end{aligned}
\tag{8.40}
$$

Remark

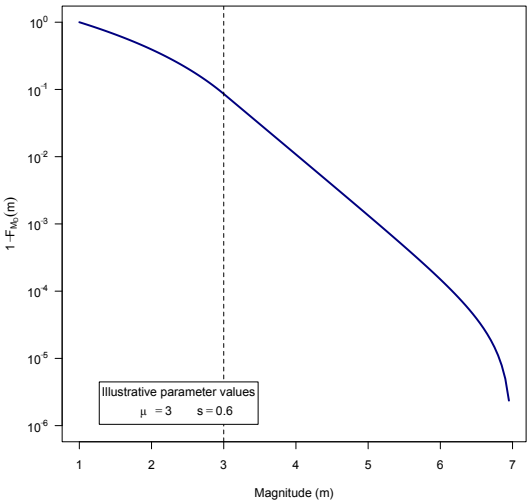
Due to the inclusion of the Gauss hyper-geometric function ${}_2F_1(a, b; c; z)$ in the expressions for the distribution and quantile functions derived in this section, an additional constraint is placed on the parameters of the distribution. This must be enforced to ensure convergence of the series that defines the function ${}_2F_1(a, b; c; z)$. These additional constraints arise from the fact that the z argument in the Gauss hyper-geometric function has the following constraint placed on it : $|z| < 1$. For the various Gauss hyper-geometric functions, it therefore follows that

$$
\begin{aligned}
& \left| -e^{\frac{m_{min} - \mu}{s}} \right| < 1 \qquad \text{and} \qquad \left| -e^{\frac{m_c - \mu}{s}} \right| < 1 \\
\Rightarrow \qquad & m_{min} < \mu \qquad \text{and} \qquad m_c < \mu
\end{aligned}
\tag{8.41}
$$

A third restriction derived from the above condition $|z| < 1$ is that $m < \mu$ for all values of m to be passed to the Gauss hyper-geometric function. However, since the equality $m_c < \mu$ must hold and computation of quantiles, through the use of the Gauss hyper-geometric function, only include $m \in [m_{min}, m_c)$, this condition is automatically satisfied and need not be considered any further for the case of a sharp threshold of completeness.



(a) *Density function of detected magnitude distribution. Illustrative values used together with* $\mu = 3$ *and* $s = 0.6$



(b) *Complementary distribution function of detected magnitude. Illustrative values used together with* $\mu = 3$ *and* $s = 0.6$

Figure 8.3: *Sensitivity illustrations of complementary distribution function of detected magnitude with detection probability, below* m_c*, modelled by the cumulative logistic distribution function. Illustrative values for parameters include :* $m_{min} = 1$ *,* $m_c = 3$*,* $m_{max} = 7$*, b-Value* $= 0.9$*. Reference lines are included at* m_c

8.2.4.2 Soft detection threshold

In this section, the case where the probability of detecting an event being modelled by the cumulative logistic distribution function will now be considered under the assumption that the detected magnitude distribution incorporates a soft detection threshold.

As mentioned at the end of the previous section, the inclusion of the Gauss hyper-geometric function ${}_2F_1(a, b; c; z)$ in the distribution and quantile function results in an additional constraint being placed on the parameters of the distribution. Considering the situation where the distribution function includes the Gauss hyper-geometric function for magnitudes in the interval $[m_{min}, m_c)$, as is the case for a sharp detection threshold, the restriction is such that $m_c < \mu$. Even with this restriction, resultant detected magnitude distributions could still accurately depict a distribution that has been influenced by a data destruction process, such as non-detection. However, when the Gauss hyper-geometric function is utilized in the detected magnitude distribution for magnitudes in the range $[m_{min}, m_{max}]$, as is the case when considering a soft detection threshold, the above restriction is replaced by $m_{max} < \mu$. This restriction produces detected magnitude distributions where the effect of data destruction can not be modelled as effectively as in the case of a sharp detection threshold. For this reason the modelling of a soft magnitude of completeness will not be computed utilizing Gauss hyper-geometric functions, but rather expressed in terms of integrals to be evaluated numerically.

Continuing, the detected magnitude distribution takes the form :

$$f_{M_D}(m) \propto \begin{cases} f_{M_G}(m) \cdot F_D(m) & \text{if } m \in [m_{min}, m_{max}] \\ 0 & \text{otherwise} \end{cases} \tag{8.42}$$

where the detection probability is modelled by (8.32).

Normalizing constant

The normalizing constant should be such that

$$1 = \int_{m_{min}}^{m_{max}} f_{M_D}(m)\, dm = C_{Norm} \int_{m_{min}}^{m_{max}} f_{M_G}(m) \cdot F_D(m)\, dm \tag{8.43}$$

Due to the functional form of the detected magnitude distribution (8.42) leads to the above integral that cannot be expressed in terms of elementary functions. Therefore numerical integration techniques are employed to further the investigation of this detected magnitude distribution.

Cumulative distribution function

$$F_{M_D}(m) = \int_{m_{min}}^{m} f_{M_D}(t)\, dt \tag{8.44}$$

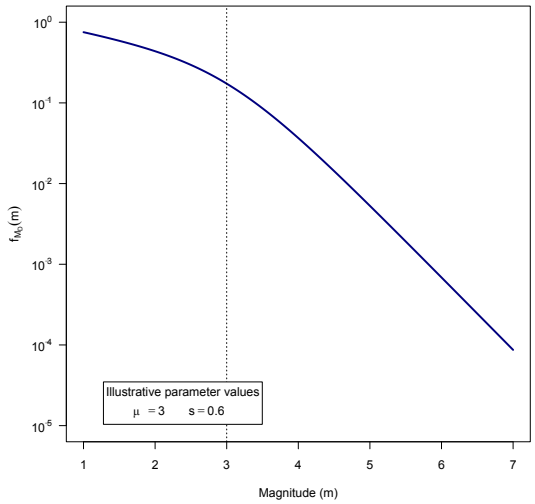
As previously stated, due to the form of the density $f_{M_D}(m)$, the cumulative distribution function must be evaluated numerically.

Quantile function

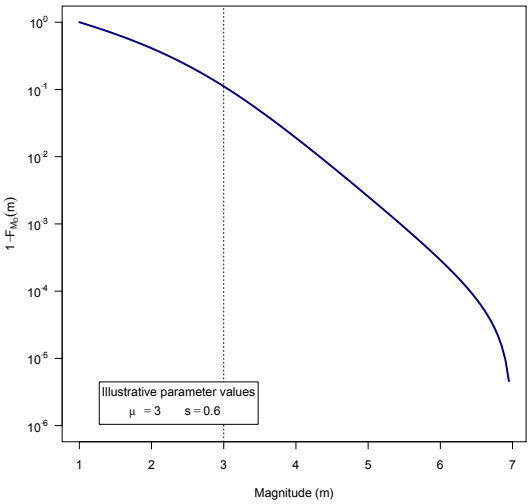
Due to the inability of the cumulative distribution function to be expressed in terms of elementary functions, values of the quantile function ($Q_{M_D}(p) = m$) must also be obtained numerically. The Newton-Raphson algorithm will be employed by defining

$$g(m) := \int_{m_{min}}^{m} f_{M_D}(t)\, dt - p \tag{8.45}$$

and solving for $g(m) = 0$.



(a) *Density function of detected magnitude distribution. Illustrative values used together with* $\mu = 3$ *and* $s = 0.6$.



(b) *Complementary distribution function of detected magnitude. Illustrative values used together with* $\mu = 3$ *and* $s = 0.6$

Figure 8.4: *Sensitivity illustrations of density function of detected magnitude distribution with detection probability, over the entire magnitude range, modelled by the cumulative logistic distribution function. Illustrative values for parameters include :* $m_{min} = 1$, $m_{max} = 7$ *and b-Value* $= 0.9$. *Reference lines are included at* μ.

8.2.5 Detection probability modelled by cumulative Pareto type II distribution

In the following two sections the probability of observing an event of size m will be represented by the value from the cumulative Pareto type II distribution. This distribution is characterized by the following continuous distribution and density functions :

$$F_D(m) = 1 - \left(1 + \frac{m-\mu}{\sigma}\right)^{-\alpha} \tag{8.46}$$

$$f_D(m) = \frac{\alpha}{\sigma}\left(1 + \frac{m-\mu}{\sigma}\right)^{-\alpha-1} \tag{8.47}$$

$$\text{where } m > \mu; \quad \mu \in \mathbb{R}; \quad \sigma > 0; \quad \alpha \in \mathbb{R}$$

8.2.5.1 Soft detection threshold

The probability of observing a seismic event of size m will be represented by the value from the cumulative Pareto type II distribution. The detection probability will be incorporated as follows:

$$f_{M_D}(m) \propto \begin{cases} f_{M_G}(m) \cdot F_D(m) & \text{for } m \in [m_{min}, m_{max}] \\ 0 & \text{otherwise} \end{cases} \tag{8.48}$$

Cumulative distribution function and normalizing constant

In order to obtain the cumulative distribution function it is worth noting that

$$\begin{aligned} &\int_a^b f_{M_G}(m) F_{M_D}(m)\, dm \\ &= \int_a^b f_{M_G}(m)\, dm - \int_a^b f_{M_G}(m)\left(1 + \frac{m-\mu}{\sigma}\right)^{-\alpha} dm \\ &= F_{M_G}(m)|_a^b - \frac{\beta e^{\beta m_{min}}}{1 - e^{-\beta(m_{max}-m_{min})}} \int_a^b e^{-\beta m}\left(1 + \frac{m-\mu}{\sigma}\right)^{-\alpha} dm \\ &= F_{M_G}(m)|_a^b - \frac{e^{-\beta(\mu - m_{min} - \sigma)}}{1 - e^{-\beta(m_{max}-m_{min})}} (\beta\sigma)^{\alpha}\, \gamma(1-\alpha; r(a), r(b)) \end{aligned}$$

from (B.43) in the appendix (p.140) where $r(t)$ is also defined. Furthermore, $\gamma(1-\alpha; r(a), r(b))$ is the incomplete gamma function with respect to a range. (8.49)

From the above, the normalizing constant (that must ensure that the density function integrates to 1 over the support of the random variable) follows to be :

$$C_{Norm}^{-1} = 1 - \frac{e^{-\beta(\mu-\sigma)}}{e^{-\beta m_{min}} - e^{-\beta m_{max}}} (\beta\sigma)^{\alpha} \gamma(1-\alpha; r(m_{min}), r(m_{max})) \tag{8.50}$$

The distribution function also follows from the above integral,

$$\begin{aligned} F_{M_D}(m) &= C_{Norm} \int_{m_{min}}^{m} f_{M_G}(t) F_D(t)\, dt \\ &= C_{Norm} \left(F_{M_G}(m) - \frac{e^{-\beta(\mu-\sigma)}}{e^{-\beta m_{min}} - e^{-\beta m_{max}}} (\beta\sigma)^{\alpha} \gamma(1-\alpha; r(m_{min}), r(m)) \right) \end{aligned} \tag{8.51}$$

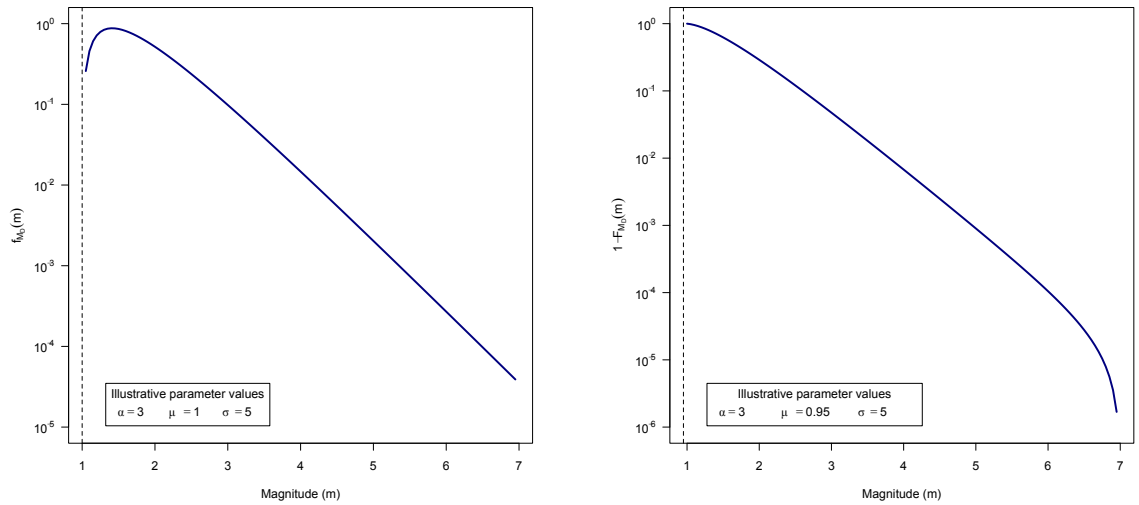
Quantile function

From (8.51), the implicit quantile function $Q_{M_D}(m) = p$ follows readily. Furthermore, it can be solved for values of p using numerical methods by defining :

$$\begin{aligned} g(m) &:= \frac{p}{C_{Norm}} - \int_{m_{min}}^{m} f_{M_G}(t) F_D(t)\, dt \qquad (8.52) \\ &:= \frac{p}{C_{Norm}} - F_{M_G}(m) + \frac{e^{-\beta(\mu-\sigma)}}{e^{-\beta m_{min}} - e^{-\beta m_{max}}} (\beta\sigma)^{\alpha} \gamma(1-\alpha; r(m_{min}), r(m)) \\ \Rightarrow g'(m) &= -f_{M_G}(m) F_{M_D}(m) \qquad \text{by (8.52)} \end{aligned} \tag{8.53}$$

from where the required quantile can be obtained by utilizing the Newton-Raphson iterative scheme :

$$m_{n+1} = m_n - \frac{g(m_n)}{g'(m_n)} \tag{8.54}$$



(a) *Density function of detected magnitude distribution. Illustrative values for parameters used.*

(b) *Complementary distribution function of detected magnitude. Illustrative values used for parameters.*

Figure 8.5: *Sensitivity illustrations of density function of detected magnitude distribution with detection probability, over entire magnitude range, modelled by the cumulative Generalized Pareto distribution function. Illustrative values for parameters include :* $m_{min} = 1$, $m_{max} = 7$, *b-Value* $= 0.9$, $\mu = m_{min}$, $\sigma = 5$ *and* $\alpha = 3$.

8.2.5.2 Sharp detection threshold

An explicit magnitude break in the detection probability can be introduced.

$$f_{M_D}(m) \propto \begin{cases} f_{M_G}(m) \cdot F_D^T(m) & \text{if } m \in [m_{min}, m_c) \\ f_{M_G}(m) & \text{if } m \in [m_c, m_{max}] \\ 0 & \text{otherwise} \end{cases} \tag{8.55}$$

Where the superscript T indicates that the detection probability distribution is truncated from the right at the point m_c.

Normalizing constant

The normalizing constant should be such that

$$1 = C_{Norm} \int_{m_{min}}^{m_{max}} f_{M_D}(m)\, dm \tag{8.56}$$

Therefore, it follows that

$$\begin{aligned} C_{Norm}^{-1} &= \int_{m_{min}}^{m_c} f_{M_G}(m) \cdot F_D^T(m)\, dm + \int_{m_c}^{m_{max}} f_{M_G}(m)\, dm \\ &= \frac{1}{F_D(m_c)} \int_{m_{min}}^{m_c} f_{M_G}(m) \cdot F_D(m)\, dm + F_{M_G}(m)\Big|_{m_c}^{m_{max}} \end{aligned}$$

It follows that after substituting the above integral for the expression in (8.49) and subsequent simplifications that

$$C_{Norm}^{-1} = 1 - F_{M_G}(m_c) + \frac{1}{F_D(m_c)} \left(F_{M_G}(m_c) - \frac{e^{-\beta(\mu-\sigma)}}{e^{-\beta m_{min}} - e^{-\beta m_{max}}} (\beta\sigma)^\alpha \gamma(1-\alpha; r(m_{min}), r(m_c)) \right) \tag{8.57}$$

Cumulative distribution function

1. For $m \in [m_{min}, m_c)$:

$$\begin{aligned} F_{M_D}(m) &= C_{Norm} \int_{m_{min}}^{m} f_{M_G}(t) F_D^T(t)\, dt = \frac{C_{Norm}}{F_D(m_c)} \int_{m_{min}}^{m_c} f_{M_G}(t) F_D(t)\, dt \\ &= \frac{C_{Norm}}{F_D(m_c)} \left(F_{M_G}(m) - \frac{e^{-\beta(\mu-\sigma)}}{e^{-\beta m_{min}} - e^{-\beta m_{max}}} (\beta\sigma)^\alpha \gamma(1-\alpha; r(m_{min}), r(m)) \right) \end{aligned}$$

from (8.49) and subsequent simplification. (8.58)

2. For $m \in [m_c, m_{max}]$:

$$
\begin{aligned}
F_{M_D}(m) &= C_{Norm}\left(\int_{m_{min}}^{m_c} f_{M_G}(t)F_D^T(t)\,dt + \int_{m_c}^{m} f_{M_G}(t)\,dt\right) \\
&= C_{Norm}\left(\frac{1}{F_D(m_c)}\int_{m_{min}}^{m_c} f_{M_G}(t)F_D(t)\,dt\right) + F_{M_G}(t)\Big|_{m_c}^{m} \\
&= C_{Norm}\Bigg(F_{M_G}(m) - F_{M_G}(m_c) + \frac{1}{F_D(m_c)}\Big(F_{M_G}(m_c) \\
&\qquad - \frac{e^{-\beta(\mu-\sigma)}}{e^{-\beta m_{min}} - e^{-\beta m_{max}}}(\beta\sigma)^{\alpha}\gamma(1-\alpha; r(m_{min}), r(m_c))\Big)\Bigg) \\
&\qquad\qquad \text{from (8.49) and subsequent simplification.}
\end{aligned}
\tag{8.59}
$$

Quantile function

1. For $m \in [m_{min}, m_c)$:

   From (8.58), the implicit quantile function $Q_{M_D}(m) = p$ follows readily and can be solved using numerical methods by defining :

$$
\begin{aligned}
g(m) &:= \frac{p \cdot F_D(m_c)}{C_{Norm}} - \int_{m_{min}}^{m} f_{M_G}(t)F_M D(t)\,dt \\
&:= \frac{p \cdot F_D(m_c)}{C_{Norm}} - F_{M_G}(m) + \frac{e^{-\beta(\mu-\sigma)}}{e^{-\beta m_{min}} - e^{-\beta m_{max}}}(\beta\sigma)^{\alpha}\gamma(1-\alpha; r(m_{min}), r(m)) \\
\Rightarrow g'(m) &= -f_{M_G}(m)F_D(m)
\end{aligned}
\tag{8.60}
$$

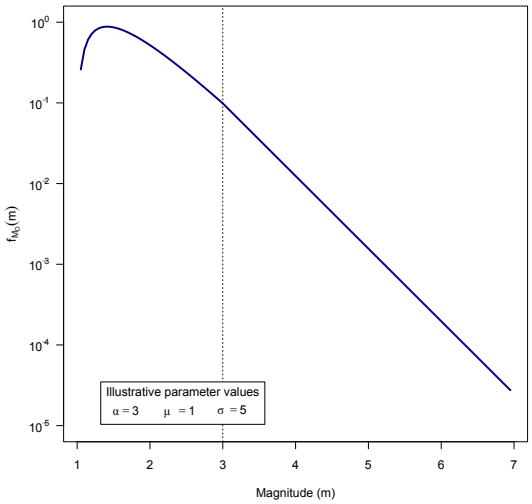
   The required quantile can be obtained by utilizing the Newton-Raphson iterative scheme :

$$
m_{n+1} = m_n - \frac{g(m_n)}{g'(m_n)} \tag{8.61}
$$

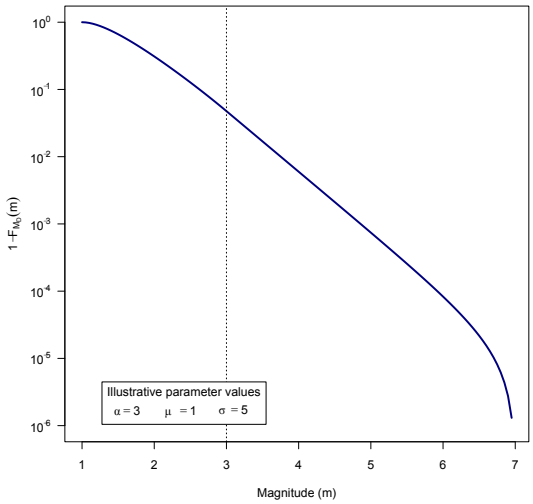
2. For $m \in [m_c, m_{max}]$:

   From (8.59), the quantile function $Q_{M_D}(m) = p$ can be expressed explicitly as

$$
\begin{aligned}
m &= Q_{M_D}(p) \\
&= Q_{M_G}\Bigg(\frac{p}{C_{Norm}} + F_{M_G}(m_c) \\
&\qquad - \frac{1}{F_D(m_c)}\left(F_{M_G}(m_c) - \frac{e^{-\beta(\mu-\sigma)}}{e^{-\beta m_{min}} - e^{-\beta m_{max}}}(\beta\sigma)^{\alpha}\gamma(1-\alpha; r(m_{min}), r(m_c))\right)\Bigg)
\end{aligned}
\tag{8.62}
$$



(a) *Density function of detected magnitude distribution. Illustrative values used for parameters.*



(b) *Complementary distribution function of detected magnitude. Illustrative values used for parameters.*

Figure 8.6: *Sensitivity illustrations of density function of detected magnitude distribution with detection probability, below m_c, modelled by the cumulative Generalized Pareto distribution function. Illustrative values for parameters include : $m_{min} = 1$, $m_c = 3$, $m_{max} = 7$, b-Value $= 0.9$, $\mu = m_{min}$, $\sigma = 5$ and $\alpha = 3$. Reference lines are included at m_c.*

8.2.6 Detection probability modelled by cumulative Log-Normal distribution

In the following two sections, the case to be considered is that where the probability of detecting an earthquake is modelled by the cumulative Log-Normal distribution function. For $m > 0$, $\mu \in \mathbb{R}$ and $\sigma > 0$ the Log-Normal distribution is characterized by the following continuous density and distribution functions

$$F_D(m) = \Phi\left(\frac{\ln m - \mu}{\sigma}\right) \tag{8.63}$$

$$f_D(m) = \frac{1}{m\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln m - \mu}{\sigma}\right)^2} \tag{8.64}$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of a standard Normal random variable. Even though the detected magnitude distribution will be include the cumulative distribution function of a Normal random variable, techniques employed such as in Section 8.2.3 do not yield similar solutions to the integrals. This is due to the specific form of the density function of the Log-Normal distribution. In order to bypass this problem, when necessary, techniques of numerical integration are implemented in the following two sections.

8.2.6.1 Soft detection threshold

When considering the case where the probability of detecting an event of magnitude m is modelled by the cumulative Log-Normal distribution function. The detected magnitude distribution can be expressed as follows :

$$f_{M_D}(m) \propto \begin{cases} f_{M_G}(m) \cdot F_D(m) & \text{if } m \in [m_{min}, m_{max}] \\ 0 & \text{otherwise} \end{cases} \tag{8.65}$$

Normalizing constant
The normalizing constant should be such that

$$1 = \int_{m_{min}}^{m_{max}} f_{M_D}(m)\, dm = C_{Norm} \int_{m_{min}}^{m_{max}} f_{M_G}(m) \cdot F_D(m)\, dm \tag{8.66}$$

The combination of the functions in the density of the detected magnitude distribution (8.65) leads to the above integral that cannot be expressed in terms of elementary functions. Therefore numerical integration techniques are employed to further the investigation of this detected magnitude distribution.

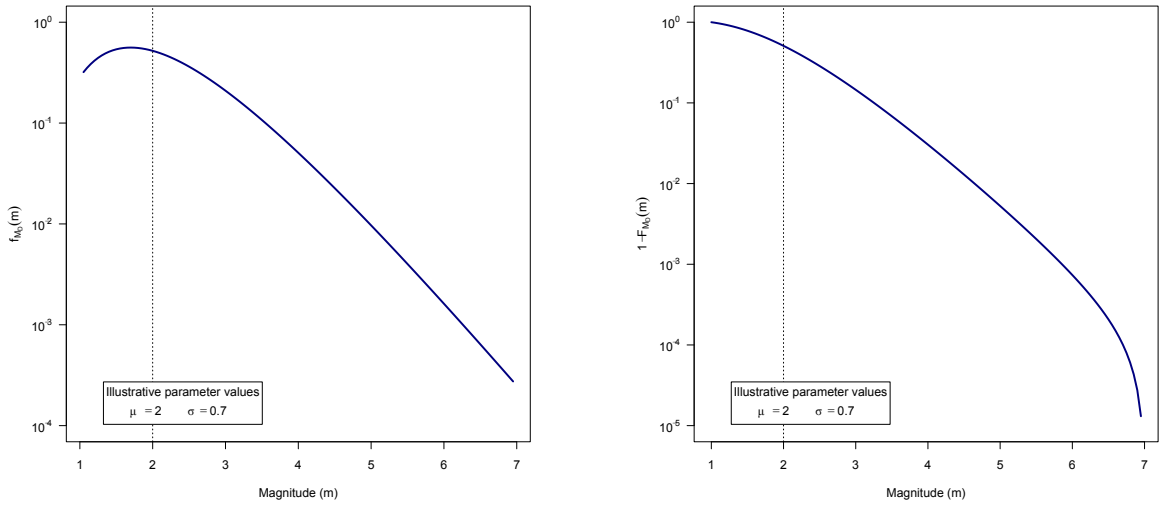
Cumulative distribution function

$$F_{M_D}(m) = \int_{m_{min}}^{m} f_{M_D}(t)\, dt \tag{8.67}$$

As stated above, values of the cumulative distribution function must be obtained through numerical methods, due to the functional form of the cumulative distribution function.

Quantile function
Due to the inability to express the cumulative distribution function in terms of elementary functions, values of the quantile function ($Q_{M_D}(p) = m$) must also be obtained numerically. The Newton-Raphson algorithm will be employed as follows to subsequently solve for $g(m) = 0$:

$$g(m) := \int_{m_{min}}^{m} f_{M_D}(t)\, dt - p \tag{8.68}$$



(a) *Density function of detected magnitude distribution. Illustrative values used together with* $\mu = 2$ *and* $\sigma = 0.7$.

(b) *Complementary distribution function of detected magnitude. Illustrative values used together with* $\mu = 2$ *and* $\sigma = 0.7$

Figure 8.7: *Sensitivity illustrations of density function of detected magnitude distribution with detection probability, over the entire magnitude range, modelled by the cumulative Log-Normal distribution function. Illustrative values for parameters include :* $m_{min} = 1$, $m_{max} = 7$ *and b-Value* $= 0.9$. *Reference lines are included at* μ.

8.2.6.2 Sharp detection threshold

An explicit magnitude break in the detection probability can be introduced.

$$f_{M_D}(m) \propto \begin{cases} f_{M_G}(m) \cdot F_D^T(m) & \text{if } m \in [m_{min}, m_c) \\ f_{M_G}(m) & \text{if } m \in [m_c, m_{max}] \\ 0 & \text{otherwise} \end{cases} \tag{8.69}$$

Where the superscript T indicates that the detection probability distribution is truncated from the right at the point m_c.

Normalizing constant

The normalizing constant should be such that

$$\begin{aligned} 1 &= C_{Norm} \int_{m_{min}}^{m_{max}} f_{M_D}(m)\, dm \\ &= C_{Norm} \left(\int_{m_{min}}^{m_c} f_{M_G}(m) \cdot F_D(m)\, dm + 1 - F_{M_G}(m_c) \right) \end{aligned} \tag{8.70}$$

As in the previous section, where a soft detection threshold has been modelled by a cumulative Log-Normal distribution function, due to the form of the constituent functions of $f_{M_D}(\cdot)$ the above integral cannot be expressed in terms of elementary functions. This necessitates the use of numerical integration methods to obtain a value of the normalizing constant.

Cumulative distribution function

$$F_{M_D}(m) = \begin{cases} 0 & \text{if } m < m_{min} \\ C_{Norm} \int_{m_{min}}^{m} f_{M_G}(t) \cdot F_D(t)\, dt & \text{if } m \in [m_{min}, m_c) \\ C_{Norm} \left(\int_{m_{min}}^{m_c} f_{M_G}(t) \cdot F_D(t)\, dt + F_{M_G}(m) - F_{M_G}(m_c) \right) & \text{if } m \in [m_c, m_{max}) \\ 1 & \text{if } m \geq m_{min} \end{cases} \tag{8.71}$$

For the same reason as for the normalizing constant, the above functions must also be evaluated numerically.

Quantile function

1. For $m \in [m_{min}, m_c)$
   Due to the inability of expressing the cumulative distribution function in terms of elementary functions, values of the quantile function ($Q_{M_D}(p) = m$) must be obtained by use of numerical methods. The Newton-Raphson algorithm will be employed by defining

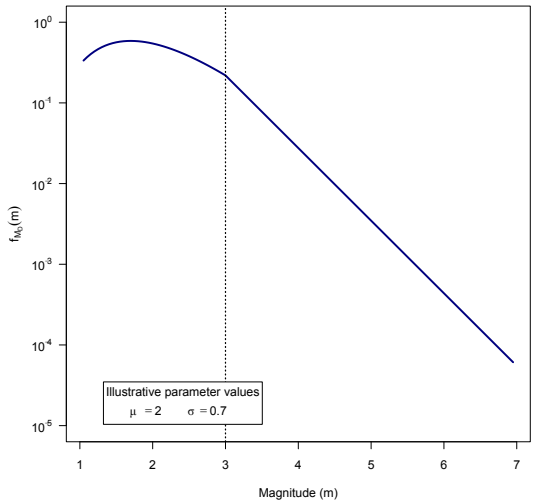
   $$g(m) := \int_{m_{min}}^{m} f_{M_D}(t)\, dt - p \tag{8.72}$$

   and solving for $g(m) = 0$.

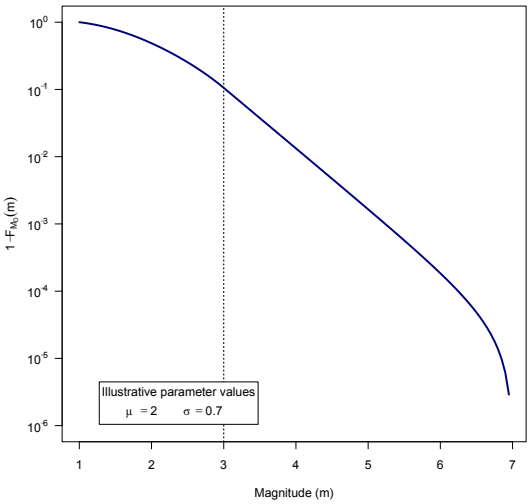
2. For $m \in [m_c, m_{max}]$

   From (8.71), the quantile function $Q_{M_D}(m) = p$ can be explicitly expressed as

   $$m = Q_{M_G} \left(\frac{p - F_D(m_c)}{C_{Norm}} + F_{M_G}(m_c) \right) \tag{8.73}$$



(a) *Density function of detected magnitude distribution. Illustrative values used together with* $\mu = 2$ *and* $\sigma = 0.7$.



(b) *Complementary distribution function of detected magnitude. Illustrative values used together with* $\mu = 2$ *and* $\sigma = 0.7$

Figure 8.8: *Sensitivity illustrations of density function of detected magnitude distribution with detection probability, over the entire magnitude range, modelled by the cumulative Log-Normal distribution function. Illustrative values for parameters include :* $m_{min} = 1$, $m_c = 3$, $m_{max} = 7$ *and b-Value* $= 0.9$. *Reference lines are included at* m_u *and at* m_c.

8.2.7 Parameter sensitivities of detected magnitude distributions

In this section some of the parameter sensitivities of the derived distributions in Sections 8.2.3 up to 8.2.6 will be investigated. As will be seen, the parameters of the distribution used to model the detection probability have an influence on the eventual detected magnitude distribution. These parameters have differing interpretations for the various detected magnitude distributions.

It can be seen that the three parameter types responsible for characterizing the detection probability are

1. **Location parameters**

   Where such parameters affect the location of a probability distribution (i.e. rigid shifts of the entire density function, either to the left or right of a specific starting value). Therefore this parameter affects the effectiveness of detecting events. Shifts in this quantity cause a uniform displacement of probability over the support of the random variable.

2. **Scale parameters**

   Scale parameters affect the statistical dispersion of probability distributions. For smaller values of scale parameters, the distribution is more localized. This in contrast to larger valued scale parameters, where the resulting distribution is spread out over a larger region. Such parameters affect the rate at which detection changes.

3. **Shape parameters**

   This is a parameter that is not classified as a scale or location parameter (not a shift or stretch of the probability distribution), but directly affects the shape of the distribution. Shape parameters determine the characteristic progression of a detection probability, e.g. a concave, versus a convex progression of a detection probability.

An interesting point to note is that the amount of influence individual parameters have on a distribution varies when considering the spectrum of values of the other remaining parameters. These profile sensitivities indicate the relationship between the distributional parameters.

For the current study the parameters of the cumulative probability distributions used to model the detection probability can be tabulated as follows:

		Distributional parameter type		
		Location	Scale	Shape
CDF	Normal distribution	μ	σ	-
	Logistic distribution	μ	s	-
	Log-Normal distribution	μ (log-scale)	-	σ
	Pareto type II distribution	μ	σ	α

Table 8.2: *Categorization of different parameter types for the various cumulative distribution functions used to model the event detection probability in the current study.*

8.2.7.1 Soft detection threshold

For distributions where a soft detection threshold has been incorporated, the physical interpretation of the above parameters can be directly translated from their respective influences on the particular probability distribution. The reason for this is that the probability of event detection is modelled by the various well-studied cumulative probability distributions, in their respective natural forms.

The sensitivities of the detection probability to the parameters can be seen when considering the quantile function of the different forms of the detection probability. Intuitively, by considering the quantile function, shifts of the probability distribution can be quantified by fixing a detection probability p and estimating the difference in possible observed magnitudes for that detection probability.

1. **Sensitivity to location parameter**

   When considering the quantile functions as described in (B.45), (B.51) and (B.63) for the Normal, Logistic and Pareto type II distributions, it can be seen that shifts in the quantity μ cause a displacement of probability, but uniformly over the support of the random variable. Furthermore, for the case of the Normal and Logistic form of the detection probability an event with magnitude μ has exactly a 50/50 chance of being detected or missed.

   Although not strictly seen as a location parameter, the value of μ in the case where the detection probability is modelled by a cumulative Log-Normal distribution, affects the detected magnitude distribution in a multiplicative exponential fashion. This can be seen by the quantile function as described in (B.56).

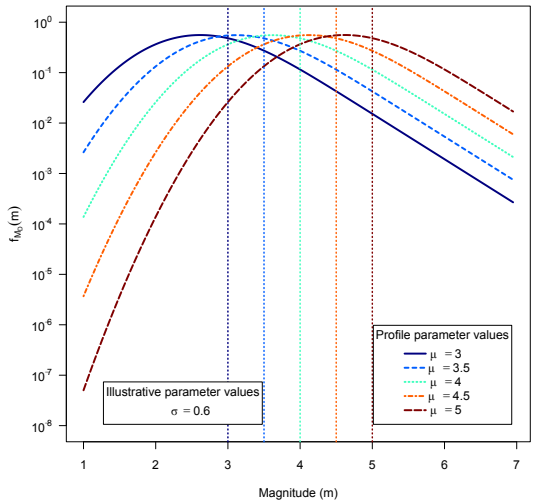
   Therefore, the sensitivity of the distributions to the location parameter can be summarized as follows:

   (a) Detected magnitude distributions where the detection probability is modelled by one of the symmetric distributions (Normal and Logistic) display a rigid shift in probability while retaining their shape to a fairly high degree due to shifts in the location parameter.

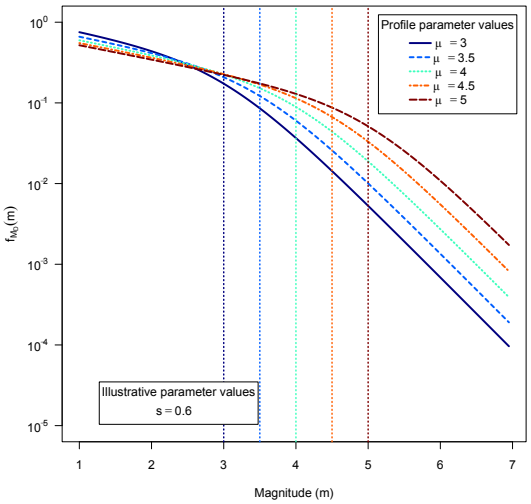
   (b) Where the detection probability is modelled by the non-symmetric Pareto type II distribution, the distribution does not retain its shape under shifts of this quantity and appears to be less sensitive to shifts than the situation described in the previous point.

   (c) The case where the detection distribution is modelled by a cumulative Log-Normal distribution is particularly susceptible to large changes when even small changes occur in the location parameter.

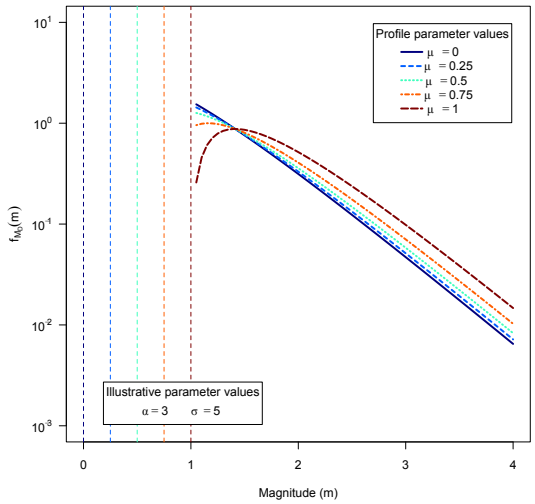
   The above summary can also be noted in the following graphic representation of the detected magnitude distributions.



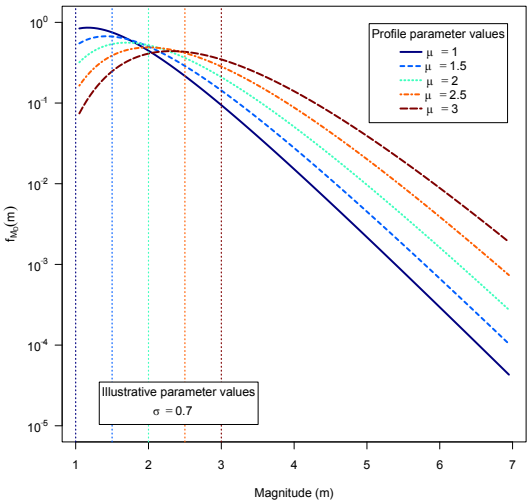
(a) *Sensitivity of density function of detected magnitude distribution to parameter μ. Illustrative values used and $\sigma = 0.6$*



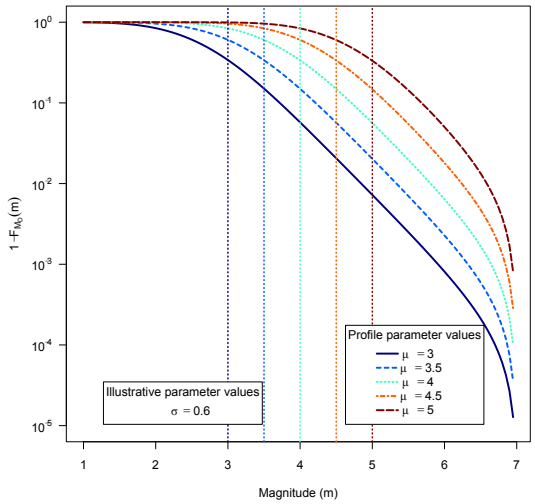
(b) *Sensitivity of density function of detected magnitude distribution to parameter μ. Illustrative values used and $s = 0.6$*



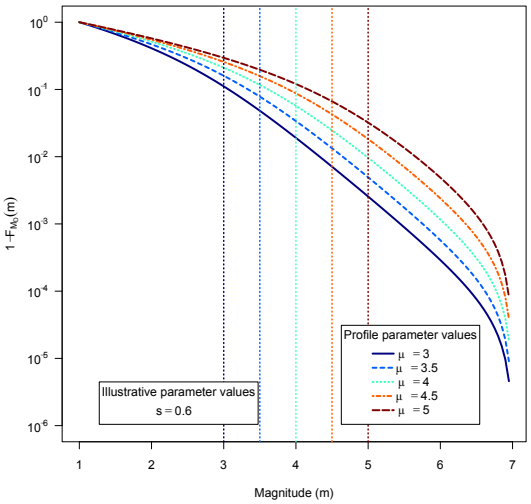
(c) *Sensitivity of density function of detected magnitude distribution to parameter μ. Illustrative values used for remaining parameters.*



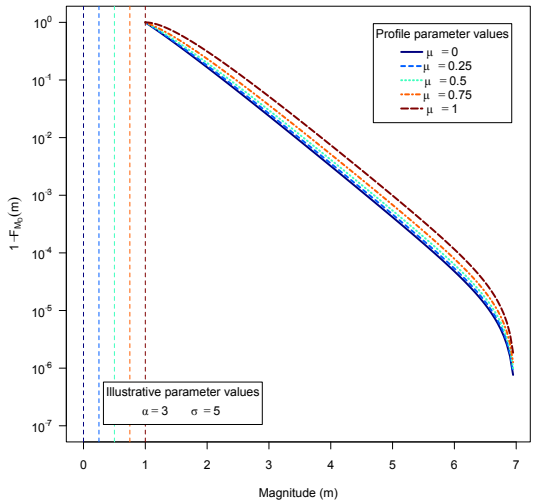
(d) *Sensitivity of density function of detected magnitude distribution to parameter μ. Illustrative values used and $\sigma = 0.7$*



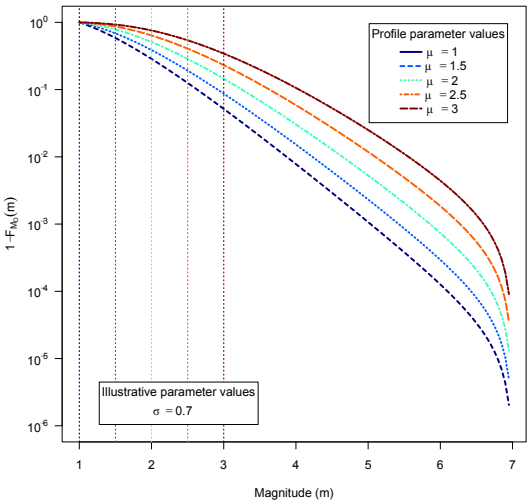
(a) *Sensitivity of complementary distribution function of detected magnitude to parameter μ. Illustrative values used and $\sigma = 0.6$*



(b) *Sensitivity of complementary distribution function of detected magnitude to parameter μ. Illustrative values used and $s = 0.7$*



(c) *Sensitivity of complementary distribution function of detected magnitude to parameter μ. Illustrative values used for remaining parameters.*



(d) *Sensitivity of complementary distribution function of detected magnitude to parameter μ. Illustrative values used and $\sigma = 0.7$*

Figure 8.10: *Sensitivity illustrations of complementary distribution function of detected magnitude with detection probability, over entire magnitude range, modelled by the cumulative Normal distribution function. Illustrative values for parameters include : $m_{min} = 1$, $m_{max} = 7$, b-Value $= 0.9$. Reference lines are included at μ*

2. **Sensitivity to scale parameter**

According to the quantile functions in (B.45), (B.51) and (B.63) for the Normal, Logistic and Pareto type II distributions, the influence of changes in the scale parameter are partly determined by the following multipliers

(a) **Cumulative Normal distribution**

$$\text{Multiplier}_1(p) = \Phi^{-1}(p) \tag{8.74}$$

(b) **Cumulative Logistic distribution**

$$\text{Multiplier}_2(p) = \ln\left(\frac{p}{1-p}\right) \tag{8.75}$$

(c) **Cumulative Pareto type II distribution**

$$\text{Multiplier}_3(p) = (1-p)^{-\frac{1}{\alpha}} - 1 \tag{8.76}$$

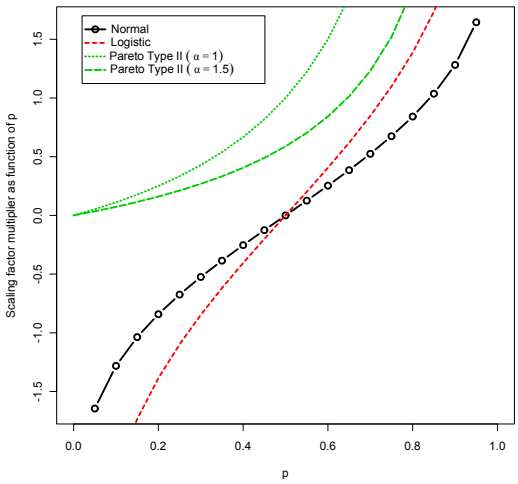


Figure 8.11: *Progression of scaling factor multiplier as function of p for different forms of detection probability as modelled by cumulative distribution functions.*

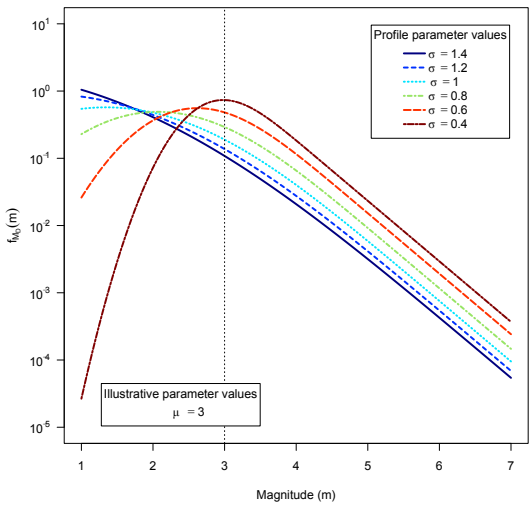
Figure 8.11 aids in the reaching of the following conclusions:

(a) As can be seen from the original quantile functions, the scale parameter affects the detection probability in a multiplicative fashion with a multiplying factor that is dependent on the exact magnitude of an event.

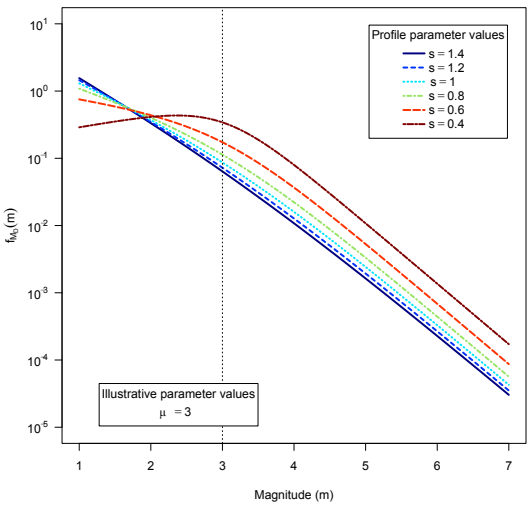
(b) The effect of the scale parameter varies over the range of the detection probability. It can be seen from Figure 8.11 that for the case of the cumulative Normal and logistic distributions, a change in scale factor is multiplied by relatively large factors and significantly influences the distribution of detected magnitudes at the low and higher end of detection probabilities.

(c) The multiplier has greater effect for the case of the Normal distribution, than for that of the Logistic distribution, for values lower than μ and the opposite is true for values greater than μ. For this reason it is believed that if a soft detection threshold is to be used, a Cumulative Normal distribution should be used, as to restrict the volatility of detection probability for larger magnitude events, since these events have a naturally bigger chance of being detected.

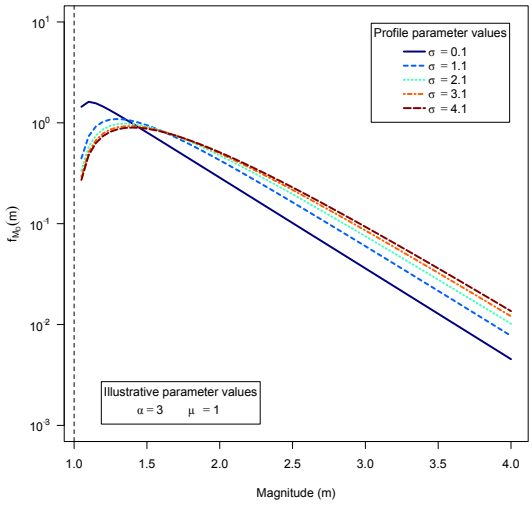
(d) If Figure 8.11 is interpreted in absolute values, it shows that the lower magnitude range (situation where the detection probability is modelled by the cumulative Pareto type II distribution) is the least susceptible to changes of the scale parameter. This due to the fact that the value of the multiplier is relatively low. Furthermore, this can be seen in Figures 8.12c and 8.13c, displaying the restricted use of this distribution.



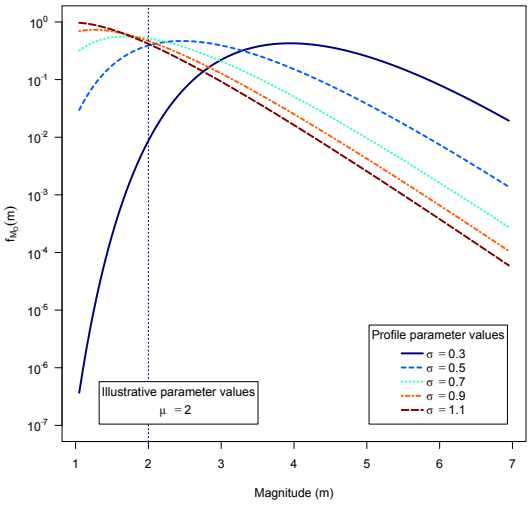
(a) *Sensitivity of density function of detected magnitude distribution to parameter σ. Illustrative values used and $\mu = 3$*



(b) *Sensitivity of density function of detected magnitude distribution to parameter s. Illustrative values used and $\mu = 3$*

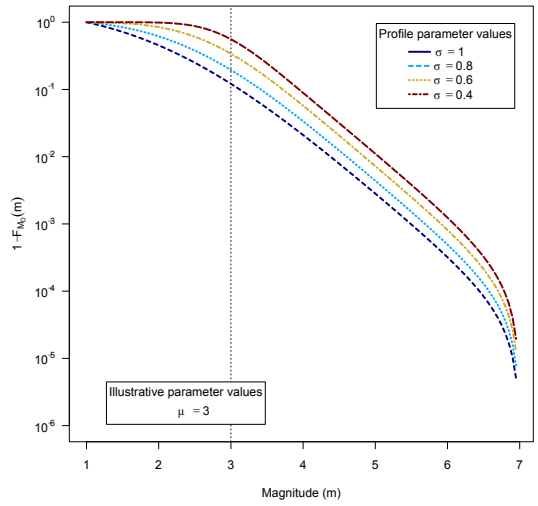


(c) *Sensitivity of density function of detected magnitude distribution to parameter σ. Illustrative values used for remaining parameters.*

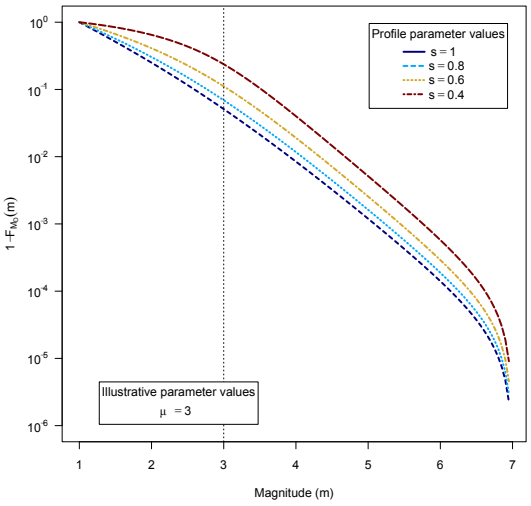


(d) *Sensitivity of density function of detected magnitude distribution to parameter σ. Illustrative values used and $\mu = 2$*

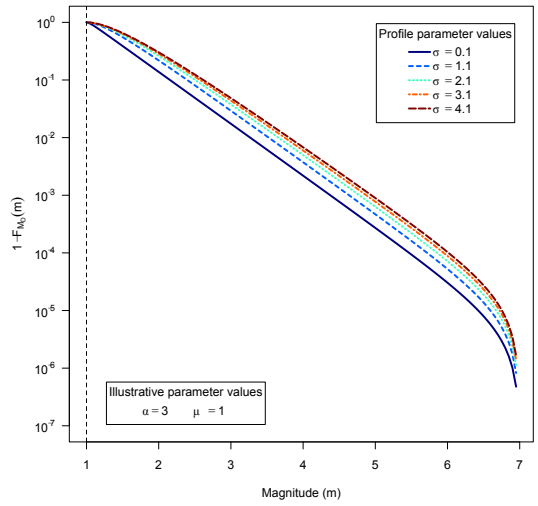
Figure 8.12: *Sensitivity illustrations of density function of detected magnitude distribution (8.15) with detection probability, over entire magnitude range, modelled by the cumulative Normal distribution function. Illustrative values for parameters include : $m_{min} = 1$, $m_{max} = 7$, b-Value $= 0.9$. Reference lines are included at μ*



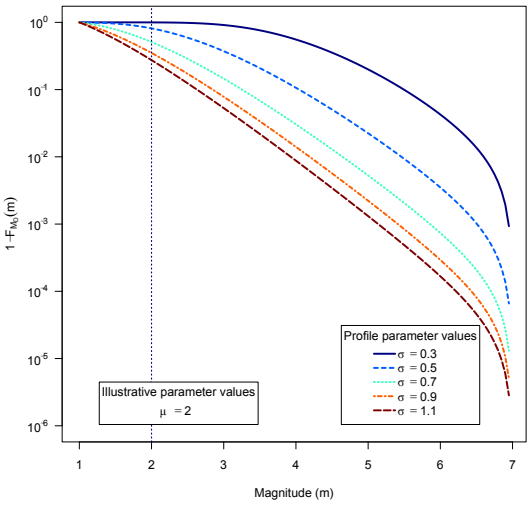
(a) *Sensitivity of complementary distribution function of detected magnitude to parameter σ. Illustrative values used and $\mu = 3$*



(b) *Sensitivity of complementary distribution function of detected magnitude to parameter s. Illustrative values used and $\mu = 3$*



(c) *Sensitivity of complementary distribution function of detected magnitude to parameter σ. Illustrative values used for remaining parameters.*



(d) *Sensitivity of complementary distribution function of detected magnitude to parameter σ. Illustrative values used and $\mu = 2$*

Figure 8.13: *Sensitivity illustrations of complementary distribution function of detected magnitude with detection probability, over entire magnitude range, modelled by the cumulative Normal distribution function. Illustrative values for parameters include : $m_{min} = 1$, $m_{max} = 7$, b-Value $= 0.9$. Reference lines are included at μ*

3. **Sensitivity to shape parameter**

Figures 8.14 and 8.15, show that the detected magnitude distribution is relatively impervious to changes in the value of the shape parameter where the detection probability is modelled by a cumulative Pareto type II distribution. On the other hand Figures 8.12d and 8.13d show the extreme flexibility of the detected magnitude distribution to changes in the shape parameter when

the detection probability is modelled by the Log-Normal distribution. This extreme sensitivity can be seen through inspection of the quantile function, where a multiplicative effect can be observed relating to changes in the shape parameter.

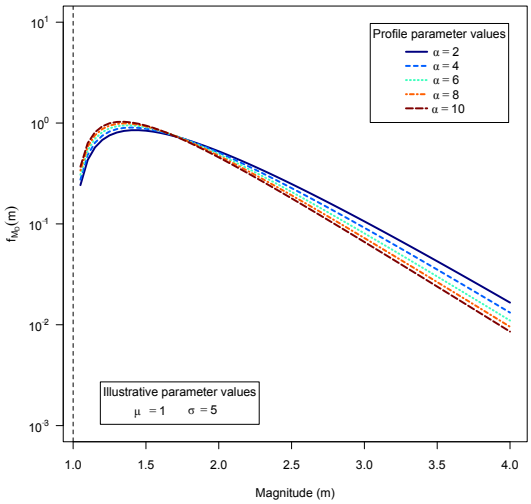


Figure 8.14: *Sensitivity illustrations of density function of detected magnitude distribution with detection probability, over entire magnitude range, modelled by the cumulative Generalized Pareto distribution function. Illustrative values for parameters include :* $m_{min} = 1$, $m_{max} = 7$, *b-Value* $= 0.9$, $\mu = m_{min}$, $\sigma = 5$ *and* $\alpha = 3$.

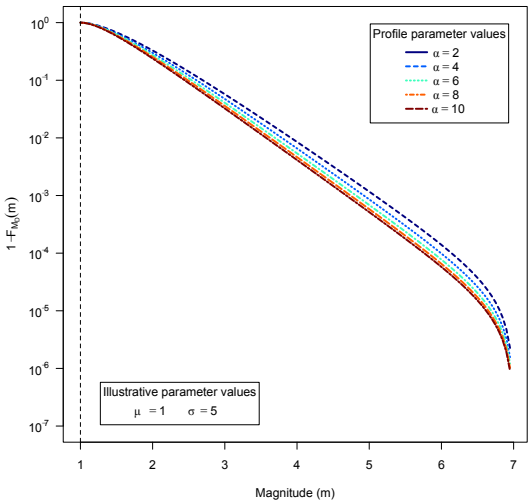


Figure 8.15: *Sensitivity illustrations of complementary distribution function of detected magnitude with detection probability, over entire magnitude range, modelled by the cumulative Pareto distribution function. Illustrative values for parameters include :* $m_{min} = 1$, $m_{max} = 7$, *b-Value* $= 0.9$, $\mu = m_{min}$, $\sigma = 5$ *and* $\alpha = 3$.

8.2.7.2 Sharp detection threshold

Pertaining to the case of a sharp detection threshold, this study has placed a restriction on the detection probability to take on the value 1 at the detection threshold. By replacing the probability p in the quantile functions the following adjusted quantile functions are applicable :

1. **Normal Distribution**

$$m = \mu + \Phi^{-1}\left(\frac{p}{\Phi(z_{mc})}\right)\sigma \tag{8.77}$$

Where $\Phi(\cdot)$ is the cumulative distribution function of a $\mathcal{N}(0,1)$ random variable and $z_{m_c} = \dfrac{m_c - \mu}{\sigma}$.

2. **Logistic Distribution**

$$m = \mu + s\left(\ln\left(\frac{p}{F(m_c)}\right) - \ln\left(1 - \frac{p}{F(m_c)}\right)\right) \tag{8.78}$$

Where $F(\cdot)$ is the cumulative distribution function of a Logistic(μ, s) random variable.

3. **Log-Normal Distribution**

$$m = \exp\left(\mu + \sigma\Phi^{-1}\left(\frac{p}{F(m_c)}\right)\right) \tag{8.79}$$

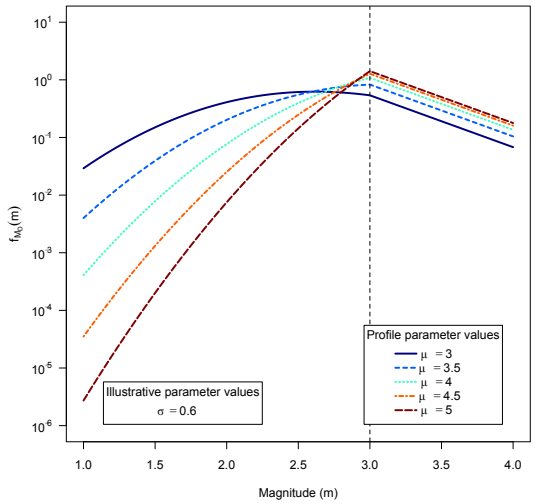
Where $F(\cdot)$ is the cumulative distribution function of a LN(μ, σ^2) random variable.

4. **Pareto type II Distribution**

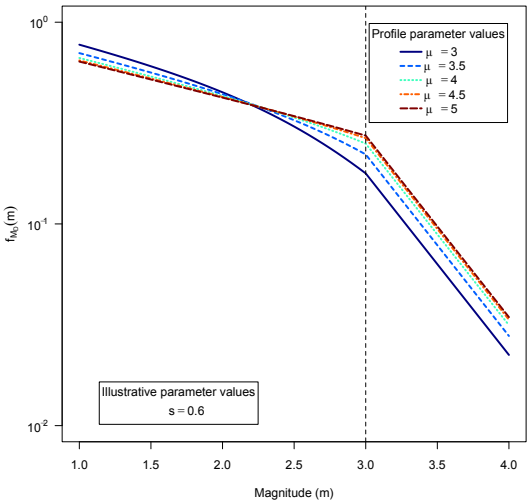
$$m = \mu + \sigma\left(\left(1 - \frac{p}{F(m_c)}\right)^{-\frac{1}{\alpha}} - 1\right) \tag{8.80}$$

Where $F(\cdot)$ is the cumulative distribution function of a P(II)(μ, σ, α) random variable.

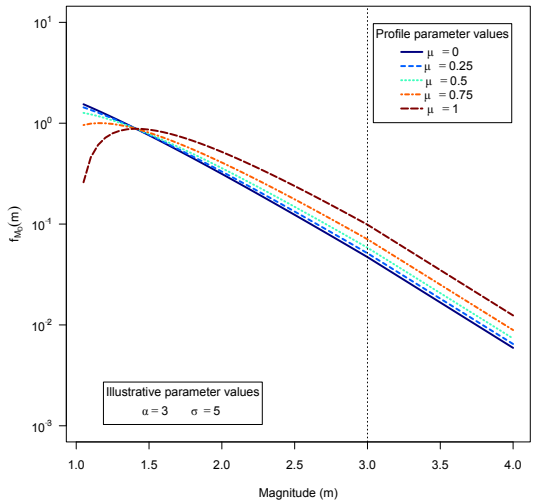
Through the replacement of the variable p by the, generic function, $\dfrac{p}{F(m_c)}$, for differing forms of the detection probability, the interpretation of the parameters have been confounded. All parameters now play a part in the shaping and scaling of the detected magnitude distribution. However, through inspection of the figures illustrating parameter sensitivities in the previous sections, it has been determined that the resulting detected magnitude threshold displays the same characteristics with regards to parameter sensitivities. The only difference is that these sensitivities only affect the region below the detection threshold. Therefore, the general interpretation of sensitivities in the section pertaining to soft detection thresholds can be utilized to this end.



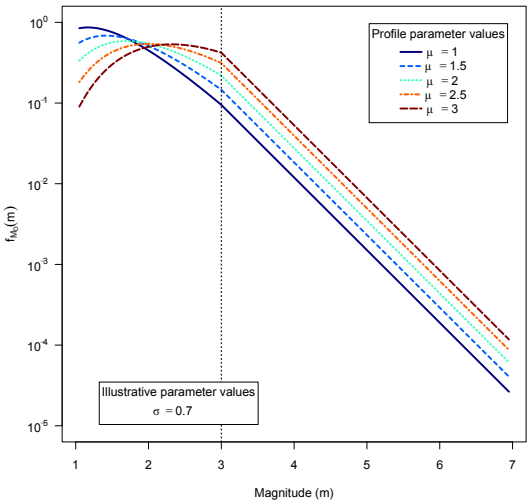
(a) *Sensitivity of density function of detected magnitude distribution to parameter μ. Illustrative values used and $\sigma = 0.6$*



(b) *Sensitivity of density function of detected magnitude distribution to parameter μ. Illustrative values used and $s = 0.6$*

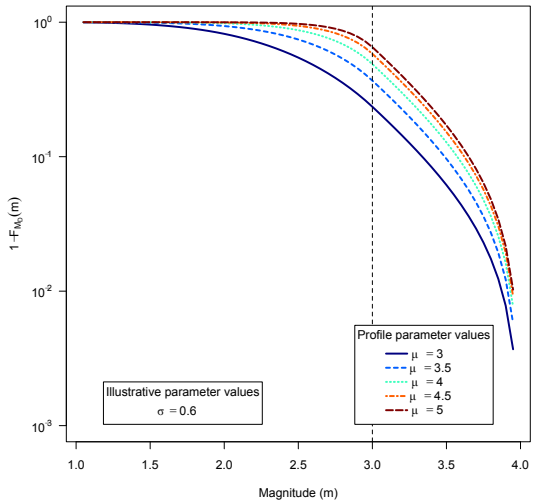


(c) *Sensitivity of density function of detected magnitude distribution to parameter μ. Illustrative values used for remaining parameters.*

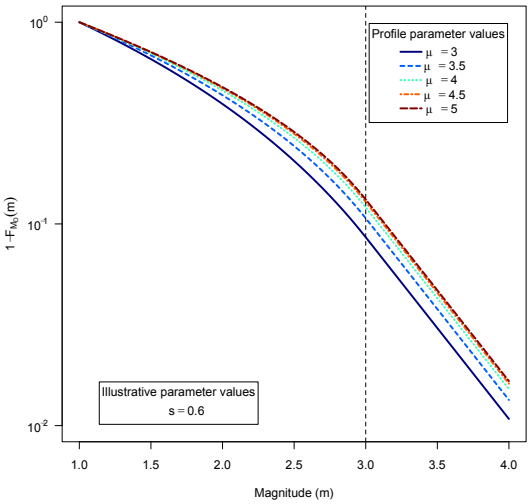


(d) *Sensitivity of density function of detected magnitude distribution to parameter μ. Illustrative values used and $\sigma = 0.7$*

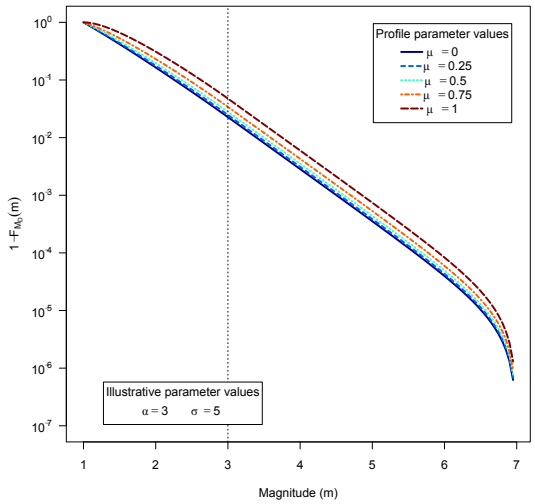
Figure 8.16: *Sensitivity illustrations of density function of detected magnitude distribution with detection probability, below m_c, modelled by the cumulative Normal distribution function. Illustrative values for parameters include : $m_{min} = 1$, $m_c = 3$, $m_{max} = 7$, b-Value $= 0.9$. Reference lines are included at m_c*



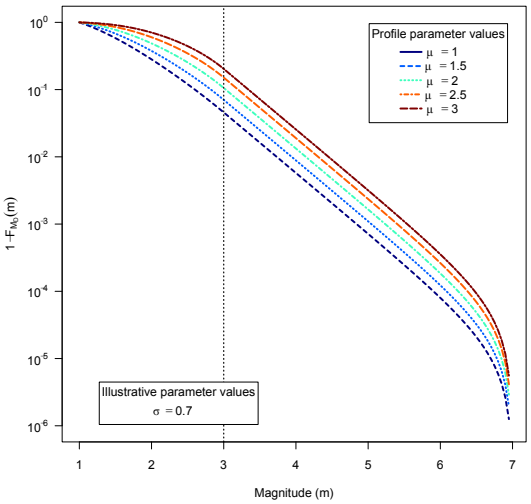
(a) *Sensitivity of complementary distribution function of detected magnitude to parameter μ. Illustrative values used and $\sigma = 0.6$*



(b) *Sensitivity of complementary distribution function of detected magnitude to parameter μ. Illustrative values used and $s = 0.6$*

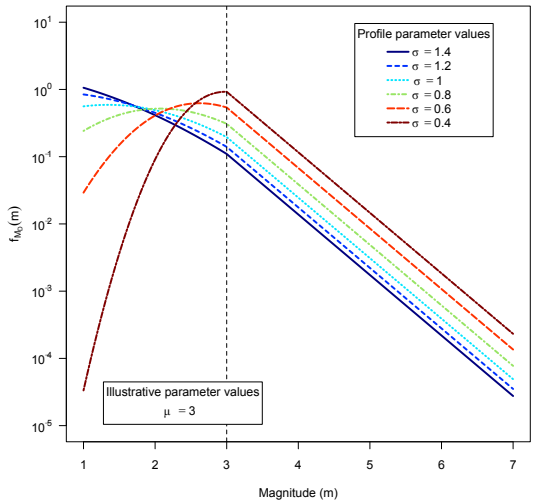


(c) *Sensitivity of complementary distribution function of detected magnitude to parameter μ. Illustrative values used for remaining parameters.*

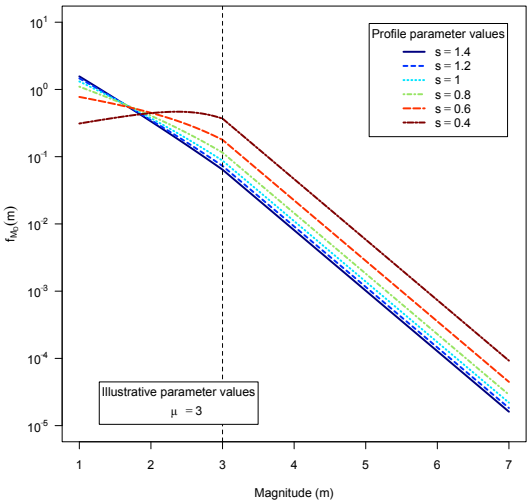


(d) *Sensitivity of complementary distribution function of detected magnitude to parameter μ. Illustrative values used and $\sigma = 0.7$*

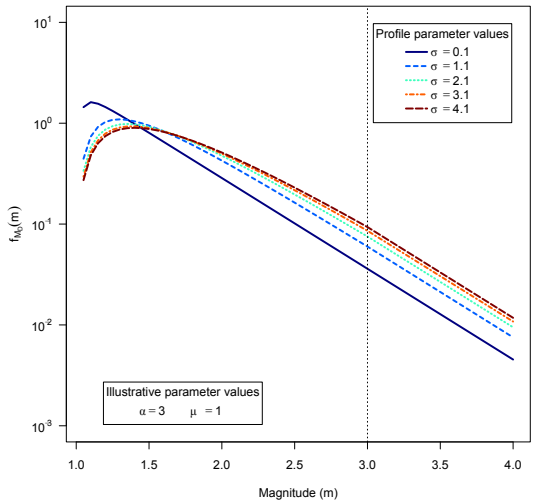
Figure 8.18: *Sensitivity illustrations of complementary distribution function of detected magnitude with detection probability, below m_c, modelled by the cumulative logistic distribution function. Illustrative values for parameters include : $m_{min} = 1$, $m_c = 3$, $m_{max} = 7$, b-Value $= 0.9$. Reference lines are included at m_c*



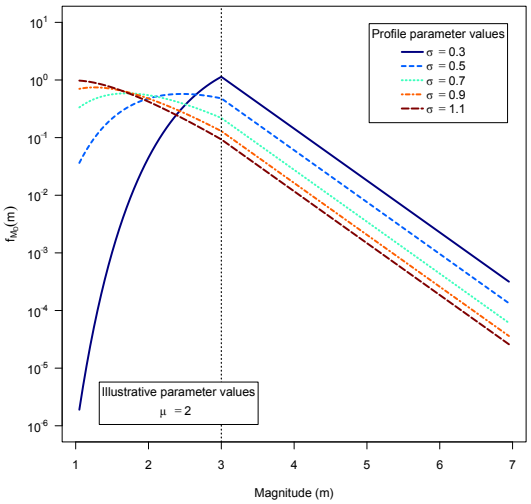
(a) *Sensitivity of density function of detected magnitude distribution to parameter σ. Illustrative values used and $\mu = 3$*



(b) *Sensitivity of density function of detected magnitude distribution to parameter s. Illustrative values used and $\mu = 3$*

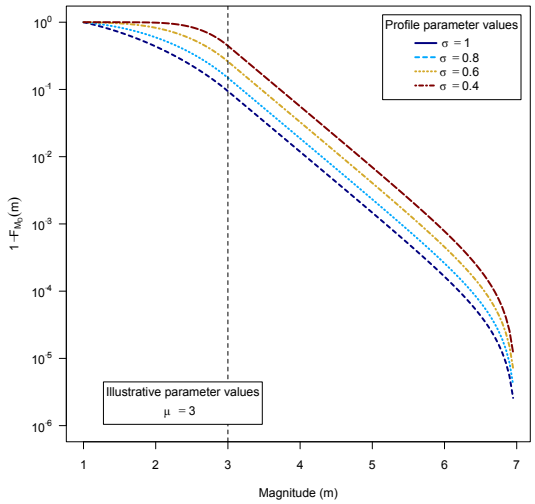


(c) *Sensitivity of density function of detected magnitude distribution to parameter σ. Illustrative values used for remaining parameters.*

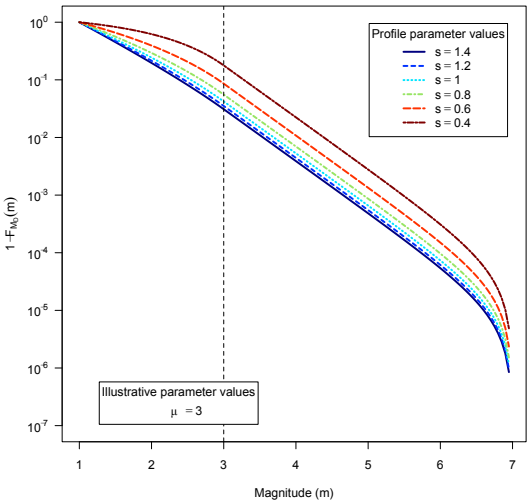


(d) *Sensitivity of density function of detected magnitude distribution to parameter σ. Illustrative values used and $\mu = 2$*

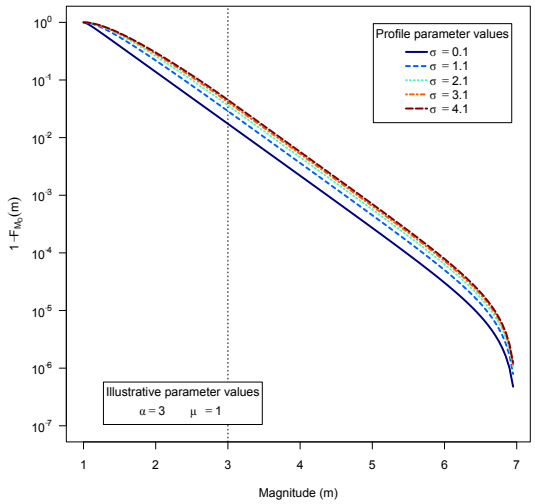
Figure 8.19: *Sensitivity illustrations of density function of detected magnitude distribution with detection probability, below m_c, modelled by the cumulative Normal distribution function. Illustrative values for parameters include : $m_{min} = 1$, $m_c = 3$, $m_{max} = 7$, b-Value $= 0.9$. Reference lines are included at m_c*



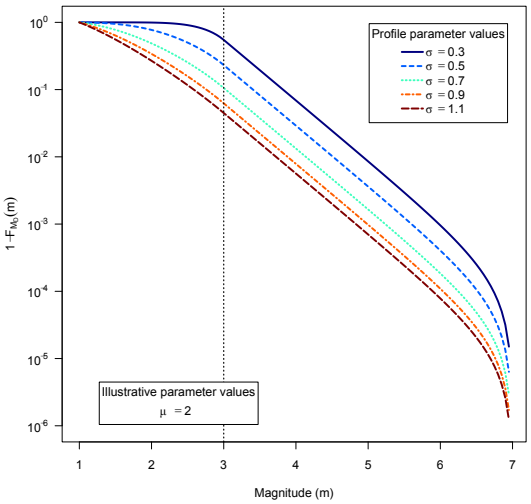
(a) *Sensitivity of complementary distribution function of detected magnitude to parameter σ. Illustrative values used and $\mu = 3$*



(b) *Sensitivity of complementary distribution function of detected magnitude to parameter s. Illustrative values used and $\mu = 3$*



(c) *Sensitivity of complementary distribution function of detected magnitude to parameter σ. Illustrative values used for remaining parameters.*



(d) *Sensitivity of complementary distribution function of detected magnitude to parameter σ. Illustrative values used and $\mu = 2$*

Figure 8.20: *Sensitivity illustrations of complementary distribution function of detected magnitude with detection probability, below m_c, modelled by the cumulative Normal distribution function. Illustrative values for parameters include : $m_{min} = 1$, $m_c = 3$, $m_{max} = 7$, b-Value $= 0.9$. Reference lines are included at m_c*

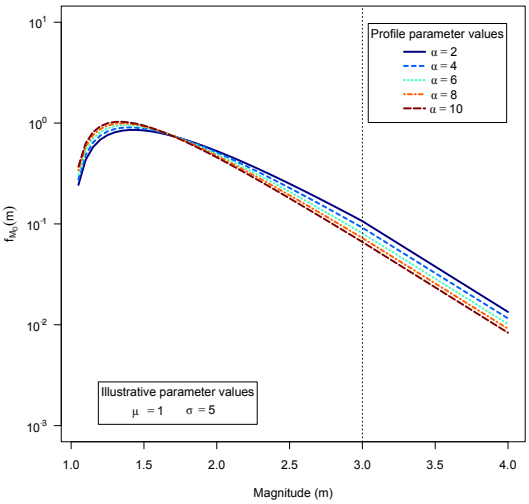


Figure 8.21: *Sensitivity illustrations of density function of detected magnitude distribution with detection probability, below m_c, modelled by the cumulative Generalized Pareto distribution function. Illustrative values for parameters include : $m_{min} = 1$, $m_c = 3$, $m_{max} = 7$, b-Value $= 0.9$, $\mu = m_{min}$, $\sigma = 5$ and $\alpha = 3$. Reference lines are included at m_c.*

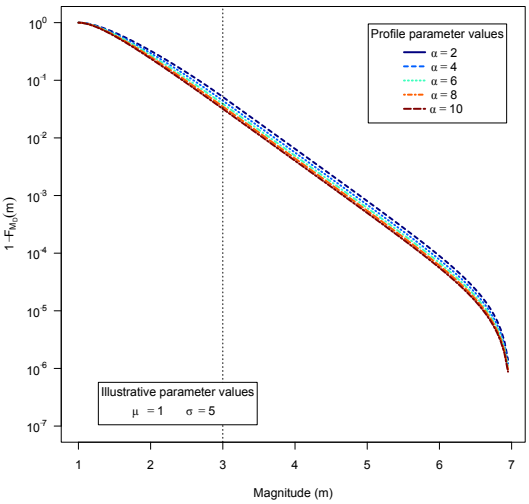


Figure 8.22: *Sensitivity illustrations of complementary distribution function of detected magnitude with detection probability, below m_c, modelled by the cumulative Pareto distribution function. Illustrative values for parameters include : $m_{min} = 1$, $m_c = 3$, $m_{max} = 7$, b-Value $= 0.9$, $\mu = m_{min}$, $\sigma = 5$ and $\alpha = 3$. Reference lines are included at m_c.*

9 Description of threshold estimation methods

In this section some of the prevailing threshold estimation methods, as found in earthquake literature, are described. The methods are described as presented in the relevant papers, however, in this section it has been aimed to describe the methods in a rigorous manner. This is done to aid in the understanding as well as objective formulations. Before continuing, a brief summary of concepts, values and functions utilized in the threshold estimation methods is provided.

Definitions

$\boldsymbol{S}$ The set of all observed events, with magnitudes in the range $[m_{min}, m_{max}]$, originating within a specified time- and space window of the dataset under investigation.

$\boldsymbol{m_i}$ The magnitude of the i^{th} recorded event of a particular dataset A, i.e. $m_i \in A$.

$\boldsymbol{\Delta m}$ The binning-width of events of dataset.

$\boldsymbol{m_i^b}$ The left end-point of the i^{th} event magnitude bin, $i = 1, 2, \ldots, k$. Furthermore the distance between consecutive m_i^b's $(m_{i+1}^b - m_i^b)$ is equal to the binning width Δm.

$\boldsymbol{P_E}$ The standard partitioning of the entire magnitude range based on the sequence of m_i^b's, where $m_{min} = m_1^b < m_2^b < \ldots m_i^b < m_{i+1}^b \ldots < m_k^b < m_{k+1}^b = m_{max}$, which leads to k subintervals defined by the successive m_i^b's.

$\boldsymbol{P_O(A)}$ The partitioning of the interval $\left[\lfloor \frac{1}{\Delta m} \min(A) \rfloor \Delta m, \lceil \frac{1}{\Delta m} \max(A) \rceil \Delta m\right]$ which contains the range of magnitudes (m_i) in the set A. The partition is a sequence (m_i^b) of the form $m_1^b < m_2^b < \ldots m_i^b < m_{i+1}^b \ldots < m_t^b < m_{t+1}^b$, where $m_1^b = \lfloor \frac{1}{\Delta m} \min(A) \rfloor \Delta m$ and $m_{t+1}^b = \lceil \frac{1}{\Delta m} \max(A) \rceil \Delta m$, which leads to t subintervals defined by the successive m_i^b's.

$\boldsymbol{H(A;P)}$ The histogram of the elements in the set A, based on the bins as described by partition P. The number of elements in the i^{th} bin will be denoted by $H(A;P)_i$.

$\boldsymbol{S_i}$ The i^{th} subset of the set S which only includes events with magnitudes greater than magnitude m_i^b, i.e. $S_i = \{m \in S : m \geq m_i^b\}$, where $i = 1, 2, \ldots k$.

$\boldsymbol{\hat{m}_c}$ The estimated magnitude of complete reporting.

For illustrative purposes, for each estimation method, an example of the implementation will be included. Such examples of implementation will be based on simulated data, where the detection probability for the incomplete portion of the distribution is modelled by a Normal CDF. Illustrative values for distributional parameters are as follows : $m_{min} = 1$, $m_c = 3$, $m_{max} = 7$, b-Value $= 0.9$, $\mu = 5$ and $\sigma = 1$.

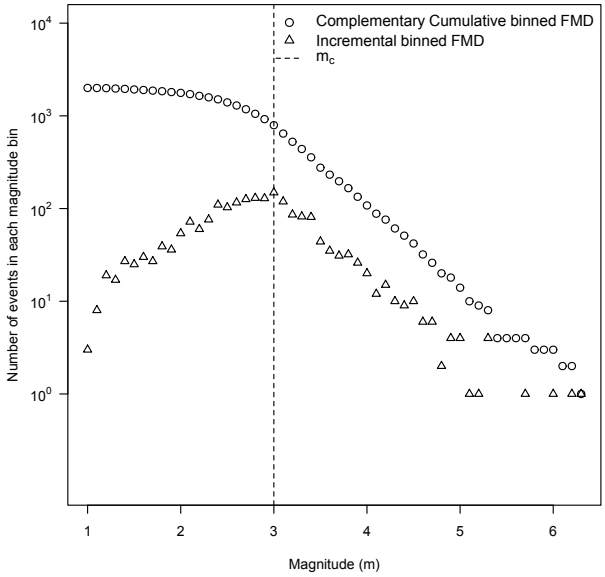


Figure 9.1: *Generated Seismic catalogue with* $m_c = 3$

All implementations of the threshold estimation methods, except MBASS, have been attached as code extracts in Appendix I. The coding of the MBASS method has been used as supplied by the author in the published academic article.

9.1 Goodnes of fit estimation method (GOF)

As described by Wiemer and Wyss [59], the basis of this method is the validity of the Gutenberg-Richter [22] relation, as seen in (2.1). The magnitude of completeness is defined as the point in the FMD that allows for the best fit of the data from the complete portion of the magnitude range to the Gutenberg-Richter model.

Assumptions made

1. The assumption is made that the data, above the magnitude of completeness, obey an exponential law.

2. The magnitude distribution below m_c is assumed not to follow the same exponential distribution as the data above m_c. Further specification of the distribution over the range $[m_{min}, m_c]$ is omitted. The result is that the minimum magnitude is specified to be m_c.

Outline of method

1. For a given set of data S, a- and b-values are computed for each subset S_i, as if reporting were complete over the magnitude range $[m_i^b, m_{max}]$. These subset specific a- and b-values will be indicated through use of a relevant subscript, i.e. a_i and b_i. The a and b values are determined, under the invariance property of maximum likelihood estimators, by using the maximum likelihood estimator $\hat{\beta}$.

2. For each pair a_i- and b_i-values calculated, a synthetic frequency-magnitude distribution is constructed for $m \in [m_{min};\ m_{max}]$, i.e. $F_{M_G}(m; b_i)$. This distribution represents a perfect fit to the estimated parameters.

3. The histogram of observed seismic events, denoted by $H_O(S_i; P_m)$, is computed for each subset S_i, where the histogram has bin counts $H_O(S_i; P_m)_j$ for $j = 1, 2, \ldots, k$.

4. The expected number of events in each interval of the partition P_m is calculated based on the synthetic distribution $F_{M_G}(m; b_i)$. This is analogous to a histogram of the expected events, denoted by $H_E(S; P; b_i)$ with bin counts $H_E(S; P; b_i)_j$ for $j = 1, 2, \ldots, k$.

5. The extent to which the observed data deviated from the expected (synthetic) data is obtained by calculating a R statistic for each S_i:

$$R(a_i, b_i, m_i^b) = 100 \left(1 - \frac{\sum_{j:m_j \in [m_i^b; m_{max}]} |H_O(S_i; P)_j - H_E(S; P; b_i)_j|}{\sum_{j:m_j \in [m_{min}, m_{max}]} H_E(S; P; b_i)_j} \right) \tag{9.1}$$

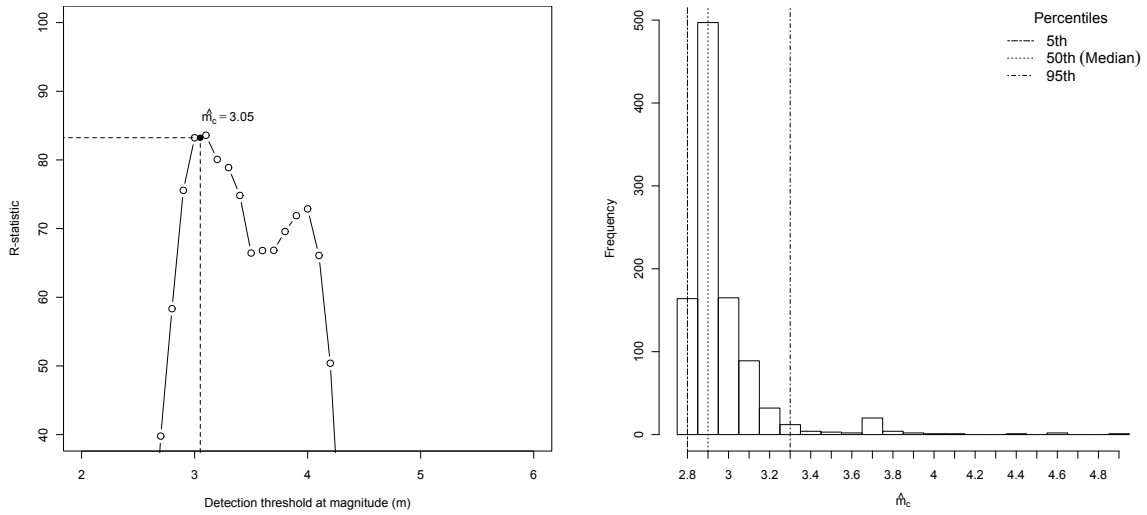
The summation in the numerator is taken over all magnitude bins starting from m_i^b up to m_{max}. The summation in the denominator is taken over all detected event magnitudes in order to, according to Wiemer and Wyss, "normalize the distribution". This is however to ensure that a lower deviation based on a smaller number of bins carries a proportionally lower weight when comparing R statistics for different m_i^b's.

6. The magnitude of completeness of the set of observations S from the particular catalogue is taken as $\hat{m}_c = \underset{m_i^b}{\operatorname{argmax}}\, R(a_i, b_i, m_i^b)$

Example of estimation algorithm

As stated, the illustrative dataset as described at the start of this chapter will be used to demonstrate the specified threshold estimation method. The R implementation of the GOF estimation method can be found in Appendix I.1.4 on page 206.

Figure 9.2a shows that the goodness of fit statistic R attains a maximum of 83.23 when the threshold of detection is at $m = 3$. Upon applying the binning width correction the threshold is estimated as $\hat{m}_c = 3.05$. The sampling distribution, based on 1000 bootstrap replications, yields a 90% confidence interval for m_c of [2.85, 3.35] with expected value of 2.95.



(a) *Graphical representation of modified goodness of fit statistic, $R(\cdot)$, for varying detection thresholds.*

(b) *Sampling distribution for estimated m_c based on 1000 bootstrap re-samples.*

Figure 9.2: *Graphical representation and estimation results of GOF method*

Comment on method

1. The sampling distribution is positively skewed with the method slightly underestimating the true magnitude of completeness. This is confirmed by Woessner and Wiemer in a 2005 comparison

study [60].

2. Due to the reliance of the method on magnitude binning, it is of utmost importance that the sample be of sufficient size such that no empty bins are present at the lower magnitude range.

3. Due to the definition of the magnitude of completeness within this method, the result is heavily influenced by binning width, as well as the location of the magnitude bins. Hence, the estimated value can merely be seen as an approximation, which can perhaps be used to validate other, more rigorous estimation methods. Alternatively, such result may be used as a starting value for iterative schemes.

9.2 Maximum curvature method (MAXC)

A non-parametric estimation method of Wiemer and Wyss [59] that attempts to determine the magnitude of completeness in a fast and reliable manner. The method determines m_c as the point of maximum curvature of the frequency-magnitude distribution.

Assumptions made

1. The method does not assume any specific distribution for the complete or incomplete portion of the catalogue. The method only assumes general characteristics of the distribution.
2. The method assumes that the frequency of events, larger than the magnitude of completeness is a decreasing function of event size. Such behaviour can typically be seen in exponentially distributed data.
3. Below m_c the frequency of events is assumed to be some non-decreasing function of event size. Further specification of the relationship between event size and frequency is not made.

Outline of method

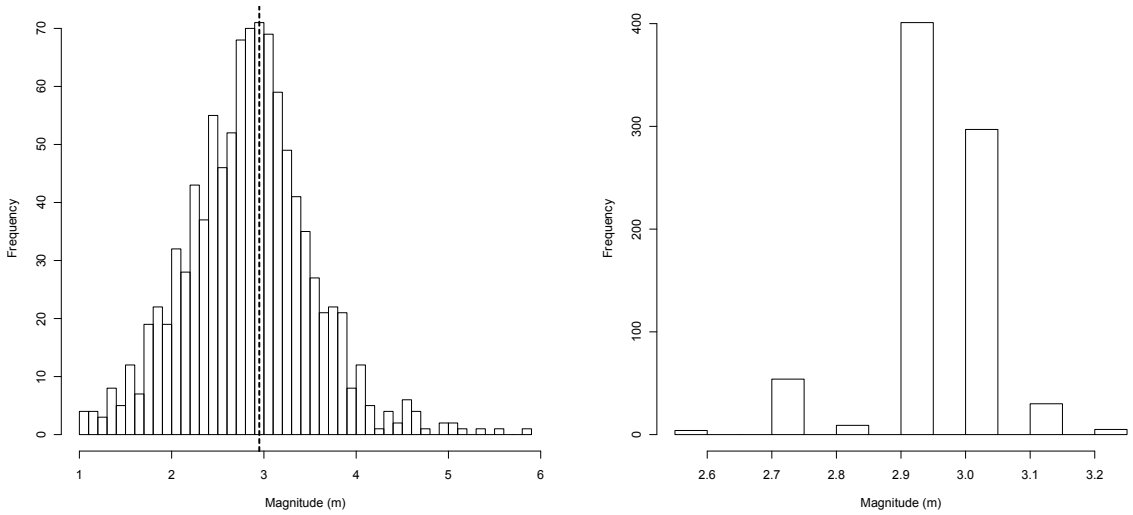
Wiemer and Wyss define curvature as the first derivative of the frequency-magnitude curve. Under the assumptions described in point 2 and 3 above, the point of maximum curvature of the FMD lies within the magnitude bin of the incremental FMD holding the highest number of events. The resulting implementation of this estimation method is relatively simple :

1. Compute the incremental histogram of observed events $H_O(S, P_m)$
2. The magnitude of completeness is estimated as $\hat{m}_c = m_j^b$, where $j = \underset{i}{\operatorname{argmax}} H_O(S; P_m)_i$.
3. Due to the fact that magnitude bins are labelled by the starting magnitude of the bin (m_i^b), modification, to take account of the binning width can be made. Under the assumption that events are uniformly distributed over each magnitude bin, the magnitude of completeness can be estimated as $\hat{m}_c = m_j^b + \frac{\Delta M}{2}$, where $j = \underset{i}{\operatorname{argmax}} H_O(S; P_m)_i$.

Example of estimation algorithm

The R implementation of the MAXC estimation method can be found in Appendix I.1.5 on page 207. As shown in Figure 9.3a, the estimated magnitude of completeness falls in the bin with the highest frequency of events, which is in the bin with left endpoint at $m = 2.9$. After incorporating the binning width correction it is found that $\hat{m}_c = 2.95$. The sampling distribution has also been obtained by 800 bootstrap samples. The resulting 90% confidence interval is [2.75, 3.05].

Comment on method



(a) *Histogram of events computed for MAXC estimation method. Dashed line at* $\hat{m}_c = 2.95$.

(b) *Distribution of* $\hat{m}_c$ *as estimated by the MAXC method. Sampling distribution estimated by 800 bootstrap replications.*

Figure 9.3: *Graphical representation of MAXC estimation method*

1. The MAXC method offers a fast alternative to estimate the magnitude of completeness, as well as offering the advantage of not being computationally intensive.

2. The reliance of the method on magnitude binning puts emphasis on adequate sample size to ensure that no empty bins are present at the lower magnitude range.

3. Partitioning of magnitude bins (bin width and location) heavily influences the value of estimation results due to the definition of the magnitude of completeness in the MAXC method. Therefore, the estimated value can merely be seen as a robust approximation. MAXC results can be used to validate other, more rigorous estimation methods, or as a starting value for other iterative schemes.

9.3 m_c by b-value stability (bVS)

As described by Cao and Gao this threshold estimation method [8] attempts to determine the magnitude of completeness by analysing the variation of the exponential parameter. The basis of this method is by mapping b-values $\left(b = \frac{\beta}{\ln 10}\right)$ for various subsets S_i of the original dataset S. These subsets are, by construction, dependent on a lower cut-off magnitude m_i^b ($m_{min} \leq m_i^b \leq m_{max}$). Subsequently, the variability of the b-value estimates are investigated.

Assumptions made

1. An exponential distribution is assumed to perfectly describe the FMD of events larger than the magnitude of completeness.
2. In a single sub-sample S_i with cut-off magnitude m_i^b the following is assumed to hold
   - For $m_i^b \leq m_c$: Due to the incomplete portion of data, b-values typically ascend for progressively higher cut-off magnitudes. This is owing to a gradually decreasing fraction of the incomplete portion being included in the studied earthquake catalogue.
   - For $m_i^b \geq m_c$: Due to the exclusion of the incomplete portion of earthquake catalogues, b-values will remain constant for progressively higher cut-off magnitudes.
   - For $m_i^b \gg m_c$: b-value estimates will ascend again for increasing magnitude cut-off values.

   In this manner the b-value of a sample can be thought of as a function of the cut-off magnitude, i.e. $\hat{b}(m_i^b)$.

Since the data in this study is generated from a doubly truncated magnitude distribution as seen in (8.2) (p. 23) the Aki-Utsu b-Value estimator, as used by Cao and Gao, will not be used. For this study, the invariance property of MLE's, together with the maximum likelihood estimator for β which satisfies the following relation, see Page [43], will be used :

$$\frac{1}{\hat{\beta}} = \bar{m} - m_{min} + \frac{(m_{max} - m_{min})}{e^{\hat{\beta}(m_{max} - m_{min})} - 1} \tag{9.2}$$

Outline of method

1. **Identifying data.**
   Subsets, S_i ($i = 1, 2, 3, \ldots, k$), of the original event dataset S must be identified.
2. **Estimating b-values for various cut-off magnitudes.**
   For each subset S_i b-value's are calculated by utilizing the the Page relation, shown in (9.2), and numerical methods.

3. **Measure variability of b-values between cut-off dependent sub-samples.**

   Cao and Gao initially expressed b-value variability ($\Delta\hat{b}(m_i^b)$) as the difference in sample estimates of successive magnitude cut-off shifts, i.e. $\Delta\hat{b}(m_i^b) = \hat{b}(m_{i+1}^b) - \hat{b}(m_i^b)$. Woessner and Wiemer in their 2005 comparison of different threshold estimation techniques [60] have extended this notion by expressing the variability as the difference between individual b-value sample estimates and the arithmetic average of these estimates over a certain range (Δm). In their study Woessner and Wiemer calculated the arithmetic average over half a magnitude range, where the binning width (Δm) is 0.1. Therefore a variability estimate is obtained over a smoothed 5 magnitude bin range, i.e.

$$\Delta\hat{b}(m_i^b) = |\hat{b}(m_i^b) - \bar{b}(m_i^b)| \text{ where } \bar{b}(m_i^b) = \frac{1}{5}\sum_{j=0}^{4}\hat{b}(m_{i+j}^b) \tag{9.3}$$

4. **Obtaining estimate of m_c.**

   Cao and Gao originally stipulated that m_c is the magnitude for which the change in b-value between two successive magnitude cut-off sub-samples, is smaller than 0.03. That is $\hat{m}_c = \inf\{m_i^b \in S : \Delta\hat{b}(m_i^b) < 0.03\}$. However, Woessner and Wiemer [60] have claimed that this definition is unstable and therefore offer an improvement, where it has been suggested that the magnitude of completeness be estimated as

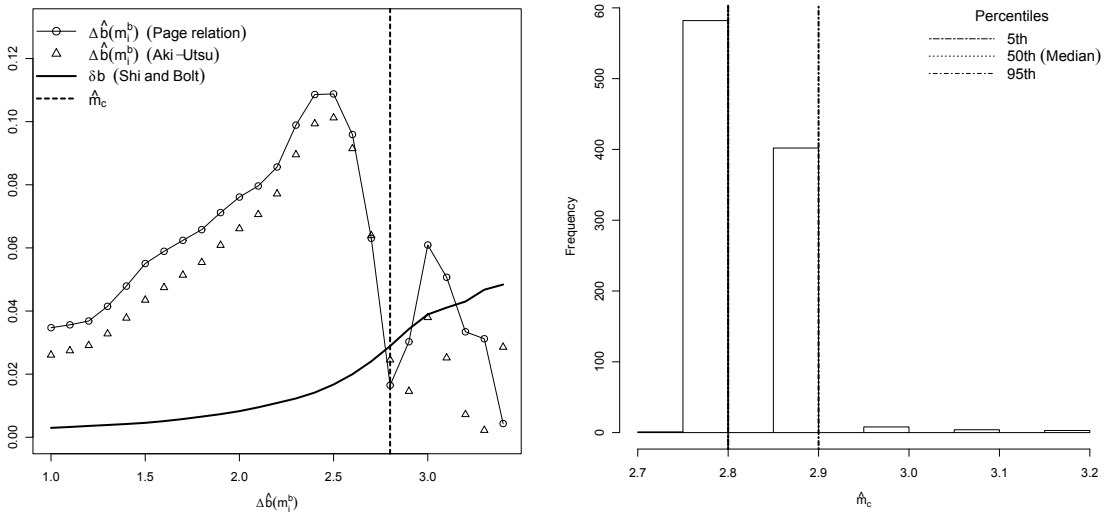
$$\hat{m}_c = \inf\{m_i^b \in S : \Delta\hat{b}(m_i^b) < \delta b\} \text{ where } \delta b = 2.3(\hat{b}(m_i^b))^2\sqrt{\frac{\sum_{j=1}^{N}(m_j - \bar{m})^2}{n(n-1)}} \tag{9.4}$$

   Here δb can be seen to be the b-value uncertainty [50].

Example of estimation algorithm

To demonstrate the method, the illustrative dataset, as described at the beginning of the chapter, will be examined by estimating m_c as well as sampling distribution. The R implementation of the bVS estimation method can be found in Appendix I.1.6 on page 208.

The following is a graphical representation of the m_c determination process. Although the process is based on the difference in b-values based on the Page relation, the b-value differences, based on the Aki-Utsu MLE estimator is also plotted in aid of the demonstration. It can therefore be seen that although the difference between consecutive b-values is smaller for the Aki-Utsu estimator, the significant change in differences occur for both estimators in the vicinity of the true threshold of completeness. Based on the initial dataset, m_c is slightly underestimated to be 2.9, with a 90% confidence interval of [2.8, 3] based on the quantiles of the sampling distribution.



(a) *Graphical representation of detection function*

(b) *Sampling distribution for estimated m_c based on 1000 bootstrap re-samples.*

Figure 9.4: Threshold estimation and estimate uncertainties of bVS method.

Comment on method

Woessner and Wiemer state that choosing Δm differently can have a large impact on the results for $\hat{m}_c$. This can be seen to hold true for two different reasons. The first being that the range of values $\hat{m}_c$ can take on is directly altered and the estimates may not be as refined as possible when larger magnitude bins are used. Furthermore, where larger binning widths are used, the b-value calculated by $\Delta\hat{b}(m_i^b)$ will be averaged over a smaller number of magnitude bins, if the averaging range is not extended. This can lead to instabilities in the method as originally described by Woessner and Wiemer [60].

9.4 Entire magnitude range method (EMR)

This threshold estimation method of Woessner and Wiemer, as introduced in their comparison study of estimation methods [60], attempts to incorporate available data from the entire magnitude range (EMR). The method aims to comprehensively address the estimation problem through explicit modelling of the incomplete portion of event size data. The magnitude of completeness is taken to be threshold value that results in the best fit of the complete as well as incomplete portion of the model to the data.

As will be seen in point 4 in the listed assumptions below, a critical assumption is altered of the EMR method. This change cannot be ignored and therefore, the method implemented in this study can be referred to as either the EMR method or the Modified EMR method (MEMR).

Assumptions made

1. The complete portion of the data is assumed to be adequately modelled by an exponential law (similar to that of the Gutenberg-Richter relation).

2. The incomplete portion of the catalogue can be modelled by applying a detection probability to the exponential distribution used to model the complete portion.

3. The probability of detecting an event of magnitude M, $q(M|\mu,\sigma)$, used to model the incomplete portion of the data is the CDF of a $\mathcal{N}(\mu,\sigma^2)$ random variable. This assumption is in line with that made by other authors [42]. In this formulation the parameter μ indicates the magnitude at which 50% of the events are detected and the variability in the detection range is further defined by the standard deviation σ.

4. Woessner and Wiemer directly attempt to model the incomplete portion of the magnitude range with the detection probability. However, the EMR method will be slightly modified in order to enforce consistency with the method of deriving detected magnitude distributions in Section 8.2.1. Furthermore, the detection probability has been scaled such that the value of 1 is attained at m_c. This is done to ensure that the probability density of the detected magnitude does not suffer from any discontinuities at m_c. Hence, the detected magnitude distribution takes the following form:

$$f_{M_D}(m)=\begin{cases} f_{M_G}(m)\cdot\frac{q(m|\mu,\sigma)}{q(m_c|\mu,\sigma)} & \text{if } m_{min}\leq m<m_c \\ f_{M_G}(m) & \text{if } m_c\leq m\leq m_{max} \\ 0 & \text{otherwise} \end{cases} \tag{9.5}$$

   As per (8.25) on page 29, in the modified version of the EMR method, $q(m|\mu,\sigma)$ in (9.5) is equivalent to $\Phi(z_m)$ where $z_m=\frac{m-\mu}{\sigma}$. This results in the equivalence of (9.5) and (8.25) in their entirety.

Outline of method

1. **Specify range of $\hat{m}_c$**

   The range of values that $\hat{m}_c$ can assume is the values in the sequence f, where $f = (m_i^b)$.

2. **Estimation of nuisance parameters**

   For a given event size dataset S and for each possible value of the magnitude of completeness in the sequence f the parameter vector $\hat{\boldsymbol{\Theta}} = (\hat{\beta}, \hat{\mu}, \hat{\sigma})$ is estimated. A maximum likelihood approach is utilized for this estimation.

$$\hat{\boldsymbol{\Theta}} = \underset{\boldsymbol{\Theta}}{\operatorname{argmax}}\, l(\boldsymbol{\Theta}; m_c = m_i^b | m_1, m_2, \ldots, m_n) \tag{9.6}$$

   where l is the log-likelihood function obtained from the likelihood function L in the following manner :

$$\begin{aligned}
L(\boldsymbol{\Theta}; m_c = m_i^b | m_1, m_2, \ldots, m_n) &= \prod_{i=1}^{n} f_{M_D}(m) = \left(\prod_{m_i < m_c} f_{M_D}(m_i) \right) \left(\prod_{m_i \geq m_c} f_{M_D}(m_i) \right) \\
l(\boldsymbol{\Theta}; m_c = m_i^b | m_1, m_2, \ldots, m_n) &= \ln L(\boldsymbol{\Theta}; m_c = m_i^b | m_1, m_2, \ldots, m_n) \\
&= \sum_{m_i < m_c} \ln f_{M_D}(m_i) + \sum_{m_i \geq m_c} \ln f_{M_D}(m_i) \\
&= \sum_{m_i < m_c} \ln f_{M_G}(m_i) \frac{\Phi(z_m)}{\Phi(z_{m_c})} + \sum_{m_i \geq m_c} \ln f_{M_G}(m_i) + n \ln C_{Norm} \\
&= \sum_{m_i < m_c} (\ln \Phi(z_m) - \ln \Phi(z_{m_c})) + \sum_{i=1}^{n} \ln f_{M_G}(m_i) + n \ln C_{Norm} \\
&= \sum_{m_i < m_c} (\ln \Phi(z_m) - \ln \Phi(z_{m_c})) + n \ln \beta - \beta \sum_{i=1}^{n} (m_i - m_{min}) \\
&\quad - n \ln \left(1 - e^{-\beta(m_{max} - m_{min})}\right) + n \ln C_{Norm}
\end{aligned} \tag{9.7}$$

   As shown $f_{M_D}(m)$ is as in (8.25) on page 29. Partial derivatives of the log-likelihood function are derived in Appendix F (p. 196).

3. **Model comparison and estimating m_c**

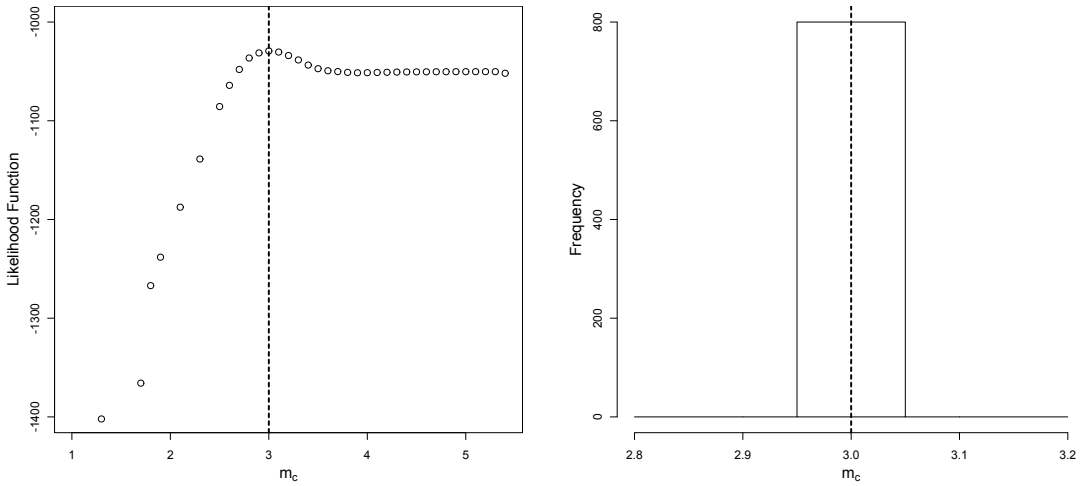
   The threshold value in the specific model that maximizes the likelihood test score for all m_i^b in sequence f is taken as the magnitude of completeness for the set of events S, i.e.

$$\hat{m}_c = \underset{m_i^b}{\operatorname{argmax}} \left(\underset{\boldsymbol{\Theta}}{\operatorname{argmax}}\, l(\boldsymbol{\Theta}; m_c = m_i^b | m_1, m_2, \ldots, m_n) \right) \tag{9.8}$$

Example of estimation algorithm

The R implementation of the MEMR estimation method can be found in Appendix I.1.7 on page 210. The following is a graphical representation of the m_c estimation process. Since the initial dataset has been generated from a distribution where the detection probability is modelled by a Normal CDF, the

MEMR method yields ideal results. These ideal conditions, also culminate in the sampling distribution of the estimator being concentrated in one point, namely the true value of $m_c = 3$.



(a) *Graphical representation of detection function* **(b)** *Sampling distribution for estimated m_c based on 1000 bootstrap re-samples.*

Figure 9.5: *Threshold estimation and estimate uncertainties as per the MEMR method*

Comment on method

Woessner and Wiemer have tested four functions (CDF of exponential, Log-Normal and Normal random variable, as well as an exponential decay function) to fit the incomplete part of actual earthquake datasets. Furthermore, they state that the the exponential decay and Normal CDF perform relatively well in comparison to the other functions. However the Normal CDF provides the best fit to the data, when the considered data ranges from regional to worldwide sources. For this reason the EMR method (as well as the MEMR method) utilizes the Normal CDF for modelling detection probabilities.

9.5 Median Based Assessment of the Segment Slope (MBASS)

The basis of the MBASS method, by D. Amorèse [3], is to estimate significant magnitude breaks by making use of a non-parametric approach adapted from the analysis of climate data [33]. Both approaches are based on the change-point test of Siegel and Castellan [51, p. 399], which seeks to establishes the statistical significance of a difference between the medians of two samples.

The magnitude of completeness is defined as the smallest magnitude point where a statistically significant break in the logarithmic slope of the frequency-magnitude distribution occurs.

Assumptions made

1. The assumption is made that the event sizes larger than the magnitude of completeness are exponentially distributed. This is not explicitly stated, but can be seen through examination of the slope value $r(m_i^b)$ as defined by (9.9) in the outline of the method below. For further elaboration on this point, see the general comment below.

2. The distribution of event sizes smaller than the magnitude of completeness is not specified, however it is assumed that this data does not follow the same distribution as that specified for the data with magnitudes greater than m_c.

Outline of method

1. **Construct histogram of event magnitudes.**
   For a given set of data S, the incremental histogram of observed events $H(S, P_O(S))$ is computed, with magnitude bins $i = 1, 2, \ldots, k$. Cumulative FMDs are not computed since successive magnitude bins would be dependent on the information contained in preceding bins. This violates the assumption of independence within groups that is required by the Wilcoxon-Mann-Whitney U test [27].

2. **Calculate piecewise slope of FMD.**
   For each magnitude segment, defined by the endpoints of the relevant bin $[m_{i-1}^b, m_i^b]$, the slope of the incremental FMD is calculated as

$$r(m_i^b) = \frac{\ln H(S, P_O(S))_{i-1} - \ln H(S, P_O(S))_i}{m_{i-1}^b - m_i^b} \qquad \text{for } i = 2, 3, \ldots, k \tag{9.9}$$

3. **Identify position of most probable change in FMD slope.**
   Ranks are assigned to the slopes $r(m_i^b)$ from 1 to k, where ties receive a rank equal to the average of the ranks they span. Thereafter, for each $i = 2, 3, \ldots, k$ in the sequence of magnitude bins,

the sum of the ranks $(SR)_i$ of the segment slopes, $r(m_j^b)\ \forall j \leq i$, is computed. This value is then adjusted as follows

$$(SA)_i = |2(SR)_i - i(k+1)| \tag{9.10}$$

The reasoning of the adjustment stems from the Wilcoxon-Mann-Whitney U-test, which has test statistic $U = (SR)_i - \frac{i(i+1)}{2}$, with expected value $\frac{i(k-i)}{2}$ for $k \in \mathbb{N}$ observations [41]. For varying i, each statistic U is centred around a different magnitude. To aid comparison of this statistic for a range of possible magnitude breaks, the expected value at each i is subtracted from U. This yields the statistic $(SR)_i - \frac{i(k+1)}{2}$, which can be multiplied by 2 for display purposes. Finally, by taking the absolute value of the statistic it is ensured that the absolute physical shift in process is considered.

The magnitude bin which yields the largest value of $(SA)_i$ for all $i = 2, 3, \ldots, k$ must be chosen, i.e.

$$i^* = \underset{i}{\operatorname{argmax}}\,(SA)_i \tag{9.11}$$

Magnitude $m_{i^*}^b$ divides the range $[m_1^b, m_{k+1}^b]$ into two intervals, namely $I_1 = [m_1^b, m_{i^*}^b]$ and $I_2 = [m_{i^*+1}^b, m_k^b]$. Subsequently the event size data in S can be split into two sets $S_1 = \{m_i \in S : m_i \in I_1\}$ and $S_2 = \{m_i \in S : m_i \in I_2\}$, with $S_1 \cup S_2 = S$. Hence, the respective median value of segment slopes of subsets S_1 and S_2 show the greatest promise of being statistically different.

4. **Assess statistical significance of magnitude break.**
   The sets S_1 and S_2 will respectfully have $k_1 = i^*$ and $k_2 = k - k_1$ numbers of elements. Continuing, the statistic $W = (SR)_{k_1}$ is defined.

   A hypothesis test is done to determine whether the segment slopes of the two subsets S_1 and S_2 are significantly different. If so, this change will be indicated at the point M_{k_1}. The Wilcoxon-Mann-Whitney U-test is performed at a suitably chosen significance level in order to determine if a difference is present. If a statistically significant magnitude break is identified, then this magnitude is taken to be the magnitude of completeness, i.e. $\hat{m}_c = m_{i^*}^b$. Binning width can be taken into account by having $\hat{m}_c = m_{i^*}^b + \frac{\Delta M}{2}$. This modification assumes that event sizes are uniformly distributed over each magnitude bin.

5. **Continuing search for additional magnitude breaks**
   If in the previous step it is found that a statistically significant difference in median FMD slopes of the subsets S_1 and S_2 exist, then this effect can be compensated for in order to continue searching for further statistically significant magnitude breaks. After subtracting the respective median FMD

slope values from S_1 and S_2, a further iteration of the above process can be carried out, until no more statistically significant magnitude breaks can be identified.

Example of estimation algorithm

The R implementation, as published by Amorèse has been used for the estimation of threshold values. This can be found in the authors original paper detailing the method [3].

A binning width (Δm) of 0.1 has been used. As in the study of Amorèse, a 10% level of significance is specified for the Wilcoxon-Mann-Whitney U-test. Figure 9.6 illustrates how the magnitude bins can be divided into two distinct groups based on the slope over the bins. For this dataset, the MBASS estimate of the detection threshold is $\hat{m}_c = 2.8$ and, after application of the binning width correction, an adjusted value of $\hat{m}_c = 2.8 + \frac{\Delta m}{2} = 2.85$ is obtained. A bootstrap implementation with 800 resamples yielded the sampling distribution, for this specific dataset, as a point distribution centred at $\hat{m}_c$.

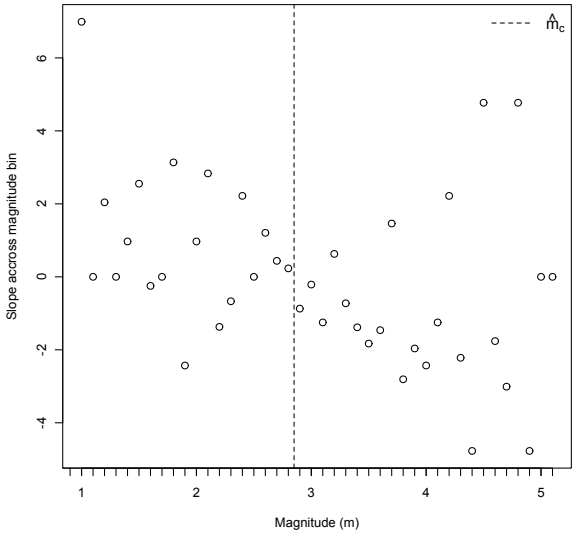


Figure 9.6: *Graphical representation of the MBASS estimation method: Plot of slopes over magnitude bins. Refernce line has been added at $\hat{m}_c$.*

Comment on method

The assumption of exponentially distributed data can arrived at as follows. For $i = 2, 3, \ldots, k$:

$$\beta = \frac{\ln f_{M_G}(m_{i-1}) - \ln f_{M_G}(m_i)}{m^b_{i-1} - m^b_i} \approx \frac{\ln H(S, P_O(S))_{i-1} - \ln H(S, P_O(S))_i}{m^b_{i-1} - m^b_i} \tag{9.12}$$

$$= r(m^b_i) \tag{9.13}$$

Since the PDF $f_{M_G}(m)$ is of the form $f_{M_G}(m) \propto C_{Norm} \cdot e^{-\beta m}$ (where C_{Norm} is the normalizing constant of the relevant exponential distribution) the above relation also states that MBASS method disregards whether the random variable M_G is truncated or not. This unmentioned strength, makes threshold estimation by use of the MBASS method extremely robust.

The reliance of the method on magnitude binning puts emphasis on adequate sample size to ensure that no empty bins are present at the lower magnitude range.

As a hypothesis test is implemented in estimating the magnitude of completeness, the possibility exists that the null hypothesis will not be rejected. The only conclusion then reached is that, based on the assessment of segment slopes, sufficient evidence has not been found to state that the event data S can be divided into complete and incomplete subsets. In such a situation, the magnitude where a threshold is most likely to exist ($m_{i^*}^b$) can be used to further an investigation on detection probabilities. Otherwise the sample can be regarded as statistically complete. For this reason the current investigation establishes $\hat{m}_c = m_{i^*}^b$.

9.6 Moment incorporating threshold computation (MITC)

In this section a new threshold estimation scheme, termed MITC, is developed. As an introduction, a brief summary of the assumptions underlying the method is given.

Assumptions made

1. The complete portion of the catalogue is assumed to be adequately modelled by an exponential law (similar to that of the Gutenberg-Richter relation).
2. The incomplete portion of the catalogue can be modelled by applying a detection probability to the exponential distribution used to model the complete portion.
3. The functional form of the probability to detect an event of magnitude M must be explicitly stated. An example of such a function, as well as that used throughout this investigation, is the CDF of a $\mathcal{N}(\mu, \sigma^2)$ random variable. This assumption is in line with that made by Ogata and Katsura [42]. Woessner and Wiemer [60] incorporate this detection probability in a similar manner, but directly apply it to the whole of the incomplete magnitude range.
4. In order to avoid a discontinuity in the probability density of the detected magnitude, this detection probability has been normalized such that the value of 1 is attained at m_c. Therefore the detected magnitude distribution takes the following form:

$$f_{M_D}(m) = \begin{cases} f_{M_G}(m) \cdot \frac{\Phi(z_m)}{\Phi(z_{m_c})} & \text{if } m \in [m_{min}\,;\, m_c) \\ f_{M_G}(m) & \text{if } m \in [m_c\,;\, m_{max}] \\ 0 & \text{otherwise} \end{cases} \tag{9.14}$$

Derivation of threshold estimation scheme

Incorporating the technique of integration by parts, the expected value of the detected magnitude distribution in (9.14) can be obtained as follows :

$$\begin{aligned} E[M_D] &= \int_{m_{min}}^{m_c} m f_{M_D}(m)\, dm + \int_{m_c}^{m_{max}} m f_{M_D}(m)\, dm \\ &= F_{M_D}(m) m \Big|_{m_{min}}^{m_c} - \int_{m_{min}}^{m_c} F_{M_D}(m)\, dm + \int_{m_c}^{m_{max}} m f_{M_D}(m)\, dm \\ &= F_{M_D}(m_c)\, m_c - \int_{m_{min}}^{m_c} F_{M_D}(m)\, dm + E[M_D \big| M_D > m_c] P[M_D > m_c] \end{aligned} \tag{9.15}$$

After rearranging and substitution of the population moments with sample moments in (9.15) the following is obtained

$$0 = \bar{m}_{[m_{min}, m_{max}]} + \int_{m_{min}}^{m_c} F_{M_D}(m)\, dm - \bar{m}_{[m_c, m_{max}]} P[M_D > m_c] - m_c\, F_{M_D}(m_c) \tag{9.16}$$

where $\bar{m}_S = \frac{1}{\sum I_S(m_i)} \sum_{m_i \in S} m_i$ is the average over the elements in the set S and $I_S(x)$ the indicator function. Based on (9.16) a new function $g()$ can be defined. In the function $g()$, the probability $P[M_D > m_c]$ has been expressed in terms of the distribution function of the random variable M_D as $(1 - F_{M_D}(m_c))$.

$$\begin{aligned} & g(\, m_c \,, m_{min}, m_{max}, \mathbf{\Psi}) \\ & \quad := \bar{m}_{[m_{min}, m_{max}]} + \int_{m_{min}}^{m_c} F_{M_D}(m;\ m_c, m_{min}, m_{max}, \mathbf{\Psi})\, dm \\ & \qquad - \bar{m}_{[m_c, m_{max}]}(1 - F_{M_D}(m_c; m_c, m_{min}, m_{max}, \mathbf{\Psi})) - m_c\, F_{M_D}(m_c; m_c, m_{min}, m_{max}, \mathbf{\Psi}) \end{aligned} \tag{9.17}$$

where $\mathbf{\Psi}$ is a vector consisting of nuisance parameters of the detected magnitude distribution. Furthermore, the distribution function $F_{M_D}(m)$ in (9.17) is explicitly stated as being variable with respect to the quantities m_c, m_{min}, m_{max} and $\mathbf{\Psi}$. Hence, all parameters of the detected magnitude distribution must be estimated. In this paper such estimation is accomplished by the method of maximum likelihood. Therefore, zeros are sought for :

$$g(m_c, m_{min}, m_{max}, \mathbf{\Psi} = \underset{\mathbf{\Theta}}{\operatorname{argmax}}\, l_{M_D}(\mathbf{\Theta}; m_c | m_1, m_2, \ldots, m_n)) \tag{9.18}$$

$l_{M_D}(\mathbf{\Theta}; m_c | m_1, m_2, \ldots, m_n)$ is the log-likelihood function of the detected magnitude distribution. It is fairly easy to envision the possibility of (9.18) having multiple roots. For this investigation it will be assumed that the first root of the equation corresponds to the sought threshold, however it is proposed that this assumption be scrutinized in further studies.

Utilizing (9.18) $\hat{m}_c$ can be obtained through a variety of numerical methods, the simplest of which is by defining a fixed point iterative scheme:

$$(\hat{m}_c)_{i+1} = \frac{1}{F_{M_D}((\hat{m}_c)_i)} \left(\bar{m}_{[m_{min}, m_{max}]} + \int_{m_{min}}^{(\hat{m}_c)_i} F_{M_D}(m)\, dm - \bar{m}_{[(\hat{m}_c)_i, m_{max}]} P[M_D > (\hat{m}_c)_i] \right) \tag{9.19}$$

For this implementation it is proposed that the MAXC estimate of the threshold [59] be used as the starting value, and a stopping criterion of $|(\hat{m}_c)_{i+1} - (\hat{m}_c)_i| < \epsilon = 0.001$. However, after a number of trials it has been ascertained that the fixed-point method does produce threshold estimates, but is highly dependent on the starting value and requires a substantial number of iterations to achieve accurate results.

Outline of proposed method

To ensure the robustness of the threshold estimation method the following algorithm is proposed:

1. **Isolate the first root of** (9.18)

   This is achieved by systematically evaluating the equation for values of m_c contained in the set

$\{m_i = m + i \cdot \Delta m : i = 0, 1, 2, \ldots, [\frac{m_{max} - m_{min}}{\Delta m}]\}$ stopping when $\text{sgn}\left(g(m_i, m_{min}, m_{max}, \mathbf{\Psi})\right) \neq \text{sgn}\left(g(m_i + \Delta m, m_{min}, m_{max}, \mathbf{\Psi})\right)$ for the smallest value of i. If the distribution of detected events is constructed to be a proper probability distribution, the mild assumption can be made that (9.18) is continuous. By the intermediate value theorem it can then be seen that a root for the equation must lie in the interval

$$[m_{i^*}\,;\, m_{i^*} + \Delta m]$$
$$\text{where } i^* = \inf\left(i : \text{sgn}\left(g(m_i, m_{min}, m_{max}, \mathbf{\Psi})\right) \neq \text{sgn}\left(g(m_i + \Delta m, m_{min}, m_{max}, \mathbf{\Psi})\right)\right) \quad (9.20)$$

Due to the estimation of nuisance parameters ($\mathbf{\Psi}$) being dependent upon the value of m_c it is possible for discontinuities to arise in the function $g()$, however this is beyond the current investigation and left for future endeavours.

2. **Utilize a robust solving algorithm to estimate the threshold.**
   For the current investigation the R implementation of Brent's method [6] will be used to solve the equation for the root in the interval $[m_{i^*}\,;\, m_{i^*} + \Delta m]$ with i^* as determined in the previous step.

3. **Obtain standard error**
   It is suggested that uncertainties in the estimated threshold be obtained through bootstrapping. Additionally, uncertainties in the estimates of the nuisance parameters can be obtained utilizing the observed Fisher information matrix, if maximum likelihood is used.

It should be noted that this method of threshold estimation can be interpreted as a modified form of moment estimation.

For the current investigation the detection probability of the detected magnitude distribution will only be modelled by the Normal CDF. The validity of modelling the detection probability by use of this function is left to be established through future work.

Example of estimation algorithm

The R implementation of the MITC estimation method can be found in Appendix I.1.8 on page 211. The illustrative dataset, as described at the beginning of the chapter, will be used for estimating m_c as well as the sampling distribution.

Figure 9.7 is a graphical representation of the (9.18) for which the root is sought in order to estimate m_c. Based on the specific dataset, m_c is slightly overestimated to be 3.1, with a 90% confidence interval of [2.97 3.21] based on the quantiles of the sampling distribution.

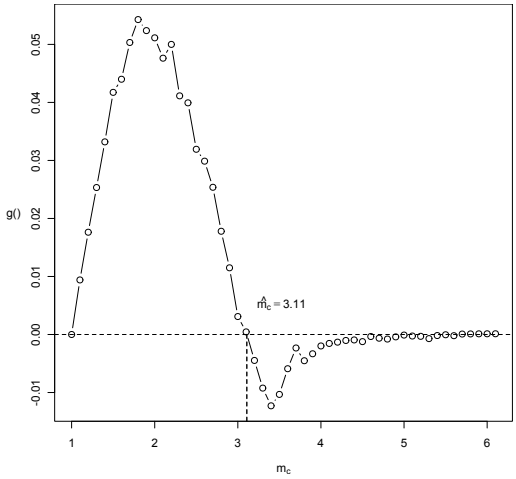
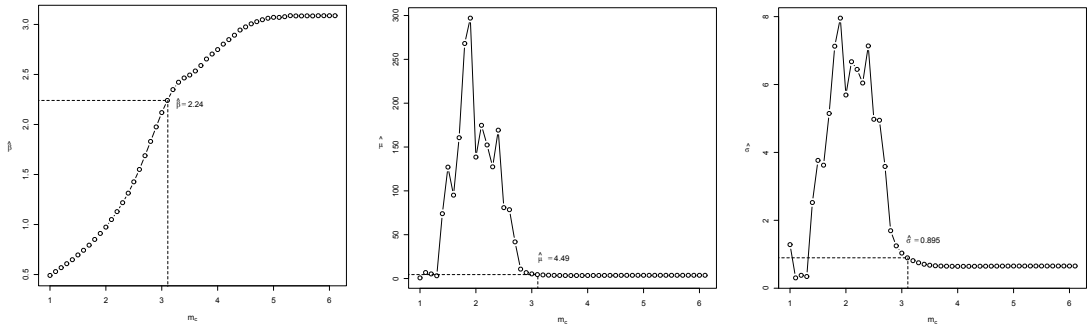


Figure 9.7: *Sketch of (9.18) that must be solved for 0 in order to estimate threshold of detection m_c as well as nuisance parameters of detected severity distribution.*



(a) *Progression of $\hat{\beta}$ estimated via ML For differing values of m_c*

(b) *Progression of $\hat{\mu}$ estimated via ML For differing values of m_c.*

(c) *Progression of $\hat{\sigma}$ estimated via ML For differing values of m_c.*

Figure 9.8: *Maximum likelihood estimation of nuisance parameters of observed magnitude distribution for varying detection threshold values (m_c). Dotted line indicates the value of $\hat{m}_c$ as obtained via MITC.*

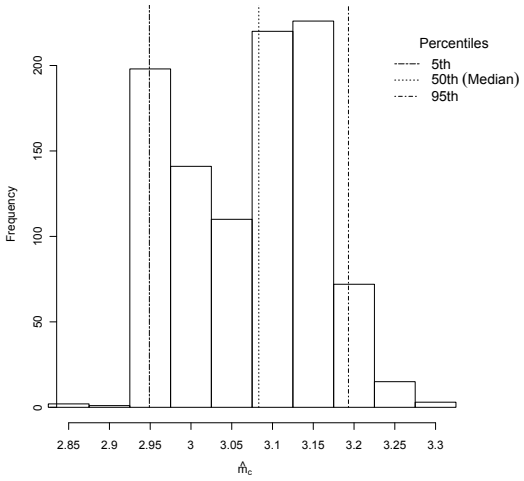


Figure 9.9: *Histogram of sampling distribution approximated through bootstrapping with 1000 repetitions of threshold estimation.*

10 Performance evaluation of estimation methods

In this section the efficiency of the previously discussed threshold estimation methods will be gauged under varying circumstances. Artificial catalogues are compiled with observations being generated from detected magnitude distributions from Section 8.2.3 (p.27) up to and including Section 8.2.6 (p.44).

Comparisons will be made based on the mean-squared error (MSE) that can be shown to be composed of the bias of the estimator, as well as the variance of the estimator.

The results will be summarized according to

1. **Magnitude of completeness**
   The characterization of the magnitude of completeness, as either sharp or soft will form a natural division for data analysis. For the case where the magnitude of completeness is characterized as sharp, i.e. being explicitly included, the level at which it has been specified will further categorize the data.

2. **Estimation Method**
   Fundamental to the current investigation is the specific method that has been implemented in order to estimate the magnitude of completeness. This will directly provide information relating to the efficacy of the relevant methods.

3. **Distribution of seismic detection probability**
   The different forms of the probability describing the detection of events further follows as a natural way of categorizing the data.

4. **Moments of detection distribution**
   The distribution characterizing the probability of seismic event detection will further be categorized according to the lower order moments of the distribution. Factors will be created based on non-overlapping and exhaustive intervals of the range of the observed moment values. These moments include the Expected Value, Variance and Third null-point moment of the detection distribution.

Furthermore, to aid in the objective comparison of methods, the parameter region for the simulated catalogues must be specified. This parameter region is set up in the following section, by enforcing consistency between in the values of the lower order moments of the detection distributions.

10.1 Parameter domain of simulated catalogues

10.1.1 General Parameter Region

The complete portion, and therefore the parameter values of the original Gutenberg-Richter Distribution, of all datasets are simulated according to the following fixed parameters : $m_{min} = 1$, $m_{max} = 7$ and $b = \dfrac{\beta}{\ln(10)} = 0.9$. The incomplete portion of the respective datasets are characterized by varying the parameters of the detection distribution. Let $\boldsymbol{\Theta} = (\theta_1, \theta_2, \dots \theta_n)$ represent the vector of parameters of each respective detection distribution. In order to consistently compare estimation results, under varying forms of the detection distribution, the set of parameters to be examined for each detection probability are specified based on restrictions of the lower order moments of the detection distribution. Within this study n is 3 for the cumulative Pareto type II distribution and 2 for the remaining detection distributions. For the case where $n = 2$ the expected value and variance of the detection distribution will be considered, whereas for $n = 3$, the third moment about the origin will be additionally be considered.

The motivation for considering the third moment about the origin is that this moment contains information about the expected value, variance and skewness of the distribution. The range of the the third moment about the origin will be determined for the distributions where $n = 2$, which will be modified to represent an interval of conceivable values. This interval will act as restriction on the third moment about the origin for case of $n = 3$, which will ensure that the different detection distributions will be of comparable nature.

Let X_D represent the random variable to which the particular CDF is attributed that is used to describe the detection probability (i.e. detection distribution). The lower order moments of the detection distribution, as functions of the distributional parameters contained in $\boldsymbol{\Theta}$, can be written as

$$h_1(\boldsymbol{\Theta}) = E[X_D | M_G; \boldsymbol{\Theta}] \quad \text{and} \quad h_2(\boldsymbol{\Theta}) = Var(X_D | M_G; \boldsymbol{\Theta}) \tag{10.1}$$

and for the case where $n = 3$:

$$h_3(\boldsymbol{\Theta}) = E\left[X_D^3 | M_G; \boldsymbol{\Theta}\right] \tag{10.2}$$

which can be written as the vector valued function $\boldsymbol{H(\Theta)} = (h_1(\boldsymbol{\Theta}), h_2(\boldsymbol{\Theta}), \dots, h_n(\boldsymbol{\Theta}))$. The admissible range of expected values and variances of a specific detection distribution can be written as the following

$$C(I_1, I_2, \dots I_n) = \{\boldsymbol{H(\Theta)} \in \mathbb{R}^n : h_1(\boldsymbol{\Theta}) \in I_1,\ h_2(\boldsymbol{\Theta}) \in I_2,\ \dots h_n(\boldsymbol{\Theta}) \in I_n\} \tag{10.3}$$

where $I_1, I_2, \dots I_n$ are intervals on the real line. From where the set of all permissible parameters can be obtained through the inverse image of $\boldsymbol{H}$

$$\boldsymbol{H^{-1}}(C(I_1, I_2, \dots I_n)) = \{\Theta \in \mathbb{R}^n : (h_1(\boldsymbol{\Theta}), h_2(\boldsymbol{\Theta}), \dots h_n(\boldsymbol{\Theta})) \in C(I_1, I_2, \dots I_n)\} \tag{10.4}$$

Empirically, it has been found that detected magnitude distributions based on $I_1 = [0.2, 6]$, $I_2 = [0.2, 4]$ (for $n = 2$) and, additionally, $I_3 = [0.2, 300]$ (for $n = 3$) produce datasets comparable to real-world earthquake data.

10.1.2 Partitioning of parameter region

Some lower order moments are non-linear functions of the parameters and therefore exhibit either concave or convex behaviour. This has the implication that if the investigated values of $\theta_1, \theta_2, \ldots, \theta_n$ are equi-spaced over the set of $\boldsymbol{H^{-1}}(C(I_1, I_2, \ldots, I_n))$ the resulting variances and expectations of the detection distribution will not be evenly spaced. Therefore, to ensure a fair representation of parameter induced forms of the distribution, rather than explicitly partitioning the parameter space $\boldsymbol{H^{-1}}(C(I_1, I_2, I_n))$, the region $C(I_1, I_2, I_n)$ will be evenly partitioned.

It is designed that the number of elements $(x, y) \in C(I_1, I_2)$ distributed over I_1 must be proportional to the length of the interval, with the number of elements (x, y) distributed over I_2 specified in the same manner. Therefore, if T is the total number of elements in the set $C(I_1, I_2)$ and T_i the number of points (x, y) distributed over the interval I_i and $d_i = |\sup(I_i) - \inf(I_i)|$, for $i = 1, 2$, it can be shown that $T_1 = \sqrt{T\frac{d_1}{d_2}}$ and $T_2 = \frac{d_2}{d_1}T_1$. Accordingly the number of subintervals over I_1 and I_2 are respectively $T_1 - 1$ and $T_2 - 1$. Initially it is aimed that 440 catalogues be generated per m_c category for the sharp-m_c models as well as for the entire investigation under the soft-m_c model.

For the case where $n = 3$ (Pareto type II distribution) the total number of events to be apportioned over I_1 and I_2 is taken to be the total number of points divided by the number of points spread over I_3, which is pre-specified to be 17. This number of points is pre-specified since a proportional allocation (similar as for the case of $n = 2$) would result in a much larger number of points being allocated to this interval. The reason for this is the large difference of d_3 when compared to d_1 and d_2.

After suitable integer rounding at certain calculation stages the following summary of parameter partitioning is obtained.

Table 10.1 provides a summary of the number of catalogues to be simulated.

	I_1	I_2	I_3
Infimum	0.2	0.2	0.2
Supremum	6	4	300
d_i	5.8	3.8	299.8
Number of points in interval	25.9148	16.9786	17
Rounded number of points in interval	26	17	17
Number of subintervals	25	16	16
Length of individual subintervals	0.232	0.2375	17.63

Table 10.1: *Parameters for traditional Gutenberg-Richter seismic event distribution, as used to model complete earthquake catalogue.*

For $n = 2$, a specific earthquake catalogue can now be defined as an element of the set

$$E = \{\boldsymbol{H}(\boldsymbol{\Theta}) = (x, y) : x = 0.2 + 0.232i \text{ and } y = 0.2 + 0.2375j \ ; \\ i = 0, 1, 2, \ldots, 25 \text{ and } j = 0, 1, 2, \ldots, 16\} \tag{10.5}$$

whereas for $n = 3$ the earthquake catalogue can be defined as an element of the modified set E :

$$E = \{\boldsymbol{H}(\boldsymbol{\Theta}) = (x, y, z) : x = 0.2 + 0.232i \ , \ y = 0.2 + 0.2375j \text{ and } z = 0.2 + 17.63k \ ; \\ i = 0, 1, 2, \ldots, 25 \ , \ j = 0, 1, 2, \ldots, 16 \ \text{ and } k = 0, 1, 2, \ldots, 16\} \tag{10.6}$$

In this way it can be seen that $E \subset C(I_1, I_2, \ldots, I_n)$ for $n = 2$ or 3. Furthermore this partitioning leads to 442 catalogues per m_c category for sharp m_c models and 442 catalogues in total for soft m_c models.

10.2 Results for magnitude of completeness for explicit detection threshold

It is readily seen that the threshold estimation methods come from diverse backgrounds. This especially when considering the individual derivations and assumptions. In order to facilitate a consistent comparison between results obtained from the varying methods, a metric must be obtained that can relate to all methods. The Mean-Squared Error (MSE) is chosen as a simple metric. Furthermore, the MSE can be shown to be composed of the variance of the estimator $\hat{\theta}$ and the bias of the estimator when considering the true value of the parameter θ_0.

$$MSE(\hat{\theta}) = E\left[\left(\hat{\theta} - \theta_0\right)^2\right] = Var(\hat{\theta}) + (E[\hat{\theta} - \theta_0])^2 \tag{10.7}$$

$$= Var(\hat{\theta}) + (Bias(\hat{\theta}, \theta_0))^2 \tag{10.8}$$

It can therefore be seen that an ideal estimator has small MSE. This consequently translates into an estimator having a small variance and a small absolute bias. Additionally, a good estimator will have a symmetric sampling distribution.

In the sense of threshold estimation a prudent scenario can be described as an estimator having a positive bias (over-estimating the threshold value). When considering the contrary situation, of underestimating the threshold value, it can be seen that this can lead to the inclusion of inhomogeneous data in the complete portion of the distribution, which will skew any parameter results obtained.

10.2.1 Significance of magnitude of completeness as factor

Table 10.2 shows the p-Value of the Shapiro Wilk test for Normality of the estimated biases. The test is performed individually for groupings of the estimated bias for each of the threshold estimation methods and then also for differing forms of the detection distribution. It can therefore be concluded that, at a 5% level of significance, the estimated biases of the various m_c estimation methods do not follow a Normal distribution. It is for this reason that the investigation be continued based on non-parametric statistical methods.

CDF	GOF		MAXC		bVS		EMR		MBASS		MITC	
	Statistic	p-Value	Statistic	p-Value	Statistic	p-Value	Statistic	p-Value	Statistic	p-Value	Statistic	p-Value
Normal	0.92	<0.001	0.95	<0.001	0.92	<0.001	0.94	<0.001	0.99	<0.001	0.97	<0.001
Log-Normal	0.96	<0.001	0.95	<0.001	0.96	<0.001	0.98	<0.001	0.98	<0.001	0.98	<0.001
Logistic	0.83	<0.001	0.96	<0.001	0.86	<0.001	0.85	<0.001	1.00	<0.001	0.94	<0.001
Pareto type II	0.95	<0.001	0.92	<0.001	0.95	<0.001	0.96	<0.001	0.98	<0.001	0.97	<0.001

Table 10.2: *Test statistic and p-Value of Shapiro-Wilk test for normality of estimated biases. Test individually performed for estimated bias of each estimation method and differing CDF of RV attributed to detection probability.*

CDF	GOF		MAXC		bVS		EMR		MBASS		MITC	
	Statistic	p-Value	Statistic	p-Value	Statistic	p-Value	Statistic	p-Value	Statistic	p-Value	Statistic	p-Value
Normal	0.91	<0.001	0.81	<0.001	0.90	<0.001	0.69	<0.001	0.98	<0.001	0.93	<0.001
Log-Normal	0.93	<0.001	0.80	<0.001	0.88	<0.001	0.55	<0.001	0.99	<0.001	0.92	<0.001
Logistic	0.91	<0.001	0.69	<0.001	0.88	<0.001	0.77	<0.001	0.96	<0.001	0.94	<0.001
Pareto type II	0.91	<0.001	0.92	<0.001	0.90	<0.001	0.63	<0.001	0.97	<0.001	0.88	<0.001

Table 10.3: *Test statistic and p-Value of Shapiro-Wilk test for normality of estimated standard error of m_c estimation method. Test individually performed for estimated standard error of each estimation method and differing CDF of RV attributed to detection probability.*

The statistical significance of the pre-specified magnitude of completeness in estimating the threshold is established by utilization of the R implementation [19] of the multiple comparison Kruskal-Wallis test procedure [51, p. 213]. The results are stated in Table 10.4. Based on the multiple comparison Kruskal-Wallis test at a 5% level of significance the following results have been obtained for the bias estimates of the competing estimation methods:

Pairwise m_c comparison		GOF		MAXC		bVS		EMR		MBASS		MITC	
		OD	CD	OD	CD	OD	CD	OD	CD	OD	CD	OD	CD
2	2.5	908.4	220.3	1331.6	220.3	974.1	220.3	512.2	220.4	1187.2	220.3	1797	219.5
2	3	1899.4	222.4	2638.1	222.4	2116.3	222.4	1105.7	222.6	2827.1	222.4	3168.1	222.7
2	3.5	2982.6	224.0	3904.1	224	3299.3	224	1711.1	224.2	4233.6	224	4387.8	224.2
2	4	4042.3	225.6	5144.9	225.6	4465.2	225.6	2173.9	225.5	5343	225.6	5522.5	223.5
2.5	3	991	223.8	1306.5	223.8	1142.1	223.8	593.5	223.9	1639.8	223.8	1371	224.1
2.5	3.5	2074.1	225.3	2572.4	225.3	2325.1	225.3	1198.9	225.5	3046.3	225.3	2590.7	225.6
2.5	4	3133.9	226.9	3813.2	226.9	3491.0	226.9	1661.7	226.8	4155.8	226.9	3725.4	224.9
3	3.5	1083.1	227.4	1265.9	227.4	1182.9	227.4	605.4	227.6	1406.5	227.4	1219.7	228.7
3	4	2142.8	229.0	2506.7	229.0	2348.8	229	1068.2	228.9	2515.9	229	2354.4	228
3.5	4	1059.7	230.5	1240.7	230.5	1165.9	230.5	462.8	230.4	1109.4	230.5	1134.7	229.5

Table 10.4: *Observed difference (OD) and critical differences (CD) for the multiple comparison Kruskal-Wallis test for estimated biases by pre-specified m_c. Test individually performed for estimation methods.*

It can be seen that no grouping of biases by magnitude of completeness is similar to another. Therefore, it can be concluded at a 5% level of significance, that the pre-specified magnitude of completeness is a statistically significant factor when analysing bias as well as the variance of the competing estimation methods.

This signifies unwanted behaviour, since no correction for the bias can be attempted. This is the case since the level of bias is heavily dependent on the true value of the threshold.

10.2.2 Results of threshold estimation methods incorporating threshold as factor

Analysis of results, with the true threshold as a factor, yields consistently smaller biases for cases with lower valued thresholds. This holds true for all analyses performed by expected value, variance and skewness of the detection distribution. Hence, much of the variation in the aggregated results are artificial and due to the value of the true threshold.

10.2.3 Aggregated results of threshold estimation methods

The situation is examined where data is aggregated solely based on the function used to model the detection probability as well as the threshold estimation method. In this way a high level analysis is performed of general robustness of estimation methods.

The main quantities of the sampling distribution will be illustrated. Instead of directly displaying quantities of the sampling distribution, the figures will be shifted, by respectively subtracting the correct value of m_c, to represent expected biases. Quantities used to characterize the sampling distribution of expected biases include the mean as well as 5% and 95% quantiles, illustrating a 90% confidence interval.

In order to smooth results for the various detection distributions, quantities of sampling distribution data have been grouped by intervals of the moment domains of the detection distributions. In order to construct an image of the skewness of the sampling distribution of expected biases, the average and median of the respective quantities have been calculated.

10.2.3.1 Threshold estimation by Goodness of fit (GOF) method

In order to consider the behaviour of the GOF method, Tables 10.5, 10.6 and 10.7 have been constructed to respectively display Mean-Squared Error (MSE), Bias and Standard Error values.

Grouping event detection distribution by	True threshold value	Detection probability modelled by ...											
		Cum. Normal Dist.			Cum. Logistic Dist.			Cum. Log-Normal Dist.			Cum. Pareto type II Dist.		
		L	M	H	L	M	H	L	M	H	L	M	H
Exp. Value	2	0.6	0.19	0.1	-	0.19	0.14	0.72	0.1	0.05	0.66	0.22	-
	2.5	1.51	0.51	0.15	-	0.44	0.25	1.73	0.23	0.09	1.68	0.89	-
	3	3.02	1.18	0.36	-	0.72	0.59	3.22	0.63	0.09	3.19	1.79	-
	3.5	4.98	2.59	0.75	-	1.97	1.61	5.34	1.48	0.11	5.17	2.89	-
	4	7.4	4.14	1.49	-	-	3.2	7.79	2.59	0.17	7.74	5.56	-
Variance	2	0.23	0.26	0.35	0.08	0.15	0.24	0.62	0.56	0.62	0.76	0.62	0.65
	2.5	0.44	0.65	0.96	0.11	0.28	0.59	1.46	1.44	1.35	1.58	1.66	1.66
	3	0.89	1.35	2.05	0.15	0.46	1.22	2.78	2.62	2.58	3.46	3.2	3.1
	3.5	1.59	2.44	3.86	0.15	1.08	3.57	4.57	4.44	4.59	5.23	5.12	5.09
	4	2.61	4.02	5.77	0.17	2.79	6.2	6.74	6.63	6.62	8.24	7.68	7.64
Skewness	2	NA			NA			0.3	0.76	0.8	0.61	0.62	0.69
	2.5							0.71	1.85	2.02	1.54	1.66	1.67
	3							1.53	3.36	3.49	3.27	3.05	3.25
	3.5							2.89	5.47	5.81	5.02	5.09	5.13
	4							4.35	8.07	8.21	7.74	7.57	7.81

Table 10.5: MSE values of GOF method

Grouping event detection distribution by		True threshold value	Detection probability modelled by ...											
			Cum. Normal Dist.			Cum. Logistic Dist.			Cum. Log-Normal Dist.			Cum. Pareto type II Dist.		
			L	M	H	L	M	H	L	M	H	L	M	H
Exp. Value		2	-0.68	-0.28	-0.1	-	-0.29	-0.2	-0.78	-0.13	0.06	-0.76	-0.35	-
		2.5	-1.18	-0.56	-0.17	-	-0.48	-0.31	-1.28	-0.31	0.02	-1.26	-0.72	-
		3	-1.71	-0.98	-0.3	-	-0.69	-0.52	-1.77	-0.64	-0.02	-1.77	-1.36	-
		3.5	-2.22	-1.46	-0.57	-	-1.08	-0.92	-2.29	-1.07	-0.1	-2.27	-1.75	-
		4	-2.72	-1.97	-0.9	-	-	-1.41	-2.78	-1.52	-0.26	-2.78	-2.35	-
Variance		2	-0.19	-0.33	-0.48	-0.07	-0.23	-0.39	-0.66	-0.64	-0.66	-0.8	-0.73	-0.74
		2.5	-0.37	-0.59	-0.84	-0.09	-0.36	-0.65	-1.09	-1.07	-1.07	-1.25	-1.25	-1.24
		3	-0.59	-0.94	-1.32	-0.03	-0.46	-1.14	-1.53	-1.51	-1.53	-1.81	-1.76	-1.75
		3.5	-0.86	-1.35	-1.85	-0.06	-0.81	-1.85	-2.02	-2	-2.02	-2.26	-2.25	-2.25
		4	-1.2	-1.81	-2.38	-0.1	-1.48	-2.46	-2.49	-2.48	-2.48	-2.81	-2.75	-2.76
Skewness		2	NA			NA			-0.34	-0.83	-0.84	-0.72	-0.74	-0.76
		2.5							-0.65	-1.32	-1.37	-1.2	-1.23	-1.27
		3							-1.02	-1.82	-1.84	-1.74	-1.74	-1.78
		3.5							-1.48	-2.33	-2.37	-2.23	-2.25	-2.26
		4							-1.9	-2.83	-2.85	-2.74	-2.75	-2.78

Table 10.6: Bias values of GOF method

Grouping event detection distribution by		True threshold value	Detection probability modelled by ...											
			Cum. Normal Dist.			Cum. Logistic Dist.			Cum. Log-Normal Dist.			Cum. Pareto type II Dist.		
			L	M	H	L	M	H	L	M	H	L	M	H
Exp. Value		2	0.28	0.27	0.25	-	0.26	0.28	0.21	0.22	0.2	0.23	0.28	-
		2.5	0.28	0.36	0.29	-	0.38	0.35	0.22	0.31	0.19	0.23	0.4	-
		3	0.27	0.38	0.39	-	0.49	0.49	0.23	0.31	0.24	0.24	0.37	-
		3.5	0.28	0.43	0.44	-	0.57	0.59	0.22	0.31	0.24	0.23	0.36	-
		4	0.26	0.4	0.5	-	-	0.6	0.22	0.35	0.26	0.23	0.35	-
Variance		2	0.24	0.27	0.28	0.23	0.28	0.29	0.21	0.22	0.21	0.18	0.22	0.23
		2.5	0.27	0.31	0.35	0.25	0.36	0.45	0.22	0.25	0.26	0.23	0.24	0.24
		3	0.29	0.35	0.4	0.31	0.49	0.61	0.24	0.24	0.25	0.17	0.23	0.25
		3.5	0.31	0.39	0.44	0.3	0.67	0.68	0.23	0.24	0.26	0.21	0.23	0.23
		4	0.29	0.42	0.46	0.31	0.8	0.56	0.24	0.25	0.27	0.2	0.24	0.23
Skewness		2	NA			NA			0.24	0.21	0.19	0.23	0.23	0.23
		2.5							0.28	0.21	0.18	0.24	0.25	0.22
		3							0.29	0.21	0.22	0.25	0.25	0.23
		3.5							0.28	0.21	0.2	0.22	0.23	0.24
		4							0.3	0.2	0.23	0.23	0.25	0.22

Table 10.7: Standard error values of GOF method

From these tables the following conclusions can be drawn :

1. Bias
   (a) Biases are consistently negative
2. m_c as factor
   (a) MSE values increase for increasing m_c
   (b) Absolute biases increase for increasing m_c
   (c) Standard error values are humped. These decrease for mid valued threshold values and then ultimately increase.

3. Expected value of detection distribution as factor
   (a) MSE values reduce for increasing expected values
   (b) Absolute biases reduce for increasing expected values
   (c) Standard error (SE)
      i. Symmetric detection probability distributions
         As expected values increase ...
         A. Standard errors decrease for small m_c
         B. Standard errors are humped (increase, then ultimately decrease) for mid values of m_c
         C. Standard errors increase for large values of m_c
      ii. Non-symmetric detection probability distributions
         SE values appear humped (increaseing, then decreasing), at least for Log-Normal detection distribution. This can however not be conclusively stated, due to lack of values under the scenario of a Pareto Type II distribution.

4. Variance of detection distribution as factor
   (a) MSE values
      i. Symmetric detection distributions : MSE generally increases as variance increases
      ii. Non-symmetric detection distributions : MSE generally decreases as variance increases
   (b) Bias values
      i. Symmetric detection distributions : Absolute biases increase for increasing variance
      ii. Non-symmetric detection distributions : Bias values generally unaffected for differing variances
   (c) Standard error values increase for increasing variances

5. Skewness of detection distribution as factor
   (a) MSE values
      i. MSE increases sharply for increasing skew of Log-Normal detection distribution.
      ii. MSE increases only slightly for increasing skew of the Pareto Type II detection distribution.
   (b) Bias values
      i. Bias values increase sharply for increasing skew of Log-Normal detection distribution.
      ii. Bias values increase only slightly for increasing skew of Pareto Type II detection distribution.
   (c) Standard error

i. SE reduces for increasing skew of Log-Normal detection distribution.

ii. SE remains constant, with somewhat of humped progression (increasing, then decreasing) for increasing skew of Pareto Type II detection distribution.

The method show that considerable biases are possible over the range of expected values of the detection distribution, especially when the expected values are small. Biases are also consistently negative, showing that the estimation methods underestimate the true value of the detection threshold. For the cases of the Normal, Log-Normal and Pareto type II detection probability the biases appear to diminish with increasing expected values of the detection distribution. This establishes that for datasets with detection thresholds that become increasingly more prominent, the bias of the methods become smaller. For the case of a Logistic detection probability, the bias appears to remain relatively constant for increasing expected values of the detection distribution, showing a systematic error. The tendency of the median of the sampling distribution to be equal or smaller than the expected value appears indicative of a negative skew distribution enforcing the tendency of underestimated threshold estimates.

Considerable biases are also possible over the range of variances of the detection distribution. For the case of the symmetric distributions such biases are possible for increasing variances. For the distributions that can exhibit skewness, Pareto type II - and Log-Normal distribution, a constant systematic bias is shown. Bias for these two distributions are shown to be invariant for different values of the variance of the detection distribution.

The skewness of the Pareto type II and Log-Normal distributions also do not impact the bias of the estimation method. It does however appear that the bias , when the detection probability is modelled by the cumulative Log-Normal distribution, tends to be lower for skewness values smaller than 10.

In general, the GOF method show a maximum bias of -2.78. Biases are smaller for distributions where the threshold is more prominently defined and therefore the success of these estimation methods is heavily dependent on the expected value and variance of the detection distribution. A rule of thumb can follow that these methods have the possibility to yield unbiased results if the expected value of the detection distribution is greater or equal to 3. If the performance of the estimation method is influenced by the variance of the detection distribution, the value should be, on average, 1.75 or smaller. The combination of these factors shows that the detection threshold must be very prominently defined.

10.2.3.2 Threshold estimation by Maximum Curvature (MAXC) method

In order to consider the behaviour of the MAXC method, Tables 10.8, 10.9 and 10.10 have been constructed to respectively display Mean-Squared Error (MSE), Bias and Standard Error values.

Grouping event detection distribution by	True threshold value	Detection probability modelled by ...											
		Cum. Normal Dist.			Cum. Logistic Dist.			Cum. Log-Normal Dist.			Cum. Pareto type II Dist.		
		L	M	H	L	M	H	L	M	H	L	M	H
Exp. Value	2	0.77	0.58	0.38	-	0.64	0.62	0.78	0.21	0.01	0.77	0.48	-
	2.5	1.89	1.47	1	-	1.56	1.54	1.9	0.74	0.02	1.91	1.34	-
	3	3.5	2.85	1.95	-	2.86	2.81	3.54	1.63	0.11	3.52	3.03	-
	3.5	5.64	4.72	3.26	-	4.48	4.49	5.68	2.98	0.4	5.65	5.04	-
	4	8.26	7.08	4.92	-	-	6.71	8.28	4.88	0.94	8.3	7.33	-
Variance	2	0.29	0.61	0.77	0.34	0.73	0.77	0.66	0.66	0.66	0.74	0.76	0.77
	2.5	0.79	1.54	1.88	0.8	1.84	1.89	1.63	1.67	1.68	1.9	1.88	1.89
	3	1.58	2.99	3.46	1.36	3.4	3.5	3.08	3.1	3.2	3.44	3.45	3.52
	3.5	2.67	4.87	5.64	2.02	5.5	5.6	5.03	5.06	5.11	5.67	5.6	5.64
	4	4.2	7.29	8.18	3.35	8.06	8.24	7.41	7.43	7.63	8.23	8.19	8.3
Skewness	2	NA			NA			0.44	0.8	0.78	0.74	0.76	0.77
	2.5							1.17	1.94	1.95	1.83	1.89	1.91
	3							2.31	3.58	3.6	3.42	3.52	3.5
	3.5							3.91	5.74	5.76	5.57	5.62	5.66
	4							5.94	8.35	8.36	8.2	8.28	8.27

Table 10.8: MSE values of MAXC method

Grouping event detection distribution by	True threshold value	Detection probability modelled by ...											
		Cum. Normal Dist.			Cum. Logistic Dist.			Cum. Log-Normal Dist.			Cum. Pareto type II Dist.		
		L	M	H	L	M	H	L	M	H	L	M	H
Exp. Value	2	-0.86	-0.68	-0.44	-	-0.73	-0.68	-0.87	-0.32	0.02	-0.86	-0.68	-
	2.5	-1.36	-1.11	-0.76	-	-1.13	-1.11	-1.37	-0.68	-0.03	-1.37	-1.18	-
	3	-1.86	-1.57	-1.09	-	-1.54	-1.52	-1.87	-1.12	-0.16	-1.86	-1.7	-
	3.5	-2.36	-2.06	-1.46	-	-1.94	-1.93	-2.37	-1.59	-0.39	-2.37	-2.22	-
	4	-2.86	-2.56	-1.87	-	-	-2.36	-2.87	-2.08	-0.72	-2.87	-2.7	-
Variance	2	-0.36	-0.7	-0.85	-0.39	-0.82	-0.86	-0.74	-0.75	-0.76	-0.85	-0.86	-0.86
	2.5	-0.63	-1.15	-1.34	-0.62	-1.32	-1.36	-1.2	-1.21	-1.23	-1.36	-1.36	-1.36
	3	-0.92	-1.62	-1.84	-0.82	-1.8	-1.85	-1.67	-1.68	-1.72	-1.85	-1.84	-1.86
	3.5	-1.26	-2.1	-2.35	-1.01	-2.3	-2.35	-2.15	-2.17	-2.21	-2.36	-2.35	-2.36
	4	-1.66	-2.62	-2.84	-1.28	-2.8	-2.85	-2.64	-2.65	-2.71	-2.86	-2.85	-2.87
Skewness	2	NA			NA			-0.53	-0.88	-0.88	-0.85	-0.86	-0.86
	2.5							-0.92	-1.38	-1.39	-1.34	-1.36	-1.37
	3							-1.34	-1.88	-1.89	-1.84	-1.86	-1.86
	3.5							-1.8	-2.38	-2.39	-2.35	-2.36	-2.36
	4							-2.28	-2.88	-2.89	-2.84	-2.86	-2.86

Table 10.9: Bias values of MAXC method

Grouping event detection distribution by	True threshold value	Detection probability modelled by ...											
		Cum. Normal Dist.			Cum. Logistic Dist.			Cum. Log-Normal Dist.			Cum. Pareto type II Dist.		
		L	M	H	L	M	H	L	M	H	L	M	H
Exp. Value	2	0.04	0.1	0.13	-	0.11	0.13	0.04	0.11	0.05	0.05	0.1	-
	2.5	0.05	0.12	0.17	-	0.14	0.15	0.04	0.14	0.08	0.04	0.09	-
	3	0.05	0.13	0.19	-	0.19	0.18	0.04	0.14	0.12	0.04	0.1	-
	3.5	0.05	0.14	0.22	-	0.18	0.22	0.04	0.15	0.18	0.04	0.1	-
	4	0.04	0.15	0.22	-	-	0.25	0.04	0.16	0.19	0.04	0.1	-
Variance	2	0.09	0.13	0.07	0.19	0.09	0.06	0.05	0.06	0.07	0.05	0.05	0.05
	2.5	0.13	0.16	0.08	0.24	0.09	0.06	0.06	0.08	0.07	0.05	0.05	0.05
	3	0.14	0.18	0.08	0.3	0.11	0.07	0.07	0.07	0.08	0.05	0.05	0.04
	3.5	0.16	0.2	0.07	0.36	0.12	0.07	0.07	0.09	0.09	0.05	0.05	0.04
	4	0.18	0.19	0.08	0.42	0.12	0.07	0.08	0.08	0.1	0.04	0.05	0.04
Skewness	2	NA			NA			0.08	0.03	0.03	0.05	0.05	0.05
	2.5							0.1	0.03	0.03	0.05	0.05	0.05
	3							0.11	0.03	0.03	0.06	0.05	0.04
	3.5							0.12	0.03	0.03	0.05	0.05	0.04
	4							0.13	0.03	0.03	0.05	0.04	0.04

Table 10.10: Standard error values of MAXC method

From these tables the following conclusions can be drawn :

1. Bias
   (a) Biases are consistently negative.
2. m_c as factor
   (a) MSE values increase for increasing m_c.
   (b) Absolute biases increase for increasing m_c.
   (c) Standard error values relatively constant for increasing m_c. Slight increases are possible for increasing values of m_c.
3. Expected value of detection distribution as factor
   (a) MSE values sharply decrease for increasing expected values.
   (b) Absolute biases decrease sharply for increasing expected values.
   (c) Standard error (SE) increases for increasing expected values for all, but low to mid valued m_c thresholds of Log-Normal detection distribution.
4. Variance of detection distribution as factor
   (a) MSE values
      i. Symmetric detection distributions : MSE generally increases as variance increases.
      ii. Non-symmetric detection distributions : MSE generally remains constant for differing variances.
   (b) Bias values
      i. Symmetric detection distributions : Absolute biases increase for increasing variance.
      ii. Non-symmetric detection distributions : Bias values generally unaffected for differing variances.
   (c) Standard error values
      i. Symmetric detection distributions : Values progress in a general humped (increasing, then decreasing) manner.
      ii. Non-symmetric detection distributions : Values remain relatively constant with slight decreases as variance increases.
5. Skewness of detection distribution as factor
   (a) MSE values : Slight increase in MSE values for increased skewness.
   (b) Bias values : Initial increase in absolute values for increasing skew, thereafter relatively constant.
   (c) Standard error : Reduces for increasing skew of the detection distribution

10.2.3.3 Threshold estimation by b-Value Stability (bVS) method

In order to consider the behaviour of the bVS method, Tables 10.11, 10.12 and 10.13 have been constructed to respectively display Mean-Squared Error (MSE), Bias and Standard Error values.

Grouping event detection distribution by	True threshold value	Detection probability modelled by ...											
		Cum. Normal Dist.			Cum. Logistic Dist.			Cum. Log-Normal Dist.			Cum. Pareto type II Dist.		
		L	M	H	L	M	H	L	M	H	L	M	H
Exp. Value	2	0.72	0.25	0.07	-	0.23	0.13	0.86	0.08	0.02	0.8	0.27	-
	2.5	1.82	0.82	0.2	-	0.64	0.45	2.02	0.27	0.03	1.96	0.81	-
	3	3.49	1.95	0.58	-	1.57	1.49	3.65	0.89	0.04	3.62	2.35	-
	3.5	5.7	3.55	1.54	-	3	3.25	5.88	1.87	0.1	5.75	3.4	-
	4	8.4	5.77	2.84	-	-	5.84	8.48	3.23	0.27	8.4	6.65	-
Variance	2	0.2	0.32	0.47	0.04	0.15	0.33	0.72	0.68	0.7	0.86	0.73	0.8
	2.5	0.5	0.83	1.34	0.07	0.41	1.05	1.69	1.67	1.64	1.89	1.91	1.93
	3	1.04	1.84	2.82	0.08	1.25	3	3.11	3.02	3.13	3.64	3.59	3.58
	3.5	1.87	3.38	5.03	0.15	3.27	5.76	5.05	4.99	5.07	5.82	5.68	5.68
	4	3.05	5.55	7.68	0.32	7.56	8.84	7.36	7.36	7.29	8.45	8.26	8.37
Skewness	2	NA			NA			0.31	0.94	0.95	0.72	0.79	0.8
	2.5							0.88	2.13	2.25	1.82	1.9	1.98
	3							1.77	3.87	3.94	3.46	3.56	3.64
	3.5							3.22	6.1	6.2	5.5	5.69	5.73
	4							4.93	8.79	8.86	8.22	8.33	8.41

Table 10.11: MSE values of bVS method

Grouping event detection distribution by	True threshold value	Detection probability modelled by ...											
		Cum. Normal Dist.			Cum. Logistic Dist.			Cum. Log-Normal Dist.			Cum. Pareto type II Dist.		
		L	M	H	L	M	H	L	M	H	L	M	H
Exp. Value	2	-0.78	-0.38	-0.17	-	-0.38	-0.28	-0.87	-0.19	0.02	-0.84	-0.44	-
	2.5	-1.3	-0.75	-0.3	-	-0.69	-0.53	-1.37	-0.4	-0.03	-1.36	-0.86	-
	3	-1.84	-1.23	-0.55	-	-1.08	-0.97	-1.87	-0.78	-0.07	-1.87	-1.51	-
	3.5	-2.36	-1.75	-0.96	-	-1.54	-1.52	-2.39	-1.22	-0.2	-2.37	-1.9	-
	4	-2.85	-2.28	-1.38	-	-	-2.09	-2.88	-1.71	-0.41	-2.87	-2.51	-
Variance	2	-0.26	-0.42	-0.58	-0.14	-0.32	-0.51	-0.74	-0.72	-0.73	-0.88	-0.8	-0.84
	2.5	-0.45	-0.74	-1.05	-0.17	-0.58	-0.98	-1.18	-1.16	-1.17	-1.33	-1.34	-1.35
	3	-0.7	-1.15	-1.61	-0.16	-1.03	-1.69	-1.64	-1.62	-1.65	-1.88	-1.85	-1.86
	3.5	-1	-1.66	-2.22	-0.25	-1.78	-2.35	-2.12	-2.11	-2.15	-2.37	-2.35	-2.35
	4	-1.34	-2.2	-2.75	-0.42	-2.69	-2.92	-2.59	-2.59	-2.63	-2.89	-2.85	-2.87
Skewness	2	NA			NA			-0.4	-0.93	-0.93	-0.79	-0.84	-0.84
	2.5							-0.74	-1.42	-1.47	-1.3	-1.34	-1.36
	3							-1.13	-1.93	-1.95	-1.83	-1.85	-1.87
	3.5							-1.59	-2.44	-2.46	-2.32	-2.35	-2.36
	4							-2.04	-2.93	-2.96	-2.83	-2.86	-2.88

Table 10.12: Bias values of bVS method

Grouping event detection distribution by		True threshold value	Detection probability modelled by ...											
			Cum. Normal Dist.			Cum. Logistic Dist.			Cum. Log-Normal Dist.			Cum. Pareto type II Dist.		
			L	M	H	L	M	H	L	M	H	L	M	H
Exp. Value		2	0.18	0.19	0.14	-	0.18	0.18	0.13	0.12	0.11	0.14	0.17	-
		2.5	0.16	0.25	0.22	-	0.27	0.3	0.13	0.19	0.1	0.14	0.2	-
		3	0.16	0.25	0.3	-	0.32	0.35	0.13	0.21	0.13	0.14	0.18	-
		3.5	0.15	0.25	0.33	-	0.34	0.37	0.12	0.22	0.13	0.13	0.21	-
		4	0.14	0.25	0.32	-	-	0.36	0.12	0.23	0.21	0.13	0.19	-
Variance		2	0.13	0.18	0.19	0.11	0.19	0.22	0.12	0.13	0.13	0.1	0.14	0.14
		2.5	0.14	0.23	0.25	0.14	0.32	0.34	0.13	0.14	0.15	0.14	0.14	0.14
		3	0.17	0.27	0.28	0.17	0.46	0.31	0.13	0.15	0.16	0.11	0.13	0.15
		3.5	0.19	0.28	0.28	0.29	0.52	0.21	0.14	0.14	0.14	0.13	0.13	0.14
		4	0.2	0.3	0.24	0.48	0.38	0.15	0.14	0.15	0.17	0.09	0.13	0.13
Skewness		2	NA			NA			0.14	0.12	0.12	0.14	0.14	0.14
		2.5							0.17	0.12	0.08	0.15	0.14	0.14
		3							0.18	0.11	0.09	0.13	0.14	0.13
		3.5							0.18	0.11	0.08	0.14	0.13	0.14
		4							0.2	0.11	0.09	0.12	0.14	0.12

Table 10.13: Standard error values of bVS method.

1. Bias
   (a) Biases are consistently negative.
2. m_c as factor
   (a) MSE values increase for increasing m_c.
   (b) Absolute biases increase for increasing m_c.
   (c) Standard error (SE) values
      i. SE decreases for low expected values of the detection distribution as m_c increases.
      ii. SE values hump (increases, then ultimately decreases) for all other cases.
3. Expected value of detection distribution as factor
   (a) MSE values decrease for increasing expected values.
   (b) Absolute biases decrease sharply for increasing expected values.
   (c) Standard error (SE) values
      i. SE humped (increaseing, then ultimately decreasing) for small m_c of Normal detection distribution as expected values increase.
      ii. SE humped for mid to large m_c of Log-Normal detection distribution as expected values increase.
      iii. Otherwise, SE increases for increasing expected values

   increases for increasing expected values for all, but low to mid valued m_c thresholds of Log-Normal detection distribution.
4. Variance of detection distribution as factor

(a) MSE values

i. Symmetric detection distributions : MSE generally increases as variance increases.

ii. Non-symmetric detection distributions : MSE generally humped (increasing, then decreasing) with decreases for larger variances.

(b) Bias values

i. Symmetric detection distributions : Absolute biases increase for increasing variance.

ii. Non-symmetric detection distributions : Bias values generally humped (increasing, then decreasing)for increasing variances.

(c) Standard error values : Generally increasing for increasing variance values.

5. Skewness of detection distribution as factor

(a) MSE values : MSE value increase for increased skewness.

(b) Bias values : Absolute biases increase for increased skewness.

(c) Standard error : SE reduces for increased skew.

From the above listing it can be seen that where the bias of the method increases, the standard error decreases. This can translate into a situation where greater errors can be made and there exists less chance to escape them. Furthermore this situation gives the opportunity to correct for known biases, since the variability in estimates are lower. However, given the high dependence of the bias on the true value of the threshold, the possibility to correct for bias is greatly reduced.

When the results are aggregated and compared only by threshold estimation method and functional form of the detection probability, these three methods consistently yield comparable results.

The methods show that considerable biases are possible over the range of expected values of the detection distribution, especially when the expected values are small. Biases are also consistently negative, showing that all three threshold estimation methods underestimate the true value of the detection threshold. For the cases of the Normal, Log-Normal and Pareto type II detection probability the biases appear to diminish with increasing expected values of the detection distribution. This establishes that for datasets with detection thresholds that become increasingly more prominent, the bias of the methods become smaller. For the case of a Logistic detection probability, the bias appears to remain relatively constant for increasing expected values of the detection distribution, showing a systematic error. The tendency of the median of the sampling distribution to be equal or smaller than the expected value appears indicative of a negative skew distribution enforcing the tendency of underestimated threshold estimates.

Considerable biases are also possible over the range of variances of the detection distribution. For the case of the symmetric distributions such biases are possible for increasing variances. For the distributions

that can exhibit skewness, Pareto type II - and Log-Normal distribution, a constant systematic bias is shown. Bias for these two distributions are shown to be invariant for different values of the variance of the detection distribution.

The skewness of the Pareto type II and Log-Normal distributions also do not impact the bias of the estimation method. It does however appear that the bias , when the detection probability is modelled by the cumulative Log-Normal distribution, tends to be lower for skewness values smaller than 10.

In general, the GOF, MAXC and bVS methods show a maximum bias of -2. Biases are smaller for distributions where the threshold is more prominently defined and therefore the success of these estimation methods is heavily dependent on the expected value and variance of the detection distribution. A rule of thumb can follow that these methods have the possibility to yield unbiased results if the expected value of the detection distribution is greater or equal to 3. If the performance of the estimation method is influenced by the variance of the detection distribution, the value should be, on average, 1.75 or smaller. The combination of these factors shows that the detection threshold must be very prominently defined.

10.2.3.4 Comment on GOF, MAXC and bVS estimation methods

AS can be seen from the preceding results. Detection capability of method is generally affected by the graphical prominence of threshold. A measure of this can be found by examining the expected value of the detection distribution. In general these methods can identify threshold values that are to some degree readily spotted by graphical inspection. Further more, for these cases where a value is identified, the bias of the method reduces, but this comes with trade-off with higher standard errors.

A different situation to consider is where biases increase and standard errors decrease. Through this, worse results are obtained, with less probability of escaping from them. However, this theoretically gives opportunity to correct for bias, since variability of the estimates are low. Unfortunately, this cannot be done, due to the biases being dependent on the true value of the threshold value.

What can be seen in in short is that these methods do not improve much on visual inspection and cannot accurately estimate less prominently defined threshold values.

10.2.3.5 Threshold estimation by Entire Magnitude Range (EMR) method

In order to consider the behaviour of the EMR method, Tables 10.14, 10.15 and 10.16 have been constructed to respectively display Mean-Squared Error (MSE), Bias and Standard Error values.

Grouping event detection distribution by		True threshold value	Detection probability modelled by ...											
			Cum. Normal Dist.			Cum. Logistic Dist.			Cum. Log-Normal Dist.			Cum. Pareto type II Dist.		
			L	M	H	L	M	H	L	M	H	L	M	H
Exp. Value		2	2.31	2.2	1.28	-	1.81	1.65	2.42	0.7	3.34	2.37	2.4	-
		2.5	1.83	1.23	1.67	-	1.19	1.61	1.97	0.45	0.93	2.05	0.49	-
		3	2.15	1.84	1.09	-	1.71	1.38	1.78	0.98	1.93	2.01	1.07	-
		3.5	2.66	2.15	2.07	-	2.11	1.8	2.45	1.33	0.91	2.56	1.71	-
		4	3.26	2.52	2.3	-	-	2.36	3	1.42	0.69	3.22	3.25	-
Variance		2	2.47	1.44	1.93	1.53	2.16	1.46	2.32	2.14	1.94	2.56	2.3	2.38
		2.5	1.36	1.55	1.75	1.55	1.17	1.58	1.75	1.42	1.96	1.79	2.12	1.99
		3	1.79	1.55	1.71	1.11	1.4	1.89	1.61	1.66	1.94	2.22	2.04	1.95
		3.5	2.3	2.02	2.51	1.68	1.48	2.38	2.34	2.06	2.03	2.84	2.16	2.63
		4	2.27	2.27	3.41	1.47	2.32	3.15	2.74	2.69	2.39	3.85	3.2	3.19
Skewness		2							2.09	2.27	2.08	2.95	2.38	2.22
		2.5							1.22	1.99	1.84	2.14	1.85	2.18
		3							1.5	1.72	1.48	1.93	2.04	1.92
		3.5		NA			NA		1.92	2.63	2.21	2.32	2.6	2.49
		4							2.43	3.09	1.87	3.24	3.22	3.23

Table 10.14: MSE values of EMR method

Grouping event detection distribution by		True threshold value	Detection probability modelled by ...											
			Cum. Normal Dist.			Cum. Logistic Dist.			Cum. Log-Normal Dist.			Cum. Pareto type II Dist.		
			L	M	H	L	M	H	L	M	H	L	M	H
Exp. Value		2	0.37	0.47	0.34	-	0.33	0.38	0.61	0.2	0.98	0.53	0.8	-
		2.5	-0.11	0.03	0.44	-	0.02	0.24	0.17	0.2	0.32	0.15	-0.12	-
		3	-0.56	-0.1	0.26	-	-0.01	0.05	-0.27	0.13	0.7	-0.52	-0.56	-
		3.5	-1.08	-0.47	0.35	-	-0.36	-0.19	-0.82	-0.07	0.44	-0.96	-0.82	-
		4	-1.32	-1.08	0.03	-	-	-0.59	-1.16	-0.58	0.28	-1.21	-1.49	-
Variance		2	0.66	0.29	0.27	0.47	0.51	0.1	0.6	0.59	0.43	0.75	0.61	0.51
		2.5	0.2	0.18	0.01	0.51	0.05	-0.07	0.18	0.11	0.31	0.29	0.25	0.1
		3	0.24	-0.17	-0.36	0.39	0.09	-0.33	-0.16	-0.11	-0.27	-0.8	-0.45	-0.53
		3.5	0.16	-0.51	-0.68	0.64	-0.34	-0.85	-0.77	-0.55	-0.36	-1.48	-0.72	-1
		4	-0.24	-0.69	-1.29	0.61	-0.74	-1.46	-1.01	-0.99	-1.01	-1.82	-1.28	-1.15
Skewness		2							0.45	0.58	0.36	0.7	0.52	0.53
		2.5							0.1	0.08	0.59	0.04	0.16	0.14
		3							-0.1	-0.34	0	-0.67	-0.53	-0.47
		3.5		NA			NA		-0.6	-0.87	-0.47	-0.82	-1.04	-0.87
		4							-0.97	-1.08	-0.88	-1.52	-1.2	-1.18

Table 10.15: Bias values of MEMR method

Grouping event detection distribution by		True threshold value	Detection probability modelled by ...											
			Cum. Normal Dist.			Cum. Logistic Dist.			Cum. Log-Normal Dist.			Cum. Pareto type II Dist.		
			L	M	H	L	M	H	L	M	H	L	M	H
Exp. Value		2	0.52	0.4	0.27	-	0.46	0.44	0.36	0.21	-	0.39	0.34	-
		2.5	0.55	0.53	0.27	-	0.53	0.49	0.42	0.28	-	0.46	0.47	-
		3	0.63	0.64	0.46	-	0.72	0.59	0.38	0.3	-	0.42	0.27	-
		3.5	0.5	0.71	0.62	-	0.87	0.67	0.38	0.32	-	0.4	0.34	-
		4	0.52	0.73	0.96	-	-	0.64	0.39	0.32	-	0.41	0.72	-
Variance		2	0.24	0.4	0.5	0.21	0.5	0.53	0.33	0.33	0.36	0.41	0.29	0.41
		2.5	0.28	0.51	0.54	0.12	0.5	0.69	0.39	0.3	0.52	0.28	0.53	0.44
		3	0.48	0.59	0.65	0.08	0.66	0.83	0.34	0.42	0.33	0.23	0.22	0.48
		3.5	0.43	0.63	0.74	0.08	0.84	0.85	0.34	0.4	0.35	0.59	0.43	0.37
		4	0.68	0.74	0.86	0.14	0.76	0.75	0.35	0.38	0.42	0.3	0.39	0.44
Skewness		2							0.4	0.32	0.25	0.42	0.38	0.38
		2.5							0.34	0.46	0.34	0.57	0.44	0.46
		3							0.42	0.36	0.06	0.25	0.47	0.38
		3.5		NA			NA		0.37	0.37	0.28	0.55	0.36	0.41
		4							0.41	0.38	0.1	0.59	0.43	0.36

Table 10.16: Standard error values of MEMR method

1. Bias
   (a) Biases are consistently negative.
2. m_c as factor
   (a) MSE values increase for increasing m_c.
   (b) Absolute biases increase for increasing m_c.
   (c) Standard error (SE) values
      i. SE decreases for low expected values of the detection distribution as m_c increases.
      ii. SE values hump (increases, then ultimately decreases) for all other cases.
3. Expected value of detection distribution as factor
   (a) MSE values decrease for increasing expected values.
   (b) Absolute biases decrease sharply for increasing expected values.
   (c) Standard error (SE) values
      i. SE humped (increaseing, then ultimately decreasing) for small m_c of Normal detection distribution as expected values increase.
      ii. SE humped for mid to large m_c of Log-Normal detection distribution as expected values increase.
      iii. Otherwise, SE increases for increasing expected values

   increases for increasing expected values for all, but low to mid valued m_c thresholds of Log-Normal detection distribution.
4. Variance of detection distribution as factor
   (a) MSE values
      i. Symmetric detection distributions : MSE generally increases as variance increases.
      ii. Non-symmetric detection distributions : MSE generally humped (increasing, then decreasing) with decreases for larger variances.
   (b) Bias values
      i. Symmetric detection distributions : Absolute biases increase for increasing variance.
      ii. Non-symmetric detection distributions : Bias values generally humped (increasing, then decreasing)for increasing variances.
   (c) Standard error values : Generally increasing for increasing variance values.
5. Skewness of detection distribution as factor
   (a) MSE values : MSE value increase for increased skewness.

(b) Bias values : Absolute biases increase for increased skewness.

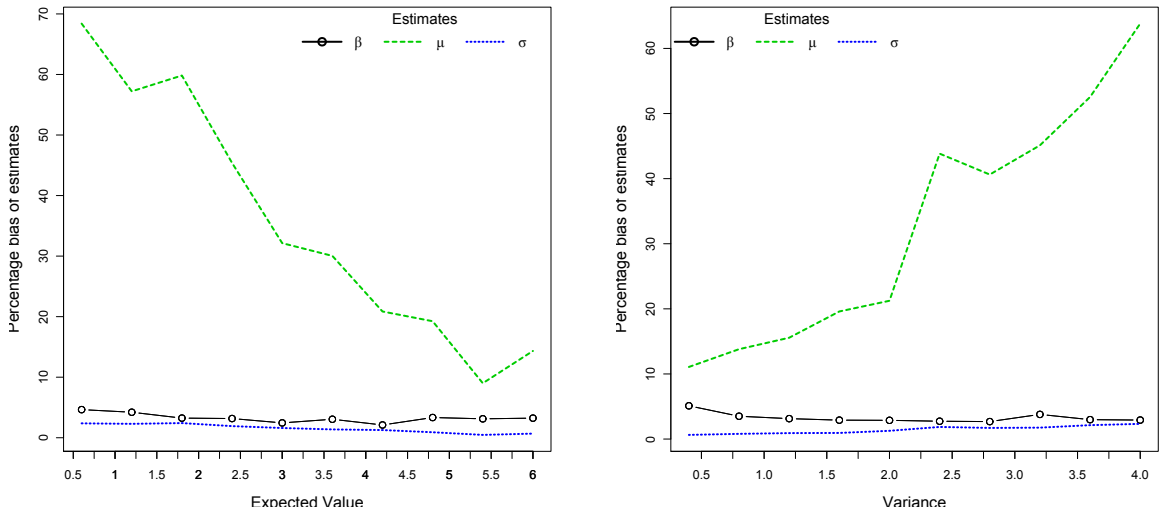
(c) Standard error : SE reduces for increased skew.

Compared to previously discussed threshold estimation methods (GOF, MAXC and bVS) this method yields an absolute expected biases of a smaller order. The sampling distribution of the EMR method is more closely centred around the true value of the threshold than the previous methods. Having a sampling distribution where the median is larger than the mean, is indicative of a positively skewed distribution and shows over-estimation of the threshold is more likely. Absolute biases decrease for increasing expected values up to an expected value of around 3.5, where after absolute biases increase again. Error bounds completely converge for expected values of non-symmetric distributions with expected value greater than 3.

Compared to previous methods, estimates form the EMR method show a markable reduced bias for differing variances of the detection distribution. The trend, however, still continues in that symmetric distributions show decreasing bias (eventually becoming negative) for progressively larger variance values, whereas the non-symmetric distributions show steadily decreasing biases in absolute terms.

The skewness of the non-symmetric impact the bias of the estimation method very slightly. Greater positively skewed distributions yield slightly smaller absolute biases. For the Log-Normal detection probability the biases are of a considerably lower order as that for the Pareto type II distribution, where the previous maximum of -2 is attained by the progression of the 95% bias quantile.

A useful consequence of this method is that the β and nuisance parameters (μ and σ from Normal distribution) of the detection distribution is estimated from the data. The situation where the simulated data incorporates a detection probability modelled by a cumulative Normal distribution is considered, as this represents ideal circumstances for recovery of the correct parameters. The bias of estimates is plotted against expected values and variances of the detection distribution in Figures 10.1, a and b with a summary of the distribution of biases given in Table 10.17. Although the EMR method has proven to be very robust for the determination of threshold values, analysis of biases shows that improvements are possible. Biases for the β and σ parameters remain stable, whereas biases for the μ parameter vary significantly. These biases do decrease for increasing expected values, but indicate the possibility of a very large maximum. Furthermore, as the variance of the Normal distribution is not dependent on μ the systematic increase in bias for increasing variance values represents unwanted behaviour. Considering the importance of the β value, the biases present in Table 10.17 is worrying. This particularly since β has been kept constant through the investigation.



(a) **(b)**

Figure 10.1: *Biases of parameter estimates obtained through EMR method where true detection probability is modelled by the cumulative Normal distribution.*

Parameter	Mean	5% Quantile	Median	95% Quantile
β	3.21	-0.39	0.15	15.16
μ	34.30	-1.95	3.57	188.26
σ	1.49	-1.17	0.20	7.07

Table 10.17: *Table of summary statistics on the distribution of biases of estimates of parameters. Estimates are through EMR method utilizing maximum likelihood estimation where true detection probability is modelled by cumulative Normal distribution.*

In general, the EMR method shows considerable improvement over the previously investigated methods when considering absolute biases, error bounds and the distribution of biases around 0. A strength of this method is that it can successfully estimate detection thresholds that are not as prominently defined as that needed by previous methods for accurate estimates. Expected values of the detection distribution as low as 0, for the Log-Normal distribution, can be handled by the method. In general, biases are of small order for varying variances. The combination of these factors shows that this estimation method is relatively robust.

10.2.3.6 Threshold estimation by Median Based Assessment of the Segment-Slope (MBASS) method

In order to consider the behaviour of the MBASS method, Tables 10.18, 10.19 and 10.20 have been constructed to respectively display Mean-Squared Error (MSE), Bias and Standard Error values.

Grouping event detection distribution by		True threshold value	Detection probability modelled by ...											
			Cum. Normal Dist.			Cum. Logistic Dist.			Cum. Log-Normal Dist.			Cum. Pareto type II Dist.		
			L	M	H	L	M	H	L	M	H	L	M	H
Grouping event detection distribution by	Exp. Value	2	1.42	1.29	0.89	-	1.43	1.15	1.33	0.61	0.13	1.33	0.86	-
		2.5	1.48	1.52	1	-	1.65	1.59	1.27	0.63	0.23	1.36	0.89	-
		3	1.51	1.78	1.66	-	1.95	2.12	1.21	0.72	0.68	1.38	0.83	-
		3.5	1.55	1.94	2.46	-	2.48	2.77	1.26	0.96	1.36	1.41	1.14	-
		4	1.62	1.96	2.92	-	-	3.88	1.32	0.96	1.86	1.37	1.06	-
	Variance	2	0.7	1.17	1.62	0.65	1.41	1.75	1.12	1.2	1.28	1.25	1.28	1.33
		2.5	0.81	1.42	1.67	0.94	1.93	1.91	1.11	1.12	1.21	1.38	1.37	1.34
		3	1.04	1.83	2	1.29	2.16	2.65	1.1	1.08	1.2	1.11	1.38	1.38
		3.5	1.68	2.02	2.27	2.18	2.99	2.92	1.18	1.24	1.38	1.34	1.39	1.41
		4	1.97	2.14	2.45	3.56	4.2	3.85	1.29	1.28	1.34	0.95	1.31	1.4
	Skewness	2							0.88	1.44	1.32	1.38	1.26	1.36
		2.5							0.93	1.28	1.28	1.44	1.32	1.36
		3							0.99	1.26	1.19	1.38	1.34	1.4
		3.5		NA			NA		1.23	1.27	1.03	1.41	1.42	1.37
		4							1.39	1.28	1.38	1.25	1.46	1.26

Table 10.18: MSE values of MBASS method

Grouping event detection distribution by		True threshold value	Detection probability modelled by ...											
			Cum. Normal Dist.			Cum. Logistic Dist.			Cum. Log-Normal Dist.			Cum. Pareto type II Dist.		
			L	M	H	L	M	H	L	M	H	L	M	H
Grouping event detection distribution by	Exp. Value	2	0.72	0.55	0.31	-	0.61	0.53	0.74	0.14	0.14	0.69	0.05	-
		2.5	0.26	0.22	0.16	-	0.31	0.33	0.23	-0.22	-0.02	0.24	-0.39	-
		3	-0.21	-0.16	-0.07	-	0	0.05	-0.27	-0.63	-0.18	-0.25	-0.84	-
		3.5	-0.67	-0.56	-0.32	-	-0.35	-0.21	-0.77	-1.09	-0.4	-0.76	-1.32	-
		4	-1.13	-1.03	-0.69	-	-	-0.45	-1.25	-1.59	-0.73	-1.24	-1.79	-
	Variance	2	0.28	0.53	0.72	0.29	0.63	0.75	0.62	0.63	0.66	0.63	0.6	0.7
		2.5	0.02	0.23	0.36	0.14	0.38	0.42	0.14	0.16	0.16	0.2	0.2	0.24
		3	-0.25	-0.12	-0.07	-0.04	0.05	0.1	-0.32	-0.32	-0.28	-0.32	-0.32	-0.25
		3.5	-0.54	-0.52	-0.47	-0.15	-0.25	-0.29	-0.81	-0.76	-0.8	-0.9	-0.81	-0.75
		4	-0.93	-0.99	-0.91	-0.21	-0.5	-0.6	-1.27	-1.3	-1.26	-1.36	-1.3	-1.24
	Skewness	2							0.35	0.79	0.77	0.62	0.67	0.7
		2.5							-0.04	0.25	0.28	0.15	0.23	0.23
		3							-0.47	-0.22	-0.23	-0.33	-0.28	-0.24
		3.5		NA			NA		-0.91	-0.74	-0.68	-0.84	-0.77	-0.77
		4							-1.38	-1.21	-1.21	-1.28	-1.26	-1.25

Table 10.19: Bias values of MBASS method

Grouping event detection distribution by		True threshold value	Detection probability modelled by ...											
			Cum. Normal Dist.			Cum. Logistic Dist.			Cum. Log-Normal Dist.			Cum. Pareto type II Dist.		
			L	M	H	L	M	H	L	M	H	L	M	H
Grouping event detection distribution by	Exp. Value	2	0.57	0.6	0.58	-	0.65	0.65	0.53	0.52	0.35	0.57	0.6	-
		2.5	0.59	0.64	0.6	-	0.65	0.66	0.54	0.52	0.19	0.57	0.7	-
		3	0.6	0.68	0.66	-	0.71	0.71	0.53	0.54	0.24	0.56	0.68	-
		3.5	0.58	0.72	0.75	-	0.79	0.77	0.54	0.56	0.32	0.57	0.68	-
		4	0.57	0.72	0.82	-	-	0.82	0.54	0.59	0.34	0.56	0.69	-
	Variance	2	0.48	0.62	0.63	0.59	0.69	0.65	0.51	0.51	0.55	0.53	0.58	0.56
		2.5	0.49	0.66	0.65	0.57	0.73	0.65	0.52	0.53	0.55	0.56	0.58	0.57
		3	0.53	0.7	0.69	0.59	0.78	0.73	0.51	0.54	0.53	0.59	0.58	0.56
		3.5	0.56	0.76	0.72	0.68	0.86	0.76	0.52	0.55	0.55	0.62	0.57	0.57
		4	0.58	0.8	0.73	0.73	0.94	0.77	0.52	0.57	0.57	0.53	0.59	0.56
	Skewness	2							0.56	0.5	0.48	0.57	0.56	0.57
		2.5							0.56	0.52	0.48	0.62	0.57	0.57
		3							0.58	0.51	0.47	0.59	0.57	0.56
		3.5		NA			NA		0.58	0.52	0.46	0.6	0.57	0.58
		4							0.59	0.51	0.48	0.58	0.57	0.57

Table 10.20: Standard error values of MBASS method

When compared to estimation techniques with graphical basis, the MBASS method also shows improved estimation capability. The method consistently underestimates the cases where the detection probability is modelled by a Normal, Log-Normal and Pareto type II distribution, but over estimates the threshold for the case of the Logistic distribution. The behaviour of expected biases grouped by expected value of the detection distribution shows some peculiarities, e.g. biases increase for larger values of the skew distributions, but, at least for the Log-Normal distribution, only up to a point. After this point biases decrease for larger values of the expected value.

The sampling distribution of this estimation method is more closely centred around the true value of the threshold than the graphical methods. Except for the case of the Pareto type II distribution, the median and mean of the sampling distribution is consistently close. This is an indication of a symmetric sampling distribution. The method also shows that absolute biases decrease for increasing expected values, up to an expected value of around 3.5 from where it increases. For the case of the non-symmetric distributions, error bounds completely converge for expected values greater than 3.

When analysed by variance, the method underestimates the threshold for all forms of the detection distribution, except for the Logistic distribution, which it overestimates. For these groups, symmetric distributions seem to have a much lower bias than for the skew distributions, which appear to have systematic biases.

Varying skewness of the non-symmetric distributions minimally impact the bias, which appears to be systematic, of the estimation method. The absolute systematic bias should be slightly greater for negatively skewed distributions than for those with positive skew.

A striking difference of this method is the large error bounds on the estimates. This indicates an extremely larger variance for the estimator and although very accurate, will increase the MSE value. The exception to this is the case where the detection distribution is modelled by the skew Log-Normal CDF and where the expected value of this distribution is very high.

In general, the MBASS method shows considerable improvement over the GOF, MAXC and bVS methods regarding bias. However, under certain circumstances, the error bounds of estimates might be exceptionally wide. The strength of this method is that it can successfully estimate detection thresholds that are not as prominently defined as that needed by other methods. Increased variance of the detection distribution translates into increased biases, but however, error bounds stay relatively constant. The combination of these factors shows that this estimation method is relatively robust, while not being as computationally intensive as parametric methods such as EMR and MITC.

10.2.3.7 Threshold estimation by Moment Incorporating Threshold Computation (MITC) method

In order to consider the behaviour of the bVS method, Tables 10.21, 10.22 and 10.23 have been constructed to respectively display Mean-Squared Error (MSE), Bias and Standard Error values.

Grouping event detection distribution by	True threshold value	Detection probability modelled by ...											
		Cum. Normal Dist.			Cum. Logistic Dist.			Cum. Log-Normal Dist.			Cum. Pareto type II Dist.		
		L	M	H	L	M	H	L	M	H	L	M	H
Exp. Value	2	1.29	1.57	1.49	-	1.61	1.41	1.38	1.42	1.51	1.29	1.84	-
	2.5	0.59	0.24	0.06	-	0.14	0.1	0.84	0.26	0.03	0.84	0.54	-
	3	0.86	0.46	0.16	-	0.29	0.31	1	0.55	0.12	1.03	0.84	-
	3.5	1.59	1.09	0.59	-	0.8	0.74	1.67	1.21	0.47	1.66	1.55	-
	4	2.92	2.31	1.17	-	-	1.53	2.76	1.91	0.94	2.82	1.81	-
Variance	2	2.12	1.46	0.85	1.92	1.15	1.51	1.46	1.35	1.18	0.84	1.33	1.34
	2.5	0.25	0.22	0.37	0.03	0.14	0.19	0.75	0.6	0.83	0.58	0.75	0.88
	3	0.46	0.39	0.57	0.04	0.34	0.6	0.9	0.84	1.09	0.79	1.15	1.01
	3.5	0.89	0.93	1.36	0.06	0.84	1.43	1.61	1.62	1.38	1.18	1.57	1.7
	4	1.7	1.87	2.61	0.11	2.08	2.67	2.59	2.62	2.88	2.81	2.53	2.88
Skewness	2	NA			NA			1.31	1.42	1.53	1.05	1.41	1.23
	2.5							0.42	0.92	0.88	0.57	0.72	1.04
	3							0.65	0.88	1.16	0.9	1.06	1.01
	3.5							1.37	1.66	1.79	1.53	1.68	1.64
	4							2.38	2.74	2.84	2.52	2.81	2.82

Table 10.21: MSE values of MITC method

Grouping event detection distribution by	True threshold value	Detection probability modelled by ...											
		Cum. Normal Dist.			Cum. Logistic Dist.			Cum. Log-Normal Dist.			Cum. Pareto type II Dist.		
		L	M	H	L	M	H	L	M	H	L	M	H
Exp. Value	2	0.54	0.52	0.5	-	0.53	0.48	0.61	0.53	0.51	0.59	0.41	-
	2.5	-0.01	-0.05	-0.02	-	-0.06	-0.04	0.08	0.05	0.04	0.07	0.02	-
	3	-0.51	-0.42	-0.15	-	-0.32	-0.25	-0.43	-0.16	0.11	-0.44	-0.42	-
	3.5	-1	-0.89	-0.44	-	-0.69	-0.6	-0.93	-0.49	0.21	-0.94	-0.92	-
	4	-1.47	-1.37	-0.82	-	-	-1.04	-1.36	-0.98	0.08	-1.42	-1.31	-
Variance	2	0.61	0.5	0.47	0.57	0.47	0.48	0.61	0.56	0.57	0.6	0.58	0.58
	2.5	0.01	-0.03	-0.05	0.01	-0.05	-0.1	0.07	0.07	0.11	0.04	0.06	0.07
	3	-0.25	-0.35	-0.46	0.02	-0.26	-0.53	-0.38	-0.37	-0.35	-0.31	-0.43	-0.45
	3.5	-0.56	-0.76	-0.96	-0.01	-0.66	-1.1	-0.85	-0.81	-0.8	-0.99	-0.94	-0.94
	4	-0.95	-1.2	-1.45	-0.04	-1.29	-1.66	-1.27	-1.29	-1.32	-1.48	-1.41	-1.41
Skewness	2	NA			NA			0.53	0.63	0.64	0.64	0.58	0.57
	2.5							0.03	0.08	0.17	0.03	0.05	0.1
	3							-0.32	-0.44	-0.38	-0.44	-0.44	-0.43
	3.5							-0.69	-0.93	-0.87	-0.95	-0.94	-0.93
	4							-1.18	-1.33	-1.38	-1.46	-1.42	-1.39

Table 10.22: Bias values of MITC method

Grouping event detection distribution by		True threshold value	Detection probability modelled by ...											
			Cum. Normal Dist.			Cum. Logistic Dist.			Cum. Log-Normal Dist.			Cum. Pareto type II Dist.		
			L	M	H	L	M	H	L	M	H	L	M	H
Exp. Value		2	0.61	0.64	0.62	-	0.65	0.63	0.62	0.69	0.65	0.6	0.57	-
		2.5	0.52	0.35	0.2	-	0.3	0.24	0.59	0.35	0.1	0.58	0.44	-
		3	0.53	0.37	0.3	-	0.33	0.36	0.6	0.47	0.22	0.59	0.49	-
		3.5	0.56	0.42	0.39	-	0.41	0.45	0.59	0.53	0.36	0.57	0.49	-
		4	0.58	0.46	0.46	-	-	0.47	0.63	0.57	0.62	0.6	0.56	-
Variance		2	0.7	0.62	0.56	0.7	0.61	0.61	0.65	0.62	0.6	0.62	0.61	0.6
		2.5	0.33	0.36	0.41	0.14	0.26	0.35	0.54	0.55	0.57	0.53	0.54	0.6
		3	0.37	0.37	0.46	0.14	0.4	0.41	0.57	0.56	0.6	0.58	0.61	0.57
		3.5	0.39	0.44	0.53	0.2	0.55	0.48	0.58	0.57	0.6	0.51	0.55	0.58
		4	0.44	0.49	0.56	0.28	0.62	0.44	0.63	0.62	0.59	0.5	0.58	0.61
Skewness		2	NA			NA			0.64	0.65	0.64	0.66	0.62	0.57
		2.5							0.46	0.58	0.58	0.55	0.57	0.61
		3							0.49	0.58	0.62	0.57	0.58	0.59
		3.5							0.53	0.59	0.62	0.54	0.57	0.58
		4							0.56	0.68	0.57	0.53	0.58	0.63

Table 10.23: Standard error values of MITC method

This method also shows superior estimation capability when compared to the previously discussed threshold estimation techniques (GOF, MAXC and bVS). For the cases where the detection probability is modelled by the Normal and Pareto type II CDFs, the method consistently underestimates the true value of the threshold. As for the situation where the detection probability is modelled by the Logistic and Log-Normal CDFs the sign of the bias is reversed at a certain point. Skewness of the sampling distribution is impacted by the expected value of the detection distribution.

The sampling distribution of this estimation method is more closely centred around the true value of the threshold than the graphical methods. Considering the relation of the median to the mean of the sampling distribution, the sampling distribution appears to be near symmetric. Also, absolute biases decrease for increasing expected values. Non-symmetric distributions detection probability distributions show that the error bounds completely converge for expected values greater than 3.

When analysed by variance of detection distribution, the method underestimates the threshold for all forms of the detection distribution, except for the Logistic distribution, which it overestimates. For these groups, symmetric distributions seem to have a much lower bias than for the skew distributions, which show to have systematic biases.

When grouped by skewness of the detection distribution, varying skewness of the non-symmetric distributions minimally impact the bias of estimates and it appears that systematic bias is present. It does however appear that the absolute systematic bias should be slightly greater for negatively skewed distributions, when compared to those with positive skew.

Error bounds on are on par with those seen from the MBASS method, which are much larger than for other methods This indicates a larger variance for the estimator and will increase the MSE value.

The exception to this is the case where the detection distribution is modelled by the skew Log-Normal distribution and where the expected value of this distribution is very high.

Comparable to the MEMR method, a useful by-product of the MITC scheme is that estimates of *beta* and nuisance parameters (μ and σ of Normal distribution) must be obtained. The ideal situation, where the simulated data incorporates a detection probability modelled by a Normal CDF is considered. The progression of biases as a function of expected value and variance of the detection distribution is illustrated in Figures 10.2 a and b with a summary of the distribution of biases given in Table 10.24. As in the case of the EMR method, biases for the β and σ parameters remain stable when compared to the μ parameter. In this case biases for the parameter μ increase for increasing expected values, i.e. increasing μ. However, the expected bias is much lower than the estimates obtained through the EMR method. Also, expected biases of the μ parameter increase for increasing variances. Comparable to results from the EMR method, this represents unwanted behaviour. Biases for estimates of the β parameter (Table 10.24) is considerably lower and the distribution thereof much more symmetric about the mean than the case of the EMR method. The importance of the β parameter affirms the usefulness of this method in estimating the complete set of distributional parameters.

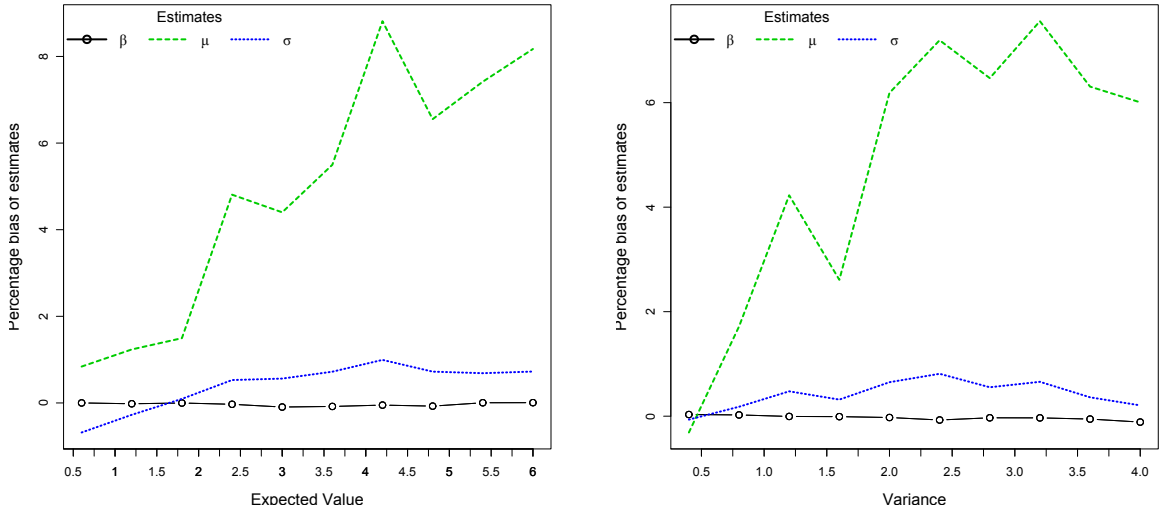


Figure 10.2: *Biases of parameter estimates obtained through MITC scheme where true detection probability is modelled by the cumulative Normal distribution.*

Parameter	Mean	5% Quantile	Median	95% Quantile
β	-0.03	-0.51	-0.04	0.41
μ	5.25	-2.98	0.10	36.88
σ	0.46	-1.49	-0.04	4.52

Table 10.24: *Table of summary statistics on the distribution of biases of estimates of parameters. Estimates are through MITC scheme utilizing maximum likelihood estimation where true detection probability is modelled by cumulative Normal distribution.*

The MITC scheme can be seen as a general improvement over the GOF, MAXC and bVS methods. Performance is comparable to that of the EMR method due to the incorporation of maximum likelihood estimation. A comparable strength of this method is that it can successfully estimate detection thresholds that are not as prominently defined as those needed by graphical methods. It is proposed that the technique of establishing maximum likelihood estimates be investigated to better the method.

10.2.3.8 Comment on EMR, MBASS and MITC estimation methods

As can be seen by the preceding analysis of the results, biases are positive in the following cases :

1. For smaller threshold values, which can be seen in general.
2. Large threshold values with large expected value of detection distribution
3. Small threshold values with high variance of detection distribution
4. Large threshold values with small variance of detection distribution

These methods will be seen as acting cautiously (obtaining an estimate with a positive bias) when

1. Maximising the quantity of complete data with respect to incomplete data
2. Expected value of detection distribution should be in vicinity of the threshold value
3. Sensitivity of detection probability (variance) is inversely related to the threshold value i.e. Small threshold values, with detection capabilities very stable, or High threshold values, with sudden changes present

10.2.4 Findings on stated study objective questions

In this section some of the fundamental questions relating to threshold estimation efficacy are addressed. These questions have been listed in the section relating to the objectives of the current study (Section 3, p. 9).

The questions will be answered based on a consideration of the MSE of estimates from the various methods, while taking the constituent components, i.e. bias and variance of the estimates into account.

The most appropriate estimation method to use under varying circumstances ...

1. **Given a dataset with a specific form of the detection probability**

   No value of the detection threshold, or values of the lower order moments of the RV whose CDF is attributable to the detection probability is specified. Therefore the Table 10.25 can be constructed by interpreting the bottom block of aggregated results in Table E.10 in the appendix.

	Cumulative distribution function of RV attributed to detection probability							
	Normal		Logistic		Log-Normal		Pareto type II	
Estimation method	MITC	(1.075)	MITC	(0.821)	MBASS	(1.186)	MBASS	(1.358)
	MBASS	(1.670)	GOF	(0.950)	MITC	(1.493)	MITC	(1.527)
	GOF	(1.871)	EMR	(1.713)	EMR	(2.099)	EMR	(2.424)
	EMR	(2.021)	bVS	(1.780)	GOF	(3.197)	GOF	(3.645)
	bVS	(2.456)	MBASS	(2.099)	bVS	(3.572)	MAXC	(4.011)
	MAXC	(3.188)	MAXC	(2.725)	MAXC	(3.582)	bVS	(4.066)

Table 10.25: *Table of threshold estimation methods sorted in descending order of appropriateness (based on MSE values), according to functional form of event detection probability.*

It can therefore be seen that the MITC scheme outperforms all methods for the cases where the detection probability is modelled by a cumulative Normal - or cumulative Logistic distribution. Furthermore, the MBASS method consistently outperforms all other methods for the case where the detection probability is modelled by a cumulative Log-Normal - or cumulative Pareto type II distribution. What is interesting to note is that the MITC method outperforms for symmetric detection probability distributions, whereas the MBASS method outperforms for cases of a non-symmetric detection distribution. It should be noted however, that for the case where the detection probability is modelled by a cumulative Logistic distribution, the MBASS method is ranked as the 5^{th} most effective method, whereas the lowest ranking that the MITC method is awarded is 2^{nd} best over all functional forms of the event detection probability.

The MAXC estimation method is awarded the lowest ranking in all but one situation, the case where the detection probability is modelled by the cumulative Pareto type II distribution, where it is ranked 2^{nd} lowest.

2. **Given a dataset with a specific level of completeness**

   No assumption is made regarding the form or characteristics of the detection probability. Therefore data has been aggregated only by the true detection threshold as well as threshold estimation method and stated in Table 10.26 below.

	Value of detection threshold									
	2		2.5		3		3.5		4	
Estimation method	GOF	(0.43)	MITC	(0.50)	MITC	(0.72)	MITC	(1.33)	MBASS	(1.91)
	bVS	(0.51)	GOF	(1.07)	MBASS	(1.50)	MBASS	(1.72)	MITC	(2.41)
	MAXC	(0.66)	bVS	(1.30)	EMR	(1.73)	EMR	(2.27)	EMR	(2.79)
	MBASS	(1.24)	MBASS	(1.34)	GOF	(2.12)	GOF	(3.73)	GOF	(5.79)
	MITC	(1.42)	MAXC	(1.63)	bVS	(2.63)	bVS	(4.50)	bVS	(6.91)
	EMR	(2.07)	EMR	(1.69)	MAXC	(3.06)	MAXC	(4.96)	MAXC	(7.36)

Table 10.26: *Table of threshold estimation methods sorted in descending order of appropriateness (based on MSE values), according to the true value of the detection threshold.*

As can be seen when MSE values are compared, the MITC method offers superior estimation results for the middle ranged values of the true detection threshold. For the case where the true detection threshold has a value of 2, the GOF method has the lowest MSE and for the case of a detection threshold of 4 the MBASS method shows superior performance. This indicates the relative strength of these two methods when smaller amounts of data are available of a specific side of the detection threshold. Furthermore, the high MSE value of the GOF method for for higher values of the true detection threshold illustrates the relatively low ranking of the method in Question 1, as discussed above.

3. **Given a dataset with unknown characteristics**

   This question requires that all available data be aggregated by threshold estimation method. Therefore the following table has been constructed :

Estimation method	MSE
bVS	3.08
EMR	2.10
GOF	2.55
MAXC	3.44
MBASS	1.53
MITC	1.27

Table 10.27: *Tabulation of aggregated MSE values for different threshold estimation procedures.*

Accordingly, it can be seen that based on all available data the MITC and MBASS methods outperform the other threshold estimation methods. Therefore, if no other data is available, such as expert opinion, it is advised to estimate the detection threshold by these methods.

10.3 Results for magnitude of completeness modelled as implicit detection threshold

Being that a detection threshold has been implicitly included in the event distribution, direct calculation of biases in threshold estimation is not possible. For this reason data will be analysed by the expected number of events not detected as a result of the relevant threshold estimate, $\hat{m}_c$. If $N_G(m)$ represents

the number of events, greater than magnitude m, that are generated in a given space-time volume, the following result can be derived :

$$\begin{aligned} N_G(m) &= N_G(m_{min})\bar{F}_{M_G}(m) = N_G(m_{min})\int_m^{m_{max}} f_{M_G}(m)\,dm \\ &= N_G(m_{min})\int_m^{m_{max}} f_{M_G}(m)\left(1 - P[M_G \in D|M_G = m] + P[M_G \in D|M_G = m]\right)\,dm \\ &= N_G(m_{min})\int_m^{m_{max}} f_{M_G}(m)P[M_G \in D|M_G = m]\,dm \\ &\quad + N_G(m_{min})\int_m^{m_{max}} f_{M_G}(m)\left(1 - P[M_G \in D|M_G = m]\right)\,dm \\ &= N_D(m) + N_{D^C}(m) \end{aligned} \tag{10.9}$$

where $N_D(m)$ and $N_{D^C}(m)$ is, respectfully, the total number of events of magnitude m and larger that have been detected and otherwise not been detected, in a given space-time volume. When the expected total number of undetected events, larger than m, is expressed as a percentage of the expected number of events with magnitude greater than m, the following metric is obtained :

$$\begin{aligned} &\text{Expected percentage of undetected events greater or equal to } m \\ &= \frac{N_{D^C}(m)}{N_G(m)} \\ &= \frac{N_G(m_{min})\int_m^{m_{max}} f_{M_G}(m)\left(1 - P[M_G \in D|M_G = m]\right)\,dm}{N_G(m_{min})\bar{F}_{G_M}(m)} \\ &= \frac{\int_m^{m_{max}} f_{M_G}(m)\left(1 - P[M_G \in D|M_G = m]\right)\,dm}{\bar{F}_{G_M}(m)} \end{aligned} \tag{10.10}$$

Per definition, for any value of $M_G > m_c$ the probability $P[M_G \in D|M_G = m]$ must be equal to one and therefore the expected percentage of undetected events greater than m must be 0 for all $m > m_c$.

10.3.1 Threshold estimation by Goodness of fit (GOF) method

The following discussion pertains to the results illustrated in Figures E.32 and E.33 of the Section "Estimated proportion of events not detected" in appendix D.1.1.2 (p. 182 & p. 183) as well as Table E.11 (p. 195).

When grouped by expected value of the detection distribution, it can be seen that the percentage of undetected events quickly increases for increasing expected values, except for the case where the probability is modelled by the Log-Normal distribution, where it seems that the percentage of missed events is bound from above at the 0.42 level.

Analysis by variance yields a striking difference between the distributions that are symmetric and those that are not. For non-symmetric distributions, the percentage of missed events stay bounded up to the 0.24 mark, while the figure quickly grows for symmetric distributions, reaching the 0.85 mark in the case of the logistic distribution.

For the non-symmetric distributions, it appears as if the percentage of missed events is slightly higher for positively skewed distributions than for negative distributions.

In general, it can be seen that the method produces a vastly varying percentage of completeness, ranging anywhere between 0.2 and 0.85 depending on the detection distribution.

10.3.2 Threshold estimation by Maximum curvature (MAXC) method

The following discussion pertains to the results illustrated in Figures E.34 and E.35 of the Section "Estimated proportion of events not detected" in appendix D.1.1.2 (p. 184 & p. 184) as well as Table E.11 (p. 195).

As can be seen from the results of the MAXC method, when grouped by expected value of the detection distribution, the method tends to underestimate the level at which only a reasonable number of events are missed. This percentage of events missed steadily increases for increasing expected values, but is bounded from above in the case of the detection probability being modelled by the Log-Normal distribution. In this case it can be seen that the percentage of missed events decrease after a certain point.

Percentage of missed events increase for increasing values of the variance of the detection distribution, but at a considerably faster pace for symmetric distributions than for non-symmetric distributions. The percentage for non-symmetric distributions is bound at 0.28, while symmetric distributions can attain values as high as 0.9.

The percentage of missed events appears to decrease for increased skewness of the Log-Normal distribution, as well as for positive skewness of the Pareto type II distribution. However, it appears that for negative skew of the Pareto type II distribution the percentage of events missed is indeed lower.

In general, the level at which the MAXC method regards a data sample as complete can vary significantly depending on the form of the detection distribution and shows that it is possible to grossly underestimate the level at which reasonable completeness is attained.

10.3.3 Threshold estimation by b-Value stability (bVS) method

The following discussion pertains to the results illustrated in Figures E.36 and E.37 of the Section "Estimated proportion of events not detected" in appendix D.1.1.2 (p. 186 & p. 187) as well as Table E.11 (p. 195).

Results for the bVS method show consistently that the greatest percentage of missed events is possible when utilizing this method, with losses as high as 0.99 in the case of the Log-Normal distribution.

Results from the bVS method, when grouped by expected value of the detection distribution, show that the expected percentage of events missed grows quickly for increasing expected values of the detection distribution. Once again, the method tends to underestimate the level at which only a reasonable number of events are missed. Not even in the case where the detection probability is modelled by the Log-Normal distribution, is the maximum number of events missed bound to a reasonable level.

Percentage of missed events increase for increasing values of the variance of the detection distribution, but at a considerably faster pace for symmetric distributions, including the Log-Normal distribution, than for the non-symmetric Pareto type II distribution.

The percentage of missed events appears to decrease for increasing skewness of the Log-Normal, as well as for positive skewness of the Pareto type II distribution. However, it appears that for negative skew of the Pareto type II distribution the percentage of events missed is indeed lower.

In general, the level at which the bVS method regards a data sample as complete can vary significantly depending on the form of the detection distribution and shows that it is possible to grossly underestimate the level at which reasonable completeness is attained.

10.3.4 Threshold estimation by Modified Entire Magnitude Range (MEMR) method

The following discussion pertains to the results illustrated in Figures E.38 and E.39 of the Section "Estimated proportion of events not detected" in appendix D.1.1.2 (p. 188 & p. 189) as well as Table E.11 (p. 195).
When grouped by expected value of the detection function, it can be seen that the percentage of undetected events quickly increases for increasing expected values. It does appear as if there is an upper bound on the percentage of missed events for all distributions at a level of 0.67.

Analysis by variance yields a striking difference between the distributions that are symmetric and those

that are not. For non-symmetric distributions, the percentage of missed events stay bounded with a maximum value of 0.18, while the figure quickly grows for symmetric distributions, reaching the 0.7 mark for the logistic distribution and 0.53 for the Normal distribution.

For the non-symmetric distributions, the percentage of missed events is slightly higher for positively skewed distributions than for those with negative skew.

In general, it can be seen that the EMR method produces levels of completeness that still vary vastly depending on the detection distribution, but are bounded to a slightly lower level than the previous graphical methods. The error bounds on the percentage of events missed can be seen to be much narrower when compared to previous methods. This method appears to be quite robust when a soft detection threshold is present in the data, since the method consistently yields some of the lowest percentages of events lost as a result of estimation.

10.3.5 Threshold estimation by Median Based Assessment of the Segment-Slope (MBASS) method

The following discussion pertains to the results illustrated in Figures E.40 and E.41 of the Section "Estimated proportion of events not detected" in appendix D.1.1.2 (p. 190 & p. 191) as well as Table E.11 (p. 195).
When grouped by expected value of the detection distribution, it can be seen that the percentage of undetected events quickly increases for increasing expected values of the Normal distribution, as well as for the Logistic distribution. However, only moderate increases are present for the non-symmetric distributions, with the maximum percentage of lost events being 0.45 for the Pareto distribution.

Analysis by variance yields a striking difference between the distributions that are symmetric and those that are not. For non-symmetric distributions, the percentage of missed events stay attain a maximum of 0.16, while the figure quickly grows for symmetric distributions, reaching the 0.55 mark in the case of the Logistic distribution.

For the non-symmetric distributions, it appears as if the percentage of missed events is slightly higher for positively skewed distributions than for those with negative skew.

In general, it can be seen that the MBASS method produces levels of completeness that still vary vastly depending on the detection distribution, but are bounded to a slightly lower level than most of the previous graphical methods. As is the case when estimating sharp magnitude thresholds, the error bounds on the percentage of events missed are very large when compared to other methods. This method consistently produces threshold estimates that result in some of the lowest percentage of events lost.

10.3.6 Threshold estimation by Moment Incorporating Threshold Computation (MITC) method

The following discussion pertains to the results illustrated in Figures E.42 and E.43 of the Section "Estimated proportion of events not detected" in appendix D.1.1.2 (p. 192 & p. 193) as well as Table E.11 (p. 195).

When grouped by expected value of the detection function, it can be seen that the percentage of undetected events quickly increases for increasing expected values. A maximum percentage of undetected events can be seen to be 0.9 in the case of a Logistic distribution.

Analysis by variance yields a striking difference between the distributions that are symmetric and those that are not. For non-symmetric distributions, the percentage of missed events stay bounded well below the 0.2 mark, while the figure quickly grows for symmetric distributions, reaching the 0.72 mark in the case of the logistic distribution.

For the non-symmetric distributions, it appears as if the percentage of missed events is slightly higher for positively skewed distributions than for negatively skewed distributions.

In general, it can be seen that the MITC method produces levels of completeness that vary vastly depending on the detection distribution. The method is ranked at the top of the bottom half of preferable soft threshold estimation methods. This should however not be so unexpected since the method is mathematically constructed around the premise of a sharp detection threshold. Therefore it is not advised that this method be used to gauge the completeness of the this type of sample.

11 Comparitive evaluation of MITC scheme for real-world data

To gauge the efficacy of the newly derived MITC scheme, the threshold of complete reporting will be estimated for real-world earthquake data.

The datasets to be used have also been reviewed by Woessner and Wiemer in their 2005 comparison study [60] and their description of the data is as follows:

Catalogue 1 Regional Catalogue : A subset of the Earthquake Catalogue of Switzerland (ECOS) of the Swiss Seismological Service (SSS) in the southern province Wallis for the period 1992-2002.

Catalogue 2 Volcanic region: A subset of the earthquake catalogue maintained by the National Research Institute for Earthquake Science and Disaster Prevention (NIED). The subset spans the small volcanic region in the Kanto province, Japan, for the period 1992-2002.

The restrictions placed on the data to be extracted from the catalogues is described in Table 11.1.

	Date		Longitude		Latitude		Magnitude		Number of events
	Start	End	Starting Point	Ending Point	Starting Point	Ending Point	Minimum Included	Maximum Recorded	
Catalogue 1	1992/01/13	2002/12/28	6.8 E	8.4 E	45.9 N	46.65 N	0.7	4.6	988
Catalogue 2	1992/01/01	2000/12/28	138.95 E	139.35 E	34.8 N	35.05 N	0	5.1	30,882

Table 11.1: *Restrictions placed on data to be extracted from catalogues. Subset of catalogue to be used for estimation of detection threshold.*

Results for Catalogue 1

Estimation method	$\hat{m}_c$	Sampling Distribution						
		Expected Value	Standard Deviation	Quantiles				
				0%	5%	50%	95%	100%
GOF	1.85	1.92	0.31	1.35	1.55	1.85	2.35	3.25
MAXC	1.45	1.41	0.07	1.25	1.35	1.45	1.45	1.75
bVS	1.30	1.38	0.11	1.30	1.30	1.30	1.50	2.00
EMR	1.90	1.89	0.03	1.80	1.80	1.90	1.90	1.90
MBASS	1.60	1.44	0.17	1.30	1.30	1.40	1.80	3.00
MITC	1.90	2.13	0.46	1.75	1.80	1.90	2.97	4.50

Table 11.2: *Results for various methods used for estimating the detection threshold, m_c, in catalogue 1 (Cat1).*

Results for Catalogue 2

Estimation method	$\hat{m}_c$	Sampling Distribution						
		Expected Value	Standard Deviation	Quantiles				
				0%	5%	50%	95%	100%
GOF	4.05	2.98	0.56	2.25	2.25	2.75	4.05	4.75
MAXC	1.05	1.05	0.00	1.05	1.05	1.05	1.05	1.05
bVS	1.80	1.75	0.74	1.00	1.00	1.80	3.20	4.00
EMR	2.20	2.16	0.05	2.10	2.10	2.20	2.20	2.20
MBASS	1.40	1.44	0.12	1.00	1.30	1.40	1.60	1.80
MITC	1.70	1.70	0.01	1.60	1.70	1.70	1.70	1.70

Table 11.3: *Results for various methods used for estimating the detection threshold, m_c, in catalogue 2 (Cat2).*

Considering the findings relating to question 3, as discussed in Section 3, the prudent approach would be to estimate the threshold by considering the EMR, MBASS and MITC estimation results. Furthermore, by the analysis in Section 10.2.3, the value of a threshold can be determined through simultaneous consideration of the estimation results. Due to poor performance of the GOF, MAXC and bVS methods, the relevant estimation results will only be shown for illustrative purposes, but will be excluded from further considerations.

Catalogue 1 : In general, the MBASS method shows a very small level of bias and through the previous discussion, is seen as a truly robust estimation method. However in this case the MBASS estimate of the threshold agrees with those of the graphical methods, which tend to grossly underestimate the value of the threshold. Furthermore, it can be seen that the EMR and MITC estimates agree and carry weight. The estimated value, erring on the side of caution might be to estimate the threshold within the interval [1.44; 2.36]. This is one standard deviation above and below the MITC estimate.

Catalogue 2 : The picture here is much less clear on how to interpret, since no 2 methods are in agreement. The MBASS method yields an estimate playing to its strength, namely an estimate that is not in the mid-values. Therefore MBASS estimates the threshold to be within the interval [1.22; 1.56] (One standard deviation above and below the MBASS estimate). However, the MITC estimate has a very low standard deviation, yielding an estimate within the interval [1.69; 1.71] (One standard deviation above and below the MITC estimate). A conservative estimate might be in the interval [1.22; 1.71], since the intervals of the MBASS and MITC methods do not overlap.

12 Threshold estimation in operational risk data

An operational risk dataset of the ORX Association (*www.orx.org*) has been utilized in a recent study relating to modelling of losses [57]. This dataset has also been obtained and will aid in the demonstration of establishing threshold values. This operational risk dataset can be obtained by contacting the administrator of the ORX Association website.

As per the description given in the original study, the dataset consists of 1 178 operational risk events that have been publicly reported since 1974. For the current investigation, the only information to be used is the stated loss amount in GBP of the events. A distribution of these losses is shown in Figure 12.1 with a summary given in Table 12.1.

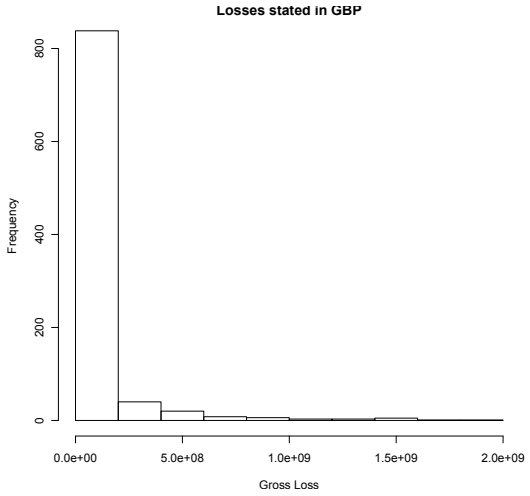


Figure 12.1: *Histogram of operational loss amounts in the considered operational risk database [57]. Amounts are stated in GBP.*

Minimum	1st Quantile	Median	Mean	3rd Quantile	Maximum
60	1 200 000	7 880 000	76 840 000	47 170 000	1 900 000 000

Table 12.1: *Summary of losses encountered in the considered operational risk dataset. Amounts stated in GBP*

As previously referenced, loss data is typically modelled by a power-law, such as that described by a Pareto distribution. For this investigation the Generalized Pareto distribution is chosen, which has a PDF and CDF of the following forms, respectively

$$f_{M_G}(m) = \frac{1}{\beta}\left(1 + \frac{\alpha\,(m-\gamma)}{\beta}\right)^{-\frac{1}{\alpha}-1} \tag{12.1}$$

$$F_{M_G}(m) = \begin{cases} 1 - \left(1 + \frac{\alpha(m-\gamma)}{\beta}\right)^{-\frac{1}{\alpha}} & \text{for } \alpha \neq 0 \\ 1 - \exp\left(-\frac{m-\gamma}{\beta}\right) & \text{for } \alpha = 0 \end{cases} \tag{12.2}$$

An analogous detected severity distribution can be derived for operational risk data that is similar to the detected magnitude distribution for earthquake data. This derivation is based on (8.9) on page 26. In this case the cumulative Normal distribution function has been chosen as the detection probability, i.e. $P[M_G \in D | \{M_G = m\} \cap \{M_G < m_c\}] = \Phi^T(z_m)$. Where $z_m = \dfrac{m-\mu}{\sigma}$ and the superscript T indicates that the Normal distribution is truncated at the point m_c. From this the detected severity distribution can be expressed as :

$$f_{M_D} = C_{Norm} \cdot \begin{cases} f_{M_G}(m) \cdot \Phi^T(z_m) & \text{if } m_{min} \leq m < m_c \\ f_{M_G}(m) & \text{if } m_c \leq m \leq m_{max} \end{cases} \tag{12.3}$$

For the current investigation $\gamma = m_{min} = 0$, $\alpha \neq 0$ and $m_{max} = 1\,900\,000\,000$. The boundaries of the distribution remain to be justified and are open to scrutiny.

Based on the findings throughout this investigation, it has been decided that the threshold estimation methods with graphical backgrounds will not be used to estimate the threshold of complete reporting. Furthermore, the MBASS method shows complications in that the nuisance parameters of the distribution explicitly enter in the calculation of segment slopes. It is left for further investigations to establish consistent manners of incorporating estimates of these parameters when calculating segment slopes. Biases in the nuisance parameter estimates of the EMR method have also lead to this method being excluded from consideration. Therefore, a MITC estimation scheme has been derived, which is analogous to that obtained for the exponential distribution.

A key difference in the current implementation of the estimation scheme is that a heavy reliance has been put on numerical integration. For this reason, as well as the magnitude of events to be directly entered into special functions, such as the CDF of the Normal distribution, numerical volatilities have been noted. Carrying forward it is proposed that investigation into optimization schemes be undertaken with specific applicability to the problem at hand. Bootstrapping of the estimates therefore also presented issues of convergence of results and was therefore not carried out at this point in time.

However, the following estimates have been obtained after empirical establishment of suitable convergence criteria when examining the entire dataset : $\hat{m}_c = 87,357,708$; $\hat{\beta} = 4,059,933$; $\hat{\alpha} = 1.361$; $\hat{\mu} = 13,745,228$ and $\hat{\sigma} = 19,750,341$.

Figure 12.2 shows a histogram of the operational loss events with values greater than the estimated level of complete reporting, $\hat{m}_c$. Furthermore, Figures 12.3 and 12.4a to 12.4d display the progression of estimated parameters as the threshold converges to the end estimation result of $\hat{m}_c = 87,357,708$. These progressions differ greatly from those encountered for previous MITC implementations, which motivate the scrutinization of numerical techniques and computational stability.

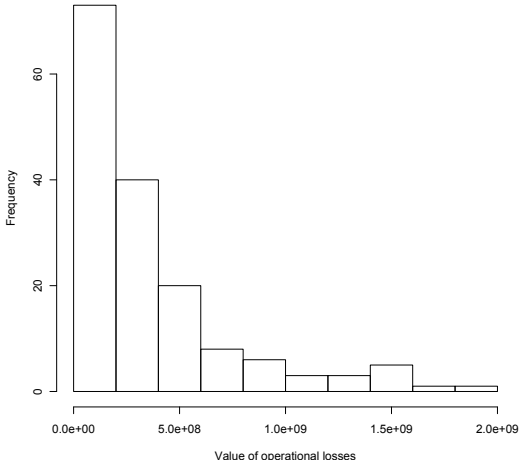


Figure 12.2: *Histogram of operational risk data with event losses greater than the estimated threshold of completeness.*

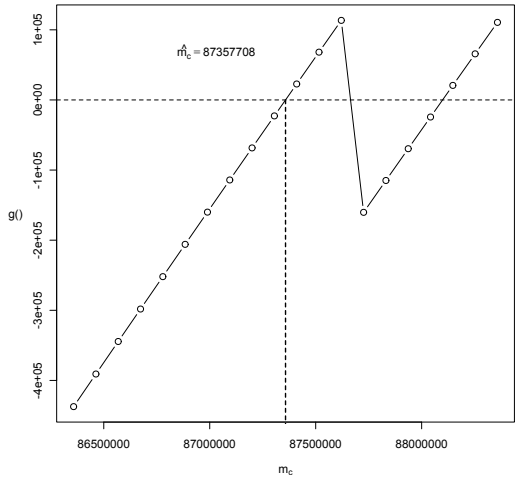
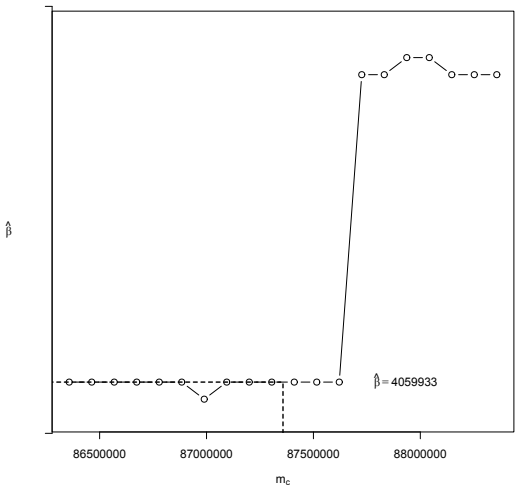
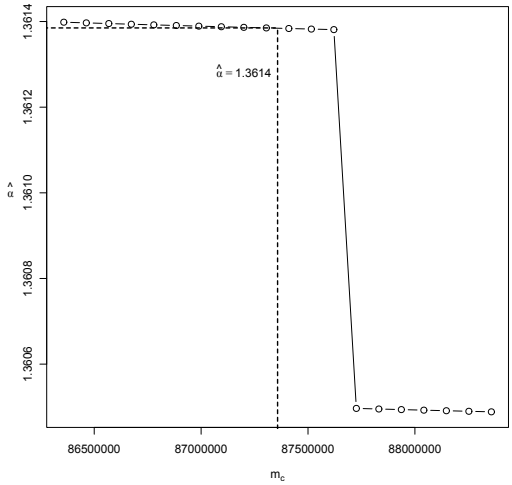


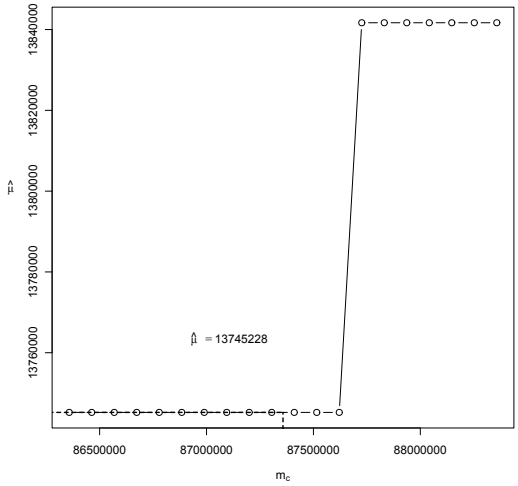
Figure 12.3: *Graphical representation of (9.18) that must be solved for 0 in order to estimate threshold of detection m_c as well as nuisance parameters of detected severity distribution.*



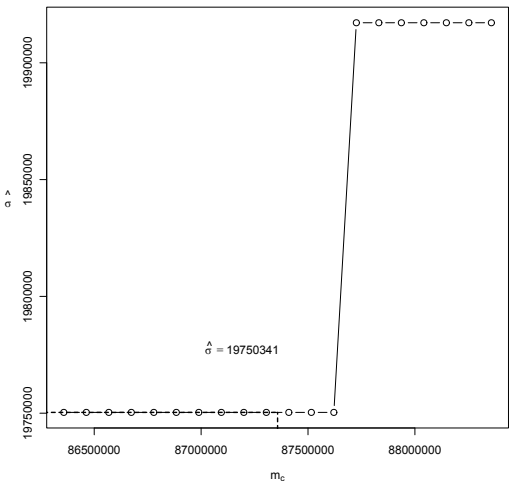
(a) *Progression of $\hat{\beta}$ estimated via ML For differing values of m_c*



(b) *Progression of $\hat{\alpha}$ estimated via ML For differing values of m_c.*



(c) *Progression of $\hat{\mu}$ estimated via ML For differing values of m_c.*



(d) *Progression of $\hat{\sigma}$ estimated via ML For differing values of m_c.*

Figure 12.4: *Maximum likelihood estimation of nuisance parameters of observed magnitude distribution for varying detection threshold values (m_c). Dotted line indicates the value of $\hat{m}_c$ as obtained via MITC.*

13 Conclusion

13.1 Findings on methodology

1. **Soft detection threshold**

   As displayed through this study, in circumstances where the process of data destruction acts upon data over the entire support of the random variable, modelling the data by a sharp-detection threshold model will be a misspecification. Hence, utilizing sharp detection threshold estimation techniques on such data samples, leads to estimates of completeness that generally warrant concern. In such circumstances, all data should be retained and parameters estimated for soft-detection threshold distributions. However, it should be stated that the EMR and MBASS techniques show reasonable promise in estimating a magnitude cut-off for complete reporting in such data.

2. **Sharp detection threshold**

   Based on the results obtained the more recent techniques of threshold estimation (EMR, MBASS and MITC) offer superior performance to the techniques with graphical background. Moving forward, it is recommended that methods with a higher degree of mathematical / statistical background be utilized. This suggestion is based on consideration of the MSE values and progression of constituent components, i.e. the bias and variance of the estimator. Through the simultaneous consideration of these factors a relative picture of the methods tendency to over- / underestimate the value of the threshold can be constructed.

   Based on the relatively small MSE values, the MITC scheme and MBASS method show promise as two of the more robust threshold estimation methods. These findings relate to various functional forms of the detection probability as well as values of the detection threshold. An alternative to be suggested is that the threshold estimation methods be used as a battery of tests to further the understanding of specific detection thresholds in data. From such a point, informed decisions can be made as to the value of a threshold. As illustration, such a deduction can be seen in the previous section.

3. **Operational loss threshold**

   Preliminary investigation into the considered dataset of operational losses, shows that the model has indicated a possible threshold value. This enables a researcher to confidently continue with either an entirely complete dataset (with a subset of loss events with values greater than the indicated threshold), or to develop and fit modified models that incorporates all data and makes adjustments for data above and below the threshold. It has been found that reasonable results have been obtained when assuming a Normal CDF as detection probability, as has been the case for the considered exponentially distributed data.

13.2 Future work to be done

1. **Operational risk data analysis**

   An in-depth investigation of the possible threshold inducing causes must be undertaken to better understand the process of data destruction. This will enable eventual modelling of detected data.

2. **Justify the use of detection probability models**

   For earthquake and operational risk data it must be seen if detection based models offer an improved fit to observed data.

3. **Derive metric to establish statistical significance of detection threshold**

   Apart from the MBASS method, none of the previously discussed methods include a method to gauge whether the estimated threshold of completeness contributes significantly to the analysed dataset. Therefore a general measure must be derived to ascertain whether it is advisable to incorporate a detection based model, or continue with a model assuming complete reporting.

4. **Investigation of optimization methods used for MLE estimates**

   Due to the biases obtained from nuisance parameter estimation for the EMR and MITC methods, it is proposed that the optimization methods be scrutinized by which maximum likelihood estimates are obtained. A particular investigation into reformulation of the distributions with orthogonal parameters can yield greater efficiency when optimizing through more accurate estimation results and, possibly, reduced computation time. Furthermore, to aid in future work on this subject partial derivatives have been obtained for the maximum likelihood function of the detected magnitude distribution where the detection probability is modelled by a cumulative Normal distribution. These equations can be found in appendix F (p. 196).

5. **Extension of estimation methods to include various forms of detection distribution**

   It has been shown that when incorporating the CDF of a Normal random variable in the MITC and EMR methods, relatively robust estimation methods are obtained. However, both of these methods can be reworked so as to incorporate various other functional forms of the detection probability. Such an investigation could yield interesting results pertaining to the biases and variability of resulting estimates.

6. **Combining estimation techniques through stacking**

   As has been proposed, the results from the various threshold estimation techniques can be seen as a battery of tests. This idea can be furthered by combining the more efficient estimation techniques through a process of stacking.

7. **Threshold estimation of Pareto type data**

   It can be seen that such threshold estimation holds considerable challenges not encountered in exponentially distributed data. However, threshold estimates have been obtained and further work

to be done on this subject must focus on refining of results. The validity of the Normal CDF as a detection probability must also be scrutinized. Furthermore, it is proposed that non-parametric methods, such as the MBASS method be adapted to take account of the Pareto distribution. This will have the obvious consequence of not being as computationally intensive while avoiding potential numerical instabilities during the estimation process.

14 Bibliography

[1] Aki, K. Maximum likelihood estimate of b in the formula log n = a - bm and its confidence limits. *Bull. Earthquake Res. Inst. Tokyo Univ. 43* (1965), 237–239.

[2] Aki, K. Magnitude-frequency relation for small earthquakes: A clue to the origin of f max of large earthquakes. *J. Geophys. Res 92(B2)* (1987), 13491355.

[3] Amorèse, D. Applying a change-point detection method on frequency-magnitude distributions. *Bulletin of the Seismological Society of America 97*, 5 (2007), 1742–1749.

[4] Bak, P., Tang, C., and Wiesenfeld, K. Self-organized criticality: An explanation of 1/f noise. *Physical Review Letters 59(4)* (1987), 381–384. http://dx.doi.org/10.1103/PhysRevLett.59.381.

[5] Bauwens, L., and Rombouts, J. V. On marginal likelihood computation in change-point models. Cahiers de recherche 0942, CIRPEE, 2009.

[6] Brent, R. *Algorithms for Minimization without Derivatives.* 1973.

[7] Brillinger, D. R. Earthquake risk and insurance. *Environmetrics 4*, 1 (1993), 1–21.

[8] Cao, A. M., and Gao, S. S. *Temporal variation of seismic b-values beneath northeastern Japan island arc.* Geophys. Res. Lett., 2002.

[9] Chen, J., and Gupta, A. K. *Parametric Statistical Change-point Analysis.* Birkhuser, 2000.

[10] Chernobai, A., Menn, C., Moscadelli, M., Rachev, S., and Trueck, S. Treatment of incomplete data in the field of operational risk: The effects on parameter estimates, el and ul figures. In *The Advanced Measurement Approach to Operational Risk.* Risk Books, Spain, 2006, pp. 145–168.

[11] Chernobai, A., Rachev, S., and Fabozzi, F. *Operational Risk: A Guide to Basel II Capital Requirements, Models, and Analysis.* Frank J. Fabozzi Series. Wiley, 2008.

[12] Coles, S. *An Introduction to Statistical Modeling of Extreme Values.* Lecture Notes in Control and Information Sciences. Springer, 2001.

[13] Committee, S. *Chronological Tables of Earthquake Data of China.* ACADEMIA SINICA, 1956. (in Chinese).

[14] Cosentino, P., Ficara, V., and Luzio, D. Truncated exponential frequency-magnitude relationship in the earthquake statistics. *Bull. Seismol. Am. 67* (1977), 16151623.

[15] Field, D., Jackson, D., and Dolan, J. A mutually consistent seismic-hazard source model for southern california. *Bull. Seismol. Soc. Am. 89* (1999), 559578.

[16] FUNCTIONS WRITTEN BY JANET E. HEFFERNAN WITH R PORT, O. S., AND DOCUMENTATION PROVIDED BY ALEC G. STEPHENSON., R. *ismev: An Introduction to Statistical Modeling of Extreme Values*, 2012. R package version 1.39.

[17] GAHALAUT, K., GAHALAUT, V., AND PANDEY, M. A new case of reservoir triggered seismicity: Govind ballav pant reservoir (rihand dam), central india. *Tectonophysics 439*, 14 (2007), 171 – 178.

[18] GIAMPICCOLO, E., DAMICO, S., PATAN, D., AND GRESTA, S. Attenuation and source parameters of shallow microearthquakes at mt. etna volcano, italy. *Bulletin of the Seismological Society of America 97*, 1B (2007), 184–197.

[19] GIRAUDOUX, P. *pgirmess: Data analysis in ecology*, 2013. R package version 1.5.7.

[20] GRADSHTEYN, I. S., AND RYZHIK, I. M. *Table of Integrals, Series, and Products, Seventh Edition*, seventh ed. Academic Press, Mar. 2007.

[21] GREGORIOU, G. *Operational Risk Toward Basel III: Best Practices and Issues in Modeling, Management, and Regulation*. Wiley Finance. Wiley, 2009.

[22] GUTENBERG, B., AND RICHTER, C. F. Frequency of earthquakes in california. *Bulletin of the Seismological Society of America 34*, 4 (1944), 185–188.

[23] HABERMANN, R. E. Man-made changes of seismicity rates. *Bulletin of the Seismological Society of America 77*, 1 (1987), 141–159.

[24] HABERMANN, R. E., AND WYSS, M. Background seismicity rates and precursory quiescence: Imperial valley, california. *Bull. Seism. Soc. Am. 74* (1984), 1743–1756.

[25] HABERMANN, R. E., AND WYSS, M. Seismic quiescence and earthquake prediction on the calaveras fault, california. *Trans. Am. Geophys. Union 65* (1984), 988.

[26] HAMILTON, R. Mean magnitude of an earthquake sequence. *Bull. Seismol. Soc. Am. 57* (1967), 11151126.

[27] HOLLANDER, M., AND WOLFE, D. *Nonparametric Statistical Analysis,*. Wiley, New York, 1973.

[28] HOTHORN, T., BRETZ, F., AND WESTFALL, P. Simultaneous inference in general parametric models. *Biometrical Journal 50*, 3 (2008), 346–363.

[29] HOTHORN, T., HORNIK, K., VAN DE WIEL, M. A., AND ZEILEIS, A. A lego system for conditional inference. *The American Statistician 60*, 3 (2006), 257–263.

[30] HOTHORN, T., HORNIK, K., VAN DE WIEL, M. A., AND ZEILEIS, A. Implementing a class of permutation tests: The coin package. *Journal of Statistical Software 28*, 8 (2008), 1–23.

[31] Jobst, A. Operational risk-the sting is still in the tail but the poison depends on the dose. IMF Working Papers 07/239, International Monetary Fund, Oct. 2007.

[32] Kijko, A. Estimation of the maximum earthquake magnitude, m_{max}. *Pure and Applied Geophysics 161* (2004), 1655–1681. 10.1007/s00024-004-2531-4.

[33] Lanzante, J. R. Resistant, robust and non-parametric techniques for the analysis of climate data: theory and examples, including applications to historical radiosonde station data. *International Journal for Climatology 16* (1996), 11971226.

[34] Lavielle, M., and Teyssire, G. Detection of multiple change-points in multivariate time series. *Lithuanian Mathematical Journal 46* (2006), 287–306. 10.1007/s10986-006-0028-9.

[35] Lee, W. H. K., and Brillinger, D. R. On chinese earthquake history - an attempt to model an incomplete data set by point process analysis. *Birkhuser Verlag, Basel 117* (1979).

[36] Lu, C., and Vere-Jones, D. Statistical analysis of synthetic earthquake catalogs generated by models with various levels of fault zone disorder. *Journal of Geophysical Research* (2001).

[37] Maechler, M. *Rmpfr: R MPFR - Multiple Precision Floating-Point Reliable*, 2013. R package version 0.5-2.

[38] Main, I. A characteristic earthquake model of the seismicity preceding the eruption of the mount st helens on 18 may 1980. *Phys. Earth Planet. Interiors 49* (1987), 283–293.

[39] Main, I. Apparent breaks in scaling in the earthquake cumulative frequency-magnitude distribution: Fact or artifact? *Bulletin of the Seismological Society of America 90*, 1 (2000), 86–97.

[40] Main, I. G. Earthquakes as critical phenomena: Implications for probabilistic seismic hazard analysis. *Bulletin of the Seismological Society of America 85*, 5 (1995), 1299–1308.

[41] Mann, H. B., and Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics 18* (1947), 50–60.

[42] Ogata, Y., and Katsura, K. Analysis of temporal and spatial heterogeneity of magnitude frequency distribution inferred from earthquake catalogs. *Geophys. J. Int. 113* (1993), 727738.

[43] Page, R. Aftershocks and microaftershocks of the great alaska earthquake of 1964. *Bulletin of the Seismological Society of America 58*, 3 (1968), 1131–1168.

[44] Peters, G. W., Johansen, A. M., and Doucet, A. Simulation of the annual loss distribution in operational risk via panjer recursions and volterra integral equations for value at risk and expected shortfall estimation. Tech. Rep. 3, 2007.

[45] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2011. ISBN 3-900051-07-0.

[46] Rydelek, P. A., and Sacks, I. S. Testing the completeness of earthquake catalogs and the hypothesis of self-similarity. *Nature 337*, 6204 (1989), 251–253.

[47] Saunders, I., M. B. J. S. D. R., and Kijko, A. The south african national seismograph network. *Seismol. Research Letters 79*, 2 (2008), 203–210.

[48] Schorlemmer, D., and Woessner, J. Probability of detecting an earthquake. *Bulletin of the Seismological Society of America 98*, 5 (2008), 2103–2117.

[49] Selvaggi, M. Analysing operational losses in insurance. *Report from ABI Research Department.* (2009). ISBN 978 190319346-X.

[50] Shi, Y., and Bolt, B. A. The standard error of the magnitudefrequency b value. *Bulletin of the Seismological Society of America 72* (1982), 16771687.

[51] Siegel, S., and Castellan, N. *Nonparametric Statistics for the Behavioral Sciences.* McGraw-Hill, New York, 1988.

[52] Speidel, D. H., and Mattson, P. H. The polymodal frequency-magnitude relationship of earthquakes. *Bulletin of the Seismological Society of America 83*, 6 (1993), 1893–1901.

[53] Stein, R., and Hanks, T. $m \geq 6$ earthquakes in southern california during the twentieth century: No evidence for a seismicity or moment deficit. *Bull. Seismol. Soc. Am. 88* (1998), 635652.

[54] Stephenson, A. G. evd: Extreme value distributions. *R News 2*, 2 (June 2002), 0.

[55] Taylor, D. W. A., Snoke, J. A., Sacks, I. S., and Takanami, T. Nonlinear frequency-magnitude relationships for the hokkaido corner, japan. *Bulletin of the Seismological Society of America 80*, 2 (1990), 340–353.

[56] Utsu, T. A method for determining the value of b in the formula $\log n = a - bm$ showing the magnitude-frequency relation for earthquakes. *Geophys. Bull. Hokkaido Univ. 13* (1965), 99103. (in Japanese with English summary).

[57] Valrie Chavez-Demoulin, Paul Embrechts, M. H. An extreme value approach for modeling operational risk losses depending on covariates. Working paper, 2013/04/09.

[58] WGCEP. Seismic hazard in southern california: Probable earthquakes, 1994 to 2024. *Bull. Seism. Soc. Am. 85* (1995), 379439. (Working Group on Central California Earthquake Probabilities).

[59] Wiemer, S., and Wyss, M. Minimum magnitude of completeness in earthquake catalogs: Examples from alaska, the western united states, and japan. *Bulletin of the Seismological Society of America 90*, 4 (2000), 859–869.

[60] Woessner, J., and Wiemer, S. Assessing the quality of earthquake catalogues: Estimating the magnitude of completeness and its uncertainty. *Bulletin of the Seismological Society of America 95*, 2 (2005), 684–698.

[61] Wolfram Research, I. *Mathematica Edition: Version 8.0*. Wolfram Research, Inc, Champaign, Illinois, 2010.

[62] Zhang, J.-Z., and Song, L.-Y. On the method of estimating b value and its standard error. *Acta Seismologica Sin. 3* (1981), 292–301.

Appendicies

Appendix A : Formulation of detected magnitude distribution

Before continuing it is worth mentioning the following general result:

Lemma A.1 *By repeatedly utilizing the definition of conditional probability for general events A, B as well as C and the complementary event C^C, $P[A|B] = \frac{P[A\cap B]}{P[B]}$, the following result can be derived*

$$\frac{P[A\cap B\cap C]+P[A\cap B\cap C^c]}{P[B]} = \frac{P[A|B\cap C]P[B\cap C]+P[A|B\cap C^c]P[B\cap C^c]}{P[B]}$$

$$= P[A|B\cap C]P[C|B]+P[A|B\cap C^c]P[C^c|B] \tag{A.1}$$

□

For notational convenience the event $\{m \leq M_G < m+h\}$ is defined as $M_G^h(m)$. Continuing, the conditional probability in (8.6) (p. 24) can expressed as

$$P[M_G \in D \mid m \leq M_G < m+h] \tag{A.2}$$

$$= P[M_G \in D|M_G^h(m)] \tag{A.3}$$

$$= \frac{P[\{M_G \in D\}\cap M_G^h(m)]}{P[M_G^h(m)]} \tag{A.4}$$

$$= \frac{P[\{M_G \in D\}\cap M_G^h(m)\cap\{\{M_G < m_c\}\cup\{M_G \geq m_c\}\}]}{P[M_G^h(m)]} \tag{A.5}$$

$$= \frac{P[\{\{M_G \in D\}\cap M_G^h(m)\cap\{M_G < m_c\}\}\cup\{\{M_G \in D\}\cap M_G^h(m)\cap\{M_G \geq m_c\}\}]}{P[M_G^h(m)]} \tag{A.6}$$

$$= \frac{P[\{M_G \in D\}\cap M_G^h(m)\cap\{M_G < m_c\}]+P[\{M_G \in D\}\cap M_G^h(m)\cap\{M_G \geq m_c\}]}{P[M_G^h(m)]} \tag{A.7}$$

Since $\{M_G < m_c\}\cap\{M_G \geq m_c\} = \emptyset$. Furthermore, utilizing Lemma (A.1) with $A = \{M_G \in D\}$, $B = M_G^h(m)$ and $C = \{M_G < m_c\}$, the above can be written as

$$= P[M_G \in D|M_G^h(m)\cap\{M_G < m_c\}]P[M_G < m_c|M_G^h(m)]$$

$$+P[M_G \in D|M_G^h(m)\cap\{M_G \geq m_c\}]P[M_G \geq m_c|M_G^h(m)] \tag{A.8}$$

Upon taking the limit as in (8.6) the following is obtained

$$\lim_{h\to 0} P[M_G \in D|M_G^h(m)] = P[M_G \in D|\{M_G = m\} \cap \{M_G < m_c\}]P[M_G < m_c|M_G = m]$$
$$+ P[M_G \in D|M_G^h(m) \cap \{M_G \geq m_c\}]P[M_G \geq m_c|M_G^h(m)] \quad \text{(A.9)}$$

and rearranging

$$P[M_G \in D|M_G^h(m)] = P[M_G \in D|M_G^h(m) \cap \{M_G < m_c\}]P[M_G < m_c|M_G^h(m)]$$
$$+P[M_G \in D|M_G^h(m) \cap \{M_G \geq m_c\}]P[M_G \geq m_c|M_G^h(m)] \quad \text{(A.10)}$$

yields

$$\begin{aligned} P[M_G \in D|M_G = m] &= \lim_{h\to 0} P[M_G \in D|m \leq M_G < m+h] \\ &= P[M_G \in D|\{M_G = m\} \cap \{M_G < m_c\}] \lim_{h\to 0} P[M_G < m_c|M_G^h(m)] \\ &\quad +P[M_G \in D|\{M_G = m\} \cap \{M_G \geq m_c\}] \lim_{h\to 0} P[M_G \geq m_c|M_G^h(m)] \quad \text{(A.11)} \end{aligned}$$

Of interest to us is the two limiting factors in (A.11). Continuing, the probability found in the first limit of the above expression can be expanded

$$P[M_G < m_c|M_G^h(m)] = \frac{P[\{M_G < m_c\} \cap \{m < M_G \leq m+h\}]}{P[m < M_G \leq m+h]} \quad \text{(A.12)}$$

$$\{M_G < m_c\} \cap \{m < M_G \leq m+h\} = \begin{cases} \{m < M_G \leq m+h\} & \text{if } m < m+h \leq m_c \\ \{m < M_G \leq m_c\} & \text{if } m < m_c \leq m+h \\ \emptyset & \text{if } m_c \leq m < m+h \end{cases}$$

$$P[M_G < m_c|M_G^h(m)] = \begin{cases} \frac{P[m<M_G\leq m+h]}{P[m<M_G\leq m+h]} & \text{if } m < m+h \leq m_c \\ \frac{P[m<M_G\leq m_c]}{P[m<M_G\leq m+h]} & \text{if } m < m_c \leq m+h \\ \frac{P[\emptyset]}{P[m<M_G\leq m+h]} & \text{if } m_c \leq m < m+h \end{cases}$$

$$= \begin{cases} 1 & \text{if } m < m+h \leq m_c \\ \frac{P[m<M_G\leq m_c]}{P[m<M\leq m+h]} & \text{if } m < m_c \leq m+h \\ 0 & \text{if } m_c \leq m < m+h \end{cases}$$

and $\lim_{h\to 0} \frac{P[m < M_G \leq m_c]}{P[m < M_G \leq m+h]} = 1$ if $m < m_c \leq m+h$ since $m < m_c \leq m+h \to m = m_c$ as $h \to 0$

Therefore it can be seen that the first limiting factor in (A.11) reduces to

$$\lim_{h\to 0} P[M_G < m_c | M_G^h(m)] = I_{[m_{min},m_c)}(m) = \begin{cases} 1 & \text{if } m \in [m_{min}, m_c) \\ 0 & \text{if } m \notin [m_{min}, m_c) \end{cases}$$

where $I_A(x)$ represents the indicator function that is 1 if $x \in A$ and 0 otherwise. Similarly, it can be shown that the second limiting factor in (A.11) reduces to

$$\lim_{h\to 0} P[M_G \geq m_c | M_G^h(m)] = I_{[m_c,m_{max}]}(m) = \begin{cases} 1 & \text{if } m \in [m_c, m_{max}] \\ 0 & \text{if } m \notin [m_c, m_{max}] \end{cases}$$

Therefore (A.11) reduces to

$$\begin{aligned} P[M_G \in D | M_G = m] &= P[M_G \in D | \{M_G = m\} \cap \{M_G < m_c\}] \cdot I_{[m_{min},m_c)}(m) \\ &\quad + P[M_G \in D | \{M_G = m\} \cap \{M_G \geq m_c\}] \cdot I_{[m_c,m_{max}]}(m) \\ &= \begin{cases} P[M_G \in D | \{M_G = m\} \cap \{M_G < m_c\}] & \text{if } m \in [m_{min}, m_c) \\ P[M_G \in D | \{M_G = m\} \cap \{M_G \geq m_c\}] & \text{if } m \in [m_c, m_{max}] \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{A.13}$$

The case where $m \in [m_c, m_{max}]$ in (A.13) can be expanded as follows

$$\begin{aligned} \lim_{h\to 0} P[M_G \in D | M_G^h(m) \cap \{M_G \geq m_c\}] &= \lim_{h\to 0} \frac{P[\{M_G \in D\} \cap M_G^h(m) \cap \{M_G \geq m_c\}]}{P[M_G^h(m) \cap \{M_G \geq m_c\}]} \\ &= \lim_{h\to 0} \frac{P[\{M_G \in D\} \cap M_G^h(m) | M_G \geq m_c] \cdot P[M_G \geq m_c]}{P[M_G^h(m) | M_G \geq m_c] \cdot P[M_G \geq m_c]} \\ &= \lim_{h\to 0} \frac{P[M_G \in D | M_G \geq m_c] P[M_G^h(m) | M_G \geq m_c]}{P[M_G^h(m) | M_G \geq m_c]} \qquad \text{(A.14)} \\ &= P[M_G \in D | M_G \geq m_c] \\ &= 1 \qquad \text{(A.15)} \end{aligned}$$

Combining this result with (A.13), the following is obtained

$$f_{D|M_G}(m) = \begin{cases} P[M_G \in D | \{M_G = m\} \cap \{M_G < m_c\}] & \text{if } m \in [m_{min}, m_c) \\ 1 & \text{if } m \in [m_c, m_{max}] \\ 0 & \text{otherwise} \end{cases} \tag{A.16}$$

which can be incorporated into (8.7) to obtain

$$f_{M_D}(m) = \begin{cases} f_{M_G}(m) \cdot P[M_G \in D | \{M_G = m\} \cap \{M_G < m_c\}] & \text{if } m \in [m_{min}, m_c) \\ f_{M_G}(m) & \text{if } m \in [m_c, m_{max}] \\ 0 & \text{otherwise} \end{cases} \tag{A.17}$$

this gives an expression for the detected magnitude distribution $f_{M_D}(m)$.

Appendix B : Mathematical results

B.1.1 Special functions

1. **Gamma function**

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\, dt \tag{B.18}$$

2. **Lower incomplete gamma function**

$$\gamma(s, x) = \int_0^x t^{s-1} e^{-t}\, dt \tag{B.19}$$

3. **Incomplete gamma function with respect to a range**

This function has been defined in this study for ease of reference. It can be seen that the main method of differentiation between (B.19) and the *incomplete gamma function with respect to a range* is the number of parameters of the function.

$$\begin{aligned}\gamma(s; a, b) &= \int_a^b t^{s-1} e^{-t}\, dt \\ &= \gamma(s, b) - \gamma(s, a)\end{aligned} \tag{B.20}$$

4. **Upper incomplete gamma function**

$$\Gamma(s, x) = \int_x^\infty t^{s-1} e^{-t}\, dt \tag{B.21}$$

5. **Gaussian density function**

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \qquad z \in \mathbb{R} \tag{B.22}$$

6. **Gaussian distribution function**

$$\Phi(z) = \int_{-\infty}^z \phi(t)\, dt \qquad \text{with } \phi(t) \text{ the Gausian density function as in (B.22).} \tag{B.23}$$

7. **Normal inverse cumulative distribution function**

$$\Phi^{-1}(p) = z \qquad z \in \mathbb{R}\ ,\ p \in [0,1]$$

such that $\Phi(z) = p$ where $\Phi(z)$ the Gaussian distribution function as in (B.23).

(B.24)

8. **Indicator function**

$$I_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases} \qquad \text{Where } x \in \mathbb{R}$$

and A some real valued set.

(B.25)

9. **Gauss hypergeometric function**

$${}_2F_1(a,b;c;z) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \cdot \frac{z^n}{n!} \qquad \text{for } |z| < 1$$

(B.26)

Where $(q)_n$ is the rising Pochhammer symbol which is defined as :

$$(q)_n = \begin{cases} 1 & \text{if } n = 0 \\ q(q+1)\dots(q+n-1) & \text{if } n > 0 \end{cases}$$

(B.27)

B.1.2 Useful Identities

1. **Lower incomplete gamma function**

$$\gamma(r+1;x) = r!\left(1 - e^{-x}\sum_{k=0}^{r}\frac{x^k}{k!}\right) \qquad \text{for } r = [0,1,2,\ldots]$$

Equation 8.352 from Table of Integrals, Series and Products [20].

(B.28)

2. **Incomplete Gamma function subject to range**

$$\begin{aligned}\gamma(r+1;a;b) &= \gamma(r+1;b) - \gamma(r+1;a)\\ &= r!\left(1 - e^{-b}\sum_{k=0}^{r}\frac{b^k}{k!}\right) - r!\left(1 - e^{-a}\sum_{k=0}^{r}\frac{a^k}{k!}\right) \qquad \text{for } r = [0,1,2,\ldots] \quad \text{from (B.28)}\\ &= r!\sum_{k=0}^{r}\frac{a^k e^{-a} - b^k e^{-b}}{k!}\end{aligned}$$

(B.29)

B.1.3 Frequently evaluated integrals

1.

$$\int_a^b m^r f_{M_G}(m)\,dm \qquad \text{with } f_{M_G}(m) \text{ as in (8.2).} \tag{B.30}$$

$$= \frac{\beta}{e^{-\beta m_{min}} - e^{-\beta m_{max}}} \int_a^b m^r e^{-\beta m}\,dm$$

By letting $\beta m = t$ and after subsequent simplification the following is obtained

$$= \frac{\beta^{-r}}{e^{-\beta m_{min}} - e^{-\beta m_{max}}} \gamma\left(r+1; \beta a, \beta b\right)$$

Where $\gamma(x, a, b)$ is the incomplete gamma function with respect to a range (B.20).

$$= \frac{1}{e^{-\beta m_{min}} - e^{-\beta m_{max}}} r! \sum_{k=0}^{r} \beta^{k-r} \frac{a^k e^{-\beta a} - b^k e^{-\beta b}}{k!} \qquad \text{for } r = [0, 1, 2, \ldots]$$

By using (B.29) and subsequent simplification.

□

2.

$$\int m f_{M_G}(m)\,dm \qquad \text{with } f_{M_G}(m) \text{ as in (8.2).} \tag{B.31}$$

This can be seen to be a restatement of (B.30) as an indefinite integral for the specific case $r = 1$. Utilizing integration by parts it can easily be seen that the expression simplifies to

$$\begin{aligned}\int m f_{M_G}(m)\,dm &= \frac{-e^{-\beta m}(m + \frac{1}{\beta})}{e^{-\beta m_{min}} - e^{-\beta m_{max}}} \\ &= -f_{M_G}(m)\frac{1}{\beta}\left(m + \frac{1}{\beta}\right)\end{aligned}$$

□

3.

$$\int_a^b e^{-\beta x} \frac{1}{\sigma}\phi(z_x)\,dx$$

$$\text{where } z_x = \frac{x-\mu}{\sigma}; \quad \mu \in \mathbb{R}; \quad \beta, \sigma > 0 \quad \text{and } 0 < a \leq b. \tag{B.32}$$

$$= \int_a^b \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2 - \beta x\right) dx$$

$$= \int_a^b \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(x^2 - 2x\mu + \mu^2 + 2\sigma^2\beta x)\right) dx$$

$$= \int_a^b \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2\beta))^2 - \beta\left(\mu - \frac{1}{2}\sigma^2\beta\right)\right) dx \quad \text{by completing the square.}$$

$$= e^{-\beta\left(\mu-\frac{1}{2}\sigma^2\beta\right)} \int_a^b \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2\beta))^2\right) dx$$

The expression inside the integral represents the PDF of a Normally distributed random variable with mean $(\mu - \sigma^2\beta)$ and variance σ^2. Therefore the variable $z_x^* = \dfrac{x-(\mu-\sigma^2\beta)}{\sigma}$ is defined.

$$= e^{-\beta\left(\mu-\frac{1}{2}\sigma^2\beta\right)} \Phi\left(z_x^*\right)\Big|_a^b$$

$$= e^{-\beta\left(\mu-\frac{1}{2}\sigma^2\beta\right)} \left(\Phi\left(z_b^*\right) - \Phi\left(z_a^*\right)\right)$$

□

4.

$$\int_a^b f_{M_G}(m)\frac{1}{\sigma}\phi(z_m)\,dm \tag{B.33}$$

where $z_m = \frac{m-\mu}{\sigma}$; $\mu \in \mathbb{R}$; $\sigma > 0$; $0 < m_{min} \leq a \leq b \leq m_{max}$ and $f_{M_G}(m)$ is defined in (8.2) on page 23.

$$= \int_a^b \frac{\beta\exp\left(-\beta m\right)}{\exp\left(-\beta m_{min}\right) - \exp\left(-\beta m_{max}\right)}\frac{1}{\sigma}\phi(z_m)\,dm$$

$$= \frac{\beta}{e^{-\beta m_{min}} - e^{-\beta m_{max}}} \int_a^b \exp\left(-\beta m\right)\frac{1}{\sigma}\phi(z_m)\,dm$$

$$= \frac{\beta}{e^{-\beta m_{min}} - e^{-\beta m_{max}}} \int_a^b \exp\left(-\beta m\right)\frac{1}{\sigma}\phi(z_m)\,dm$$

$$= \frac{\beta e^{-\beta\left(\mu-\frac{1}{2}\sigma^2\beta\right)}}{e^{-\beta m_{min}} - e^{-\beta m_{max}}} \left(\Phi\left(z_b^*\right) - \Phi\left(z_a^*\right)\right)$$

By utilizing (B.32) and defining $z_m^* = \dfrac{m-(\mu-\sigma^2\beta)}{\sigma}$

□

5.

$$\int_a^b F_{M_G}(m)\frac{1}{\sigma}\phi(z_m)\,dm \tag{B.34}$$

where $z_m = \frac{m-\mu}{\sigma}$; $\mu \in \mathbb{R}$; $\sigma > 0$; $0 < m_{min} \leq a \leq b \leq m_{max}$ and $F_{M_G}(m)$ is defined in (8.4) on page 23.

$$= \int_a^b \frac{1-\exp\left(-\beta(m-m_{min})\right)}{1-\exp\left(-\beta(m_{max}-m_{min})\right)}\frac{1}{\sigma}\phi(z_m)\,dm$$

$$= \frac{1}{1-\exp\left(-\beta(m_{max}-m_{min})\right)}\left(\int_a^b \frac{1}{\sigma}\phi(z_m)\,dm - e^{\beta m_{min}}\int_a^b e^{-\beta m}\frac{1}{\sigma}\phi(z_m)\,dm\right)$$

$$= \frac{1}{1-\exp\left(-\beta(m_{max}-m_{min})\right)}\left(\Phi(z_b)-\Phi(z_a) - e^{-\beta\left(\mu-m_{min}-\frac{1}{2}\sigma^2\beta\right)}\left(\Phi\left(z_b^*\right)-\Phi\left(z_a^*\right)\right)\right)$$

By utilizing (B.32) and rearranging, also by defining $z_m^* = \dfrac{m-(\mu-\sigma^2\beta)}{\sigma}$

□

6.

$$\int_a^b f_{M_G}(m)\Phi(z_m)\,dm \tag{B.35}$$

where $z_m = \frac{m-\mu}{\sigma}$; $\mu \in \mathbb{R}$; $\sigma > 0$; $0 < m_{min} \leq a \leq b \leq m_{max}$ and $F_{M_G}(m)$ is defined in (8.4) on page 23.

$$= F_{M_G}(m)\Phi(z_m)\Big|_{m=a}^{m=b} - \int_a^b F_{M_G}(m)\frac{1}{\sigma}\phi(z_m)\,dm$$

Integration by parts.

$$= F_{M_G}(b)\Phi(z_b) - F_{M_G}(a)\Phi(z_a) - \frac{1}{1-\exp\left(-\beta(m_{max}-m_{min})\right)}\left(\Phi(z_b)-\Phi(z_a) - e^{-\beta\left(\mu-m_{min}-\frac{1}{2}\sigma^2\beta\right)}\left(\Phi\left(z_b^*\right)-\Phi\left(z_a^*\right)\right)\right)$$

By utilizing (B.34).

$$= \frac{-1}{\beta}\left(\Phi(z_b)f_{M_G}(b)-\Phi(z_a)f_{M_G}(a)\right) - \frac{e^{-\beta\left(\mu-\frac{1}{2}\sigma^2\beta\right)}}{e^{-\beta m_{min}}-e^{-\beta m_{max}}}\left(\Phi\left(z_b^*\right)-\Phi\left(z_a^*\right)\right)$$

where $z_m^* = \dfrac{m-(\mu-\sigma^2\beta)}{\sigma}$

□

7.

$$\int_{-\infty}^{a} x \frac{1}{\sigma}\phi(z_x)\,dx \tag{B.36}$$

where $z_x = \frac{x-\mu}{\sigma}$; $\mu \in \mathbb{R}$; $\sigma > 0$.

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{a} x \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

if the substituition $z = \dfrac{x-\mu}{\sigma}$ is made, it holds that $x = \sigma z + \mu$ and therefore $dx = \sigma\, dz$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{a-\mu}{\sigma}} (\sigma z + \mu) \exp\left(-\frac{z^2}{2}\right) dz$$

$$= \mu\Phi\left(\frac{a-\mu}{\sigma}\right) + \sigma \int_{-\infty}^{\frac{a-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} z \exp\left(-\frac{z^2}{2}\right) dz$$

by letting $u = \dfrac{z^2}{2}$ it holds that $z\,dz = du$

$$= \mu\Phi\left(\frac{a-\mu}{\sigma}\right) + \sigma \frac{1}{\sqrt{2\pi}} \int_{\infty}^{\frac{1}{2}\left(\frac{a-\mu}{\sigma}\right)^2} e^{-u}\,du$$

$$= \mu\Phi\left(z_a\right) - \sigma\phi\left(z_a\right)$$

□

8.

$$\int e^{ax} \tanh(bx)\,dx \tag{B.37}$$

(B.38)

$$= \frac{e^{(a+2b)x}}{(a+2b)} {}_2F_1\left(1 + \frac{a}{2b}, 1; 2 + \frac{a}{2b}; -e^{2bx}\right) - \frac{1}{a} e^{ax} {}_2F_1\left(1, \frac{a}{2b}; 1 + \frac{a}{2b}; -e^{2bx}\right) + K$$

; for some constant $K \in \mathbb{R}$ and ${}_2F_1$ the Gauss hyper-geometric function.

Evaluated using Mathematica symbolic integration [61]

□

9.

$$\int e^{-\beta(m-m_{min})} \tanh\left(\frac{m-\mu}{2s}\right) dm \tag{B.39}$$

(B.40)

Simplification

By making the substitution $x = \left(\frac{m-\mu}{2s}\right) \Rightarrow m = 2sx + \mu$ and $dm = 2s\,dx$ is is found that

$$\int e^{-\beta(m-m_{min})} \tanh\left(\frac{m-\mu}{2s}\right) dm \tag{B.41}$$

$$= 2s \int e^{-\beta(2sx+\mu-m_{min})} \tanh(x)\, dx$$

$$= 2se^{-\beta(\mu-m_{min})} \int e^{-2\beta sx} \tanh(x)\, dx$$

with $a = -2\beta s$ and $b = 1$, the results of integral (B.38) can be used

$$= 2se^{-\beta(\mu-m_{min})} \left(\frac{e^{(-2\beta s+2)x}}{(-2\beta s+2)} {}_2F_1\left(1 + \frac{-2\beta s}{2}, 1; 2 + \frac{-2\beta s}{2}; -e^{2x}\right) \right.$$

$$\left. - \frac{1}{-2\beta s} e^{-2\beta sx} {}_2F_1\left(1, \frac{-2\beta s}{2}; 1 + \frac{-2\beta s}{2}; -e^{2x}\right) + K_1 \right) \quad ; \text{for some constant } K_1 \in \mathbb{R}$$

$$= se^{-\beta(\mu-m_{min})} \left[\frac{e^{2(1-\beta s)x}}{1-\beta s} {}_2F_1\left(1-\beta s, 1; 2-\beta s; -e^{2x}\right) + \frac{1}{\beta s} e^{-2\beta sx} {}_2F_1\left(1, -\beta s; 1-\beta s; -e^{2x}\right) + K_2 \right]$$

; for some constant $K_2 \in \mathbb{R}$

$$= se^{-\beta(m-m_{min})} \left[\frac{e^{\frac{m-\mu}{s}}}{1-\beta s} {}_2F_1\left(1-\beta s, 1; 2-\beta s; -e^{\frac{m-\mu}{s}}\right) + \frac{1}{\beta s} {}_2F_1\left(1, -\beta s; 1-\beta s; -e^{\frac{m-\mu}{s}}\right) \right] + K_3$$

; for some constant $K_3 \in \mathbb{R}$

□

10.

$$\int_a^b e^{-\beta m} \left(1 + \frac{m-\mu}{\sigma}\right)^{-\alpha} dm \tag{B.42}$$

(B.43)

Simplification

By making the substitution $x = 1 + \frac{m-\mu}{\sigma} \Rightarrow m = \mu + \sigma(x-1)$ and $dm = \sigma\,dx$ it is found, after some simplification, that

$$\int_a^b e^{-\beta m} \left(1 + \frac{m-\mu}{\sigma}\right)^{-\alpha} dm$$

$$= \sigma e^{-\beta(\mu-\sigma)} \int_{a_2}^{b_2} e^{-\beta\sigma x} x^{-\alpha}\, dx$$

from where a further substitution can be made $t = \beta\sigma x \Rightarrow x = \frac{1}{\beta\sigma} t$ and $dx = \frac{1}{\beta\sigma}\, dt$. After subsequent simplification, it is found that

$$= \beta^{\alpha-1}\sigma^{\alpha} e^{-\beta(\mu-\sigma)} \int_{a_3}^{b_3} e^{-t} t^{(1-\alpha)-1}\, dt$$

□

Translating through the above substitutions, the new integration bounds are found to be

$$a_3 = \beta\sigma \cdot a_2 = \beta\sigma \left(1 + \frac{a-\mu}{\sigma}\right) = \beta\left(\sigma + a - \mu\right)$$
$$b_3 = \beta\sigma \cdot b_2 = \beta\sigma \left(1 + \frac{b-\mu}{\sigma}\right) = \beta\left(\sigma + b - \mu\right)$$

This forms the basis for the following definition, $r(t) := \beta\left(\sigma + t - \mu\right)$.

Due to the fact that the principle integral originates from the combination of a detection probability and the kernel of a exponential distribution, some parameter restrictions may exist. For this case, where the detection probability is modelled by a Pareto type II distribution function, the parameter restriction, $m \geq \mu \forall\, m$ must be satisfied. From this restriction, the following conclusion can be drawn :

$$m \geq \mu \Rightarrow m - \mu \geq 0$$
$$\text{therefore } r(m) = \beta\left(\sigma + m - \mu\right) \geq 0 \text{ since } \sigma, \beta > 0$$
$$\text{and } a_3,\, b_3 \geq 0$$

Due the positivity of the integration bounds, as well as the functional form, the above stated integral can be written in terms of the incomplete gamma function with respect to a range (B.20).

$$\int_a^b e^{-\beta m} \left(1 + \frac{m-\mu}{\sigma}\right)^{-\alpha} dm = \beta^{\alpha-1}\sigma^{\alpha} e^{-\beta(\mu-\sigma)} \gamma(1-\alpha; r(a), r(b))$$

B.1.4 Distributions used to describe detection probability

1. **Normal Distribution**

   If $X \sim \mathcal{N}(\mu, \sigma^2)$, then the following relations can be stated :

   (a) **Quantile function**

   if $p \in [0, 1]$ represents a probability such that $P[X \leq x] = p$, $x \in \mathbb{R}$, then

   $$x = \mu + \Phi^{-1}(p)\sigma \tag{B.45}$$

   where $\Phi^{-1}(x)$ is the Normal inverse cumulative distribution function.

   (b) **Moments of distribution**

   $$E[X] = \mu \tag{B.46}$$
   $$Var(X) = \sigma^2 \tag{B.47}$$
   $$E\left[X^3\right] = \mu^3 + 3\mu\sigma^2 \tag{B.48}$$

   (c) **Mapping moments to distribution parameters**

   It follows directly from (B.46) and (B.47) that

   $$\mu = E[X] \tag{B.49}$$
   $$\sigma = \sqrt{Var(X)} \tag{B.50}$$

2. **Logistic Distribution**

   If $X \sim \text{Logistic}(\mu, s)$, then the following relations can be stated :

   (a) **Quantile function**

   if $p \in [0, 1]$ represents a probability such that $P[X \leq x] = p$, $x \in \mathbb{R}$, then

   $$x = \mu + s \ln\left(\frac{p}{1-p}\right) \tag{B.51}$$

   (b) **Moments of distribution**

   $$E[X] = \mu \tag{B.52}$$
   $$Var(X) = \frac{1}{3}(s\pi)^2 \tag{B.53}$$

   (c) **Mapping moments to distribution parameters**

   From (B.52) and (B.53) it follows directly that

   $$\mu = E[X] \tag{B.54}$$
   $$s = \frac{1}{\pi}\sqrt{3\,Var(X)} \tag{B.55}$$

3. **Log-Normal Distribution**

If $\ln(X) \sim \mathcal{N}(\mu, \sigma^2)$, then the following relations can be stated :

(a) **Quantile function**

if $p \in [0, 1]$ represents a probability such that $P[X \leq x] = p$, $x \in \mathbb{R}$, then

$$x = \exp\left(\mu + \sigma\Phi^{-1}(p)\right) \tag{B.56}$$

where $\Phi^{-1}(x)$ is the Normal inverse cumulative distribution function.

(b) **Moments of distribution**

$$E\left[X^r\right] = \exp\left(\mu r + \frac{1}{2}r^2\sigma^2\right) \tag{B.57}$$

$$E[X] = \exp\left(\mu + \frac{1}{2}\sigma^2\right) \tag{B.58}$$

$$Var(X) = \exp\left(2\mu + \sigma^2\right)\left(\exp\left(\sigma^2\right) - 1\right) \tag{B.59}$$

$$E\left[X^3\right] = \exp\left(3\mu + \frac{9}{2}\sigma^2\right) \tag{B.60}$$

(c) **Mapping moments to distribution parameters**

By substituting (B.59) into (B.58) and solving for σ the following is obtained

$$\sigma = \sqrt{\ln\left(1 + \frac{Var(X)}{E[X]^2}\right)} \tag{B.61}$$

Thereafter (B.61) can be combined with (B.58) and solved for μ to yield

$$\mu = \ln\frac{E[X]^2}{\sqrt{E[X]^2 + Var(X)}} \tag{B.62}$$

4. **Pareto type II Distribution**

If $X \sim \mathrm{P(II)}(\mu, \sigma, \alpha)$, then the following relations can be stated :

(a) **Quantile function**

if $p \in [0, 1]$ represents a probability such that $P[X \leq x] = p$, $x \in \mathbb{R}$, then

$$x = \mu + \sigma\left((1-p)^{-\frac{1}{\alpha}} - 1\right) \tag{B.63}$$

(b) **Moments of distribution**

$$E[X^r] = \alpha\sigma^\alpha(\mu - \sigma)^{r-\alpha}B\left(1 - \frac{\sigma}{\mu}; \alpha - r; -\alpha\right) \tag{B.64}$$

$$E[X] = \mu + \frac{\sigma}{\alpha - 1} \tag{B.65}$$

$$E[X^2] = \mu^2 + \frac{2\mu\sigma}{\alpha - 1} + \frac{2\sigma^2}{(\alpha - 1)(\alpha - 2)} \tag{B.66}$$

$$E[X^3] = \mu^3 + \frac{3\mu^2\sigma}{\alpha - 1} + \frac{6\mu\sigma^2}{(\alpha - 1)(\alpha - 2)} + \frac{6\sigma^3}{(\alpha - 1)(\alpha - 2)(\alpha - 3)} \tag{B.67}$$

(c) **Mapping moments to distribution parameters**

When solving for μ from (B.65) and (B.66) the following is respectively obtained

$$\mu = E[X] - \frac{\sigma}{\alpha - 1} \tag{B.68}$$

$$\mu = -\frac{\sigma}{\alpha - 1} \pm \sqrt{E[X^2] - \frac{\sigma^2 \alpha}{(\alpha - 1)^2(\alpha - 2)}} \tag{B.69}$$

on combining (B.68) and (B.69) and solving for σ it is found that

$$\sigma = (\alpha - 1)\sqrt{\frac{\alpha - 2}{\alpha} Var(X)} \tag{B.70}$$

which can again be substituted into (B.68) to obtain

$$\mu = E[X] - \sqrt{\frac{\alpha - 2}{\alpha} Var(X)} \tag{B.71}$$

Thereafter, (B.70) and (B.71) can be substituted into (B.67) yielding an equation involving the lower order moments of the distribution and the distributional parameter α. This relation will be used to define the function $f(\alpha)$ as

$$\begin{aligned} f(\alpha) := &-E\left[X^3\right] + (E[X])^3 - 3E[X]Var(X)\frac{\alpha - 2}{\alpha} + 2\left(Var(X)\frac{\alpha - 2}{\alpha}\right)^{\frac{3}{2}} \\ &+ 6\frac{\alpha - 1}{\alpha} Var(X)\left(E[X] - \sqrt{Var(X)\frac{\alpha - 2}{\alpha}}\right) + 6\frac{(\alpha - 1)^2}{(\alpha - 3)(\alpha - 2)}\left(Var(X)\frac{\alpha - 2}{\alpha}\right)^{\frac{3}{2}} \end{aligned} \tag{B.72}$$

where $f(\alpha) = 0$ will be solved numerically for α. Upon obtaining a value for α, values for μ and σ can be found through (B.70) and (B.71).

Appendix C :
Numerical computation of quantities

The computation of values from various quantile functions must be done numerically. Values of m, for which $F_{M_D}(m) = p$ or equivalently $g(m) := p - F_{M_D}(m) = 0$, can be obtained by implementing the Newton-Raphson iterative scheme as follows

$$m_{n+1} = m_n - \frac{g(m_n)}{g'(m_n)} \qquad \text{for } n > 0 \tag{C.73}$$

For some suitable starting values m_0, $m_n \xrightarrow[n]{\infty} m$. In order to briefly examine the adequacy of the starting value regarding convergence, the function $g(\cdot)$ will be plotted for some of the quantile functions. In this illustrative section the detection probability is modelled by the cumulative distribution function of a Normal random variable.

C.1.1 Soft detection threshold

Quantiles from the distribution $F_{M_D}(m) = p$, as seen in (8.22) (p. 28), can only be solved using numerical methods. Therefore it is defined that

$$\begin{aligned} g(m) &:= \frac{p}{C_{Norm}} - F_{M_G}(m) \cdot \Phi(z_m) + c_1 \left[\Phi(z_m) - c_2 \cdot \Phi(z_m^*)\right] - c_3 \\ \Rightarrow g'(m) &= -f_{M_G}(m) \cdot \Phi(z_m) - F_{M_G}(m) \cdot \frac{1}{\sigma}\phi(z_m) + \frac{c_1}{\sigma}[\phi(z_m) - c_2 \cdot \phi(z_m^*)] \end{aligned} \tag{C.74}$$

and in summary, $\Phi(t)$ represents the cumulative distribution function of a standard Normal random variable. Additionally, $z_m = \frac{m-\mu}{\sigma}$ and $z_m^* = \frac{m-(\mu-\sigma^2\beta)}{\sigma}$.

$$c_1 = \left[1 - \exp\left(-\beta(m_{max} - m_{min})\right)\right]^{-1} ; \qquad c_2 = \exp\left(-\beta(\mu - m_{min} - \frac{1}{2}\sigma^2\beta)\right)$$

$$c_3 = c_1\left[\Phi(z_{m_{min}}) - c_2\Phi(z^*_{m_{min}})\right]$$

Summary of function $g(\cdot)$

To investigate sensitivities of the parameters, the surface generated by (C.74) can be plotted. Without loss of generality all subsequent plots are made for $p = 0.5$.

As seen by examining Figures C.1 to C.3, the function $g(m)$ has a general shape for the parameter space when considered over the magnitude range. For various $p \in [0, 1]$ it appears that the function is, generally, well-behaved in the neighbourhood of the root at $g(m) = 0$ and it is seen that a starting value of $m_0 = m_{min}$ is a suitable choice in a number of circumstances. However this starting value

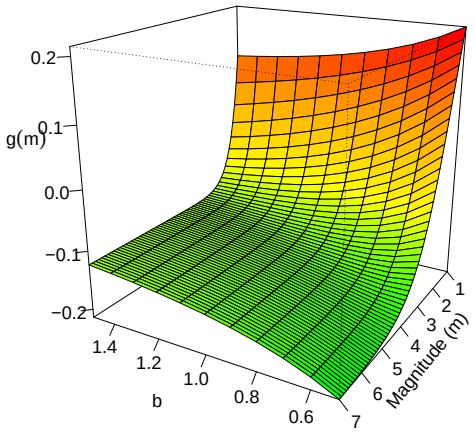
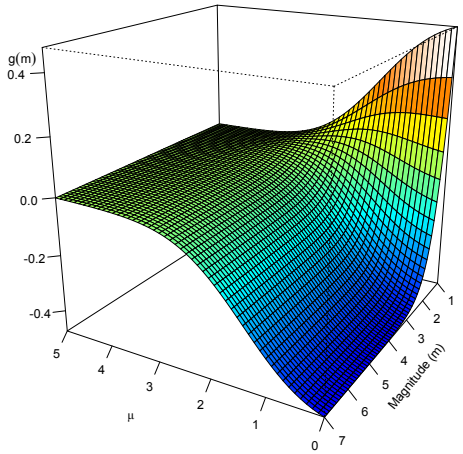
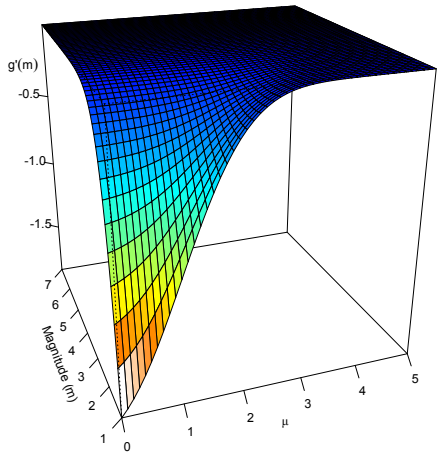


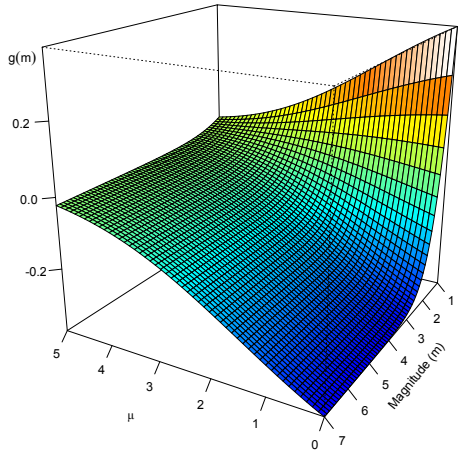
Figure C.1: *Plot of function $g(m)$ (from (C.74)) with illustrative values of $p = 0.5$; $m_{min} = 1$; $m_{max} = 7$; $\mu = 2$ and $\sigma = 0.8$.*



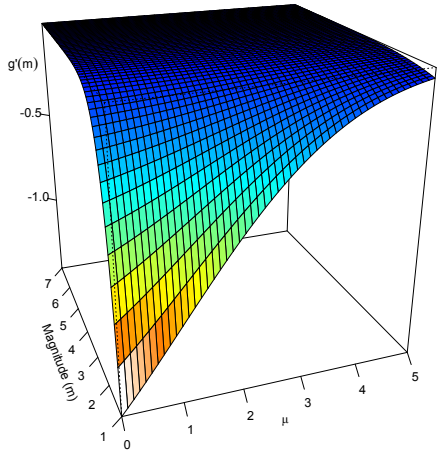
(a) *Function of $g(M)$ for varying μ and $\sigma = 0.8$*



(b) *Function of $\frac{dg}{dm}$ for varying μ and $\sigma = 0.8$*

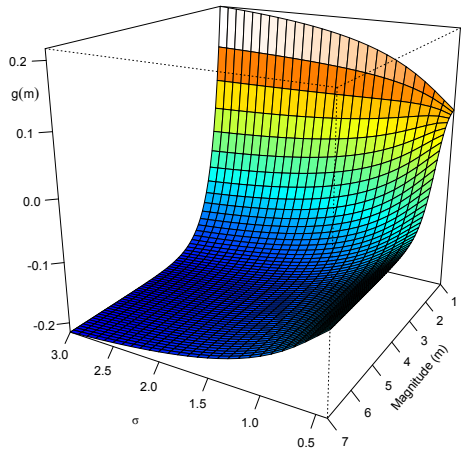


(c) *Function of $g(M)$ for varying μ and $\sigma = 2$*

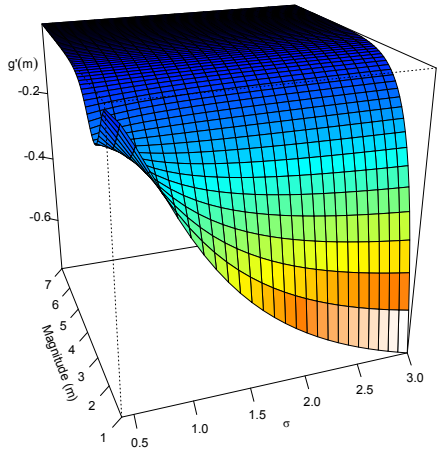


(d) *Function of $\frac{dg}{dm}$ for varying μ and $\sigma = 2$*

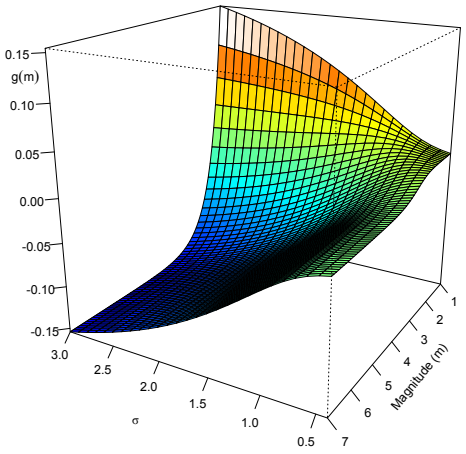
Figure C.2: *Illustration of sensitivities of functions $g(m)$ and first derivative $g'(m)$ (from (C.74)) to the parameter μ. Illustrative values for parameters : $p = 0.5$; $m_{min} = 1$; $m_{max} = 7$ and $b = 0.9$.*



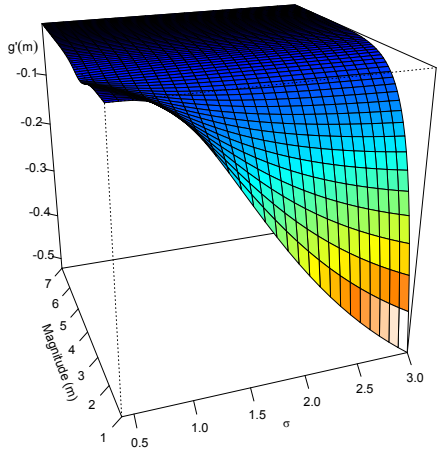
(a) *Function of* $g(M)$ *for varying* m_{min} *and* $\mu = 2$



(b) *Function of* $\frac{dg}{dm}$ *for varying* m_{min} *and* $\mu = 2$



(c) *Function of* $g(M)$ *for varying* m_{min} *and* $\mu = 3$



(d) *Function of* $\frac{dg}{dm}$ *for varying* m_{min} *and* $\mu = 3$

Figure C.3: *Illustration of sensitivities of functions* $g(m)$ *and first derivative* $g'(m)$ *(from (C.74)) to the parameter* σ*. Illustrative values for parameters :* $p = 0.5$*;* $m_{min} = 1$*;* $m_{max} = 7$ *and* $b = 0.9$*.*

does not guarantee convergence under all circumstances. One such example is the detected magnitude distribution with the following parameter values :

Parameter	Value
m_{min}	1
m_{max}	7
b-Value	0.9 ($\beta \approx 2.072$)
μ	3
σ	0.8

Table C.1: *Parameters of detected magnitude distribution under which numerical computation of the quantile function, implementing the Newton-Raphson method, does not converge for all values when the starting value is m_{min}.*

When evaluating the quantile function for probabilities in the approximate range 0.597 to 0.974 (resultant magnitudes of approximately 2.44 to 4.14), the Newton-Raphson method does not yield converging results for the particular starting value of m_0. In this case the starting value $m_0 = \mu$ is chosen and indeed convergence of the quantiles is obtained.

Hence, the following has been empirically added to the current application of the Newton-Raphson algorithm:

If it is established that the Newton-Raphson method does not converge for a specific starting value, the algorithm is re-implemented with a subsequent starting value as found in Table C.2.

Implementation of Newton-Raphson method after number of previous failed attempts	Starting Value
1	m_{min}
2	μ
3	$\frac{1}{3}(m_{max} + m_{min})$
4	$\frac{1}{2}(m_{max} + m_{min})$
5	m_{max}

Table C.2: *List of starting values to be used in the Newton-Raphson method when evaluating the quantile function $Q_{M_D}(p)$.*

C.1.2 Sharp detection threshold

In order to obtain values of the quantile function $Q_{M_D}(p) = p$ where $p < F_{M_D}(m_c)$ the following can be defined

$$g(m) := p \cdot \frac{\Phi(z_{m_c})}{C_{Norm}} - F_{M_G}(m) \cdot \Phi(z_m) + c_1 \left[\Phi(z_m) - c_2 \cdot \Phi(z_m^*)\right] - c_3 \tag{C.75}$$

Where upon the iterative scheme presented in (8.24) can be used to solve for specific values of m.

It can be seen that (8.29) is simply a shift of (C.74) by an amount of $\left(1 - \frac{p}{C_{Norm}}\right)\Phi(z_{m_c})$. Because of this, the previous Newton-Raphson analysis of (C.74) is valid in this case as well.

A critical difference though, is that the domain of the function g, and subsequently its first derivative, is restricted to $[m_{min}, m_c)$, where the previous function had domain $[m_{min}, m_{max}]$. The use of Table C.2 as reference lookup for starting values is also applied.

Appendix D : b-Value estimation

D.1.1 Page relation

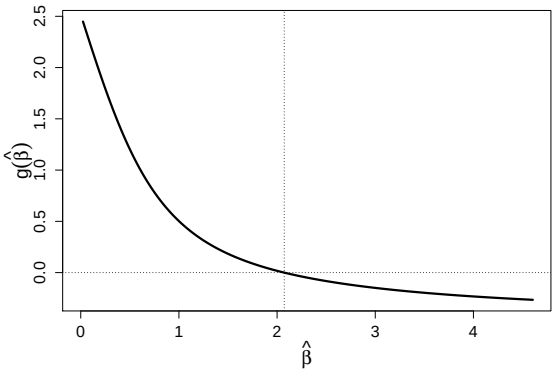
β values (as well as subsequent b-values) can only be extracted from eqrefeq:Page, on page 69, by utilizing numerical methods. Therefore, let

$$g(\hat{\beta}) := \frac{1}{\hat{\beta}} - (\bar{m} - m_{min}) - \frac{m_{max} - m_{min}}{e^{\hat{\beta}(m_{max} - m_{min})} - 1} \tag{D.76}$$

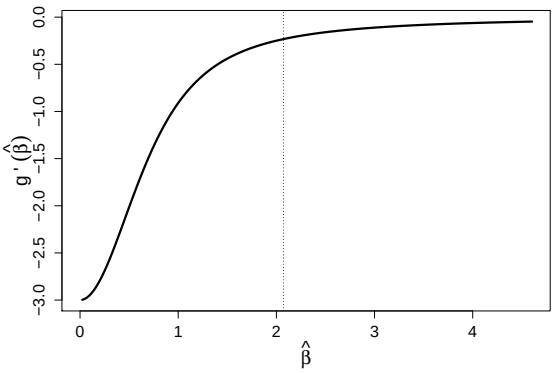
$$\Rightarrow g'(\hat{\beta}) = \frac{-1}{\hat{\beta}^2} + \left(\frac{m_{max} - m_{min}}{e^{\hat{\beta}(m_{max} - m_{min})} - 1} \right)^2 e^{\hat{\beta}(m_{max} - m_{min})} \tag{D.77}$$

Summary of function $g(\cdot)$

Without loss of generality and to further the current investigation, $g(\cdot)$ is plotted for illustrative values of the parameters.



(a) *Function of* $g(\hat{\beta})$*. Reference lines added at* $g(\hat{\beta}) = 0$ *as well as the estimated location of the root.*



(b) *First derivative of function of* $g(\hat{\beta})$*. Reference line added at estimated location of the root.*

Figure D.4: *Plot of functions* $g(\hat{\beta})$ *and* $g'(\hat{\beta})$ *((D.76) and (D.77)) with illustrative values of* $m_{min} = 1$*;* $m_{max} = 7$ *and* $\bar{m} = E[M]$ *(calculated from (8.3), p. 23, with consistent values for parameters. Additionally* $\beta = b \ln 10 = 0.9 \ln 10 \approx 2.072$*).*

As seen in Figure D.4 above, function $g(\cdot)$ as well as it's derivative, is well-behaved in the neighbourhood of the root. Therefore, added to the fact that $g(\cdot)$ has a simple mathematical form, the application of the Newton-Raphson algorithm seems to be suited.

D.1.1.1 Numerical computation of estimates

The following scheme can subsequently be used to solve for $g(\hat{\beta}) = 0$:

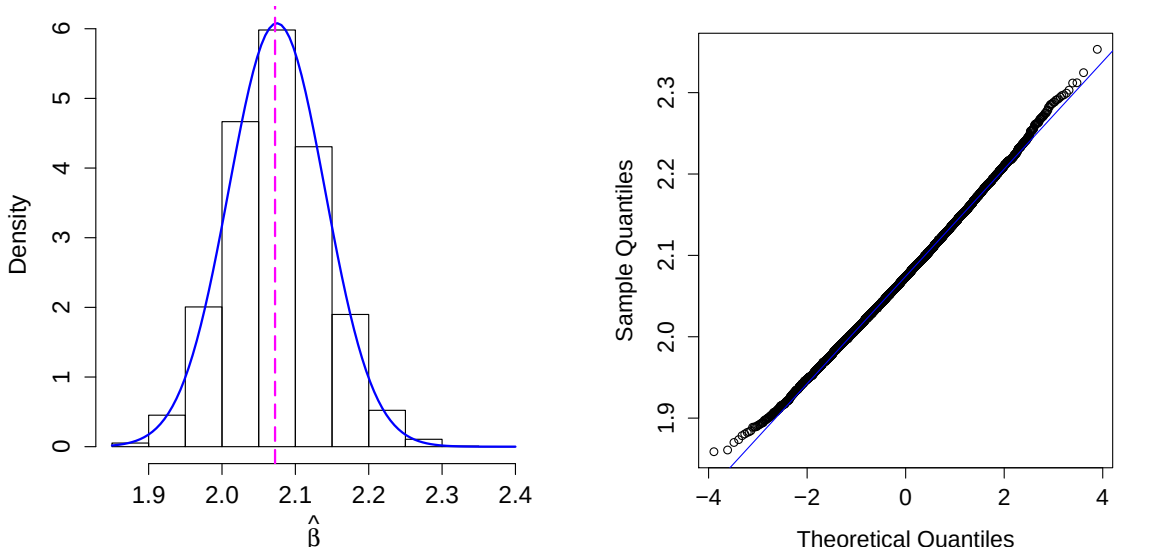
$$\hat{\beta}_{n+1} = \hat{\beta}_n - \frac{g(\hat{\beta}_n)}{g'(\hat{\beta}_n)} \tag{D.78}$$

$$\hat{\beta}_0 = \frac{1}{\bar{m} - m_{min}}$$

Because of it's resemblance to the Page relation, it should be clear why the Aki-Utsu [1, 56] maximum-likelihood estimator is an apt choice for the starting value $\hat{\beta}_0$. Utilizing this algorithm, the quadratic convergence of the Newton-Raphson method, together with the well behaved nature of $g(\hat{\beta})$ and suitable starting value $\hat{\beta}_0$, yields stable values of $\hat{\beta}$, usually, in as little as 2 to 3 iterations.

D.1.1.2 Sampling distribution

Asymptotically the method of maximum likelihood specifies that the estimator will be unbiased and Normally distribution, however due to the size of the simulated dataset, further investigation is warranted. Since there is no closed form solution for $\hat{\beta}$ stochastic simulations will be resorted to in order to gauge the performance of the estimator. A sampling distribution is estimated based on the following parameters : b-Value $= 0.9$ ($\beta \approx 2.072$); $m_{min} = 1$ and $m_{max} = 7$. The number of simulations have been specified as 10,000 with 1,000 events per simulation. The resulting sampling distribution can be seen in Figure D.5.



(a) *Histogram of estimated sampling distribution along with N(2.07, 0.066) overlay.*

(b) *Quantile-Quantile (QQ) plot of estimated sampling distribution.*

Figure D.5: *Estimated sampling distribution of $\hat{\beta}$.*

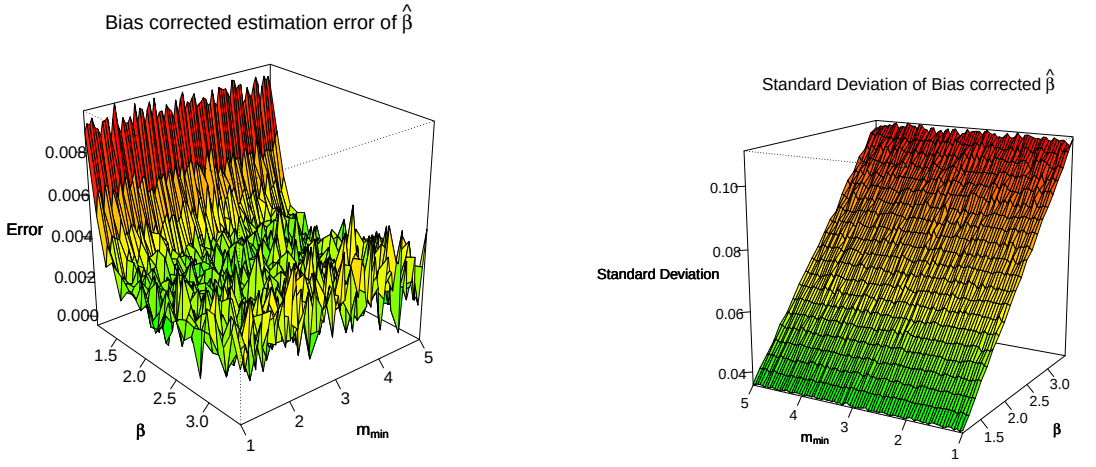
Normality Test	p-Value
D'Agostino test for skewness in Normally distributed data	0.0005445
t-Test for location	0.0001400
Anderson-Darling test for normality	0.0001679

Table D.3: *Tests performed on estimated sampling distribution*

According to the QQ-plot in Figure D.5, although most of the quantiles coincide with those of a theoretical Normal distribution with same mean and standard deviation, the quantiles relating to the tails of the distribution seem to deviate from that expected under normality. This is further supported by the statis-

tically significant p-values obtained from the D'Agostino skewness test, as well as the Anderson-Darling test for normality. The distribution of sample errors further demonstrate that there seems to exist a slight bias in the estimator. This is to some extent expected, since Zhang and Song [62] concluded, from the exact distribution of the Aki-Utsu estimator $\hat{\beta}$, that $E[\hat{\beta}] = \beta\frac{n}{n-1}$. Using this result, the currently examined estimator can also be corrected for bias.

In order to measure the effectiveness of the estimator over the entire range of possibilities a detailed simulation of the error distribution is undertaken according to the framework as set out in in Table **??**. For this exercise, only the mean and standard deviation of the estimation error from each individual simulation will be entered into the comparison. During this simulation the range over which events are generated and the β value estimated has constantly been kept at a magnitude interval of length 6. Therefore the full specification of parameters are as follows : b-Value $\in [0.5; 1.5]$,i.e. $\beta \in [1.151; 3.453]$; $m_{min} \in [1; 5]$; $m_{max} = m_{min} + 6$. Furthermore, the number of simulations have been fixed at 10,000 with 1,000 events per simulation. The resultant sampling distribution can be seen in Figure D.6.



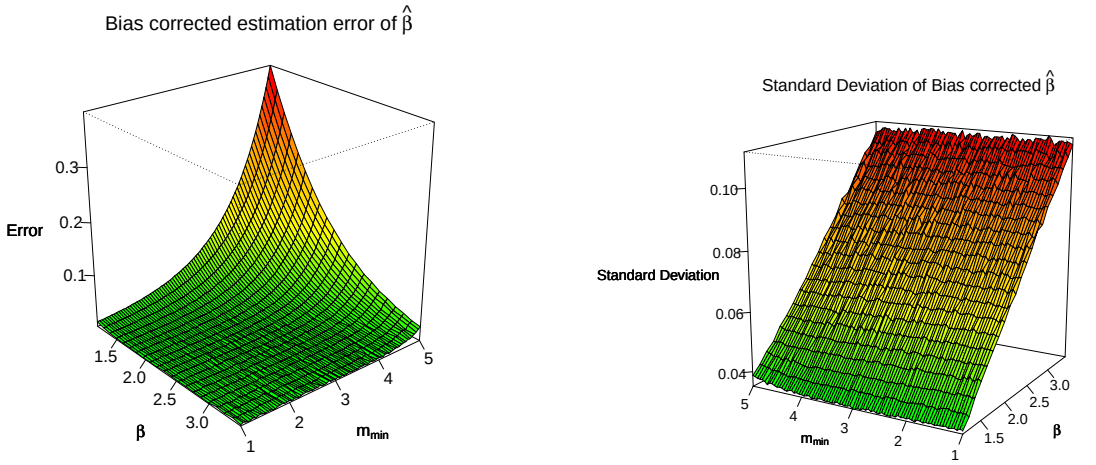
(a) *Mean estimation error* **(b)** *Standard deviation of estimation error*

Figure D.6: *Statistics for sampling distribution of* $\hat{\beta}$.

From this simulation it appears that the bias-correcting factor only makes a slight improvement, but does not correct for the full bias of the estimator. Further investigation of this must be undertaken, but currently does not fall in the scope of this investigation.

Alternatively, when the maximum point of the simulated Gutenberg-Richter distribution remains constant (i.e. the range over which seismic events are generated and the β value estimated varies), a slightly different scenario presents itself.
The following parameters have been specified in order to estimate the sampling distribution of the β estimator as seen in Figure D.7 : b-Value $\in [0.5; 1.5]$, i.e. $\beta \in [1.151;\ 3.453]$; $m_{min} \in [1; 5]$; $m_{max} = 7$.

The number of simulations have been specified as 10,000 with 1,000 events per simulation.



(a) *Mean estimation error* **(b)** *Standard deviation of estimation error*

Figure D.7: *Statistics for sampling distribution of* $\hat{\beta}$.

This evident systematic bias can only be the result of narrowing the magnitude range of the investigation for increasing m_{min}. For this reason m_{min} is assumed fixed throughout the investigation. Furthermore, this indicates the need to consider the length of permissible magnitudes, i.e. $m_{max} - m_{min}$. Therefore, throughout the investigation this range is assumed fixed at a length of 6 ($m_{min} = 1$ and $m_{max} = 7$).

Appendix E : Statistical Results

E.1.1 Explicit modelling of detection threshold

E.1.1.1 Aggregated results of threshold estimation methods

Goodness of fit estimation method

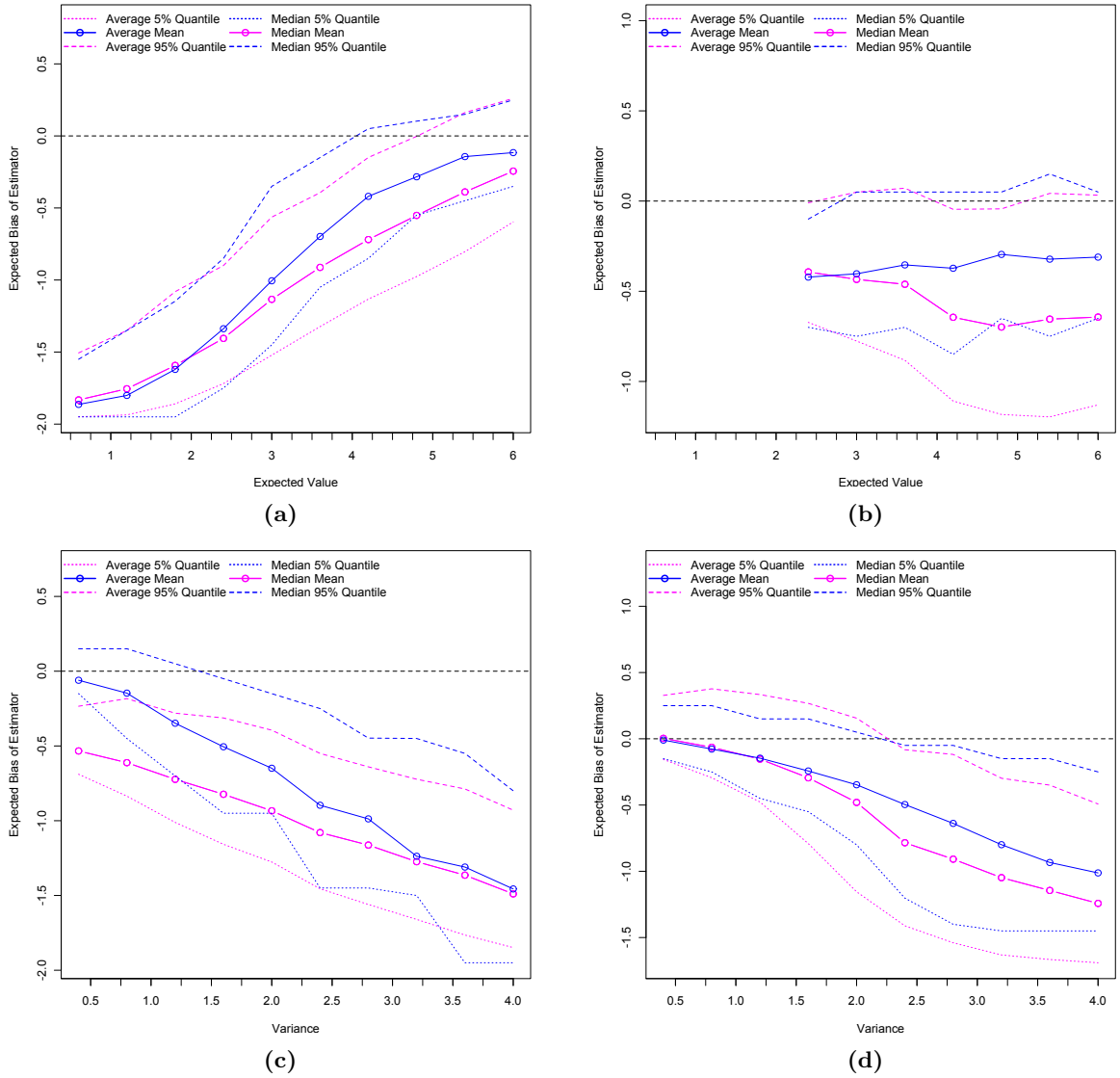


Figure E.8: *Expected bias of sampling distribution together with error bounds for GOF estimation method illustrated as a function of expected value (figs. a and b) and variance (figs. c and d) of distribution used to model detection probability. Detection probability in Figures a and c modelled by a cumulative Normal distribution, while modelled by a cumulative logistic distribution in Figures b and d.*

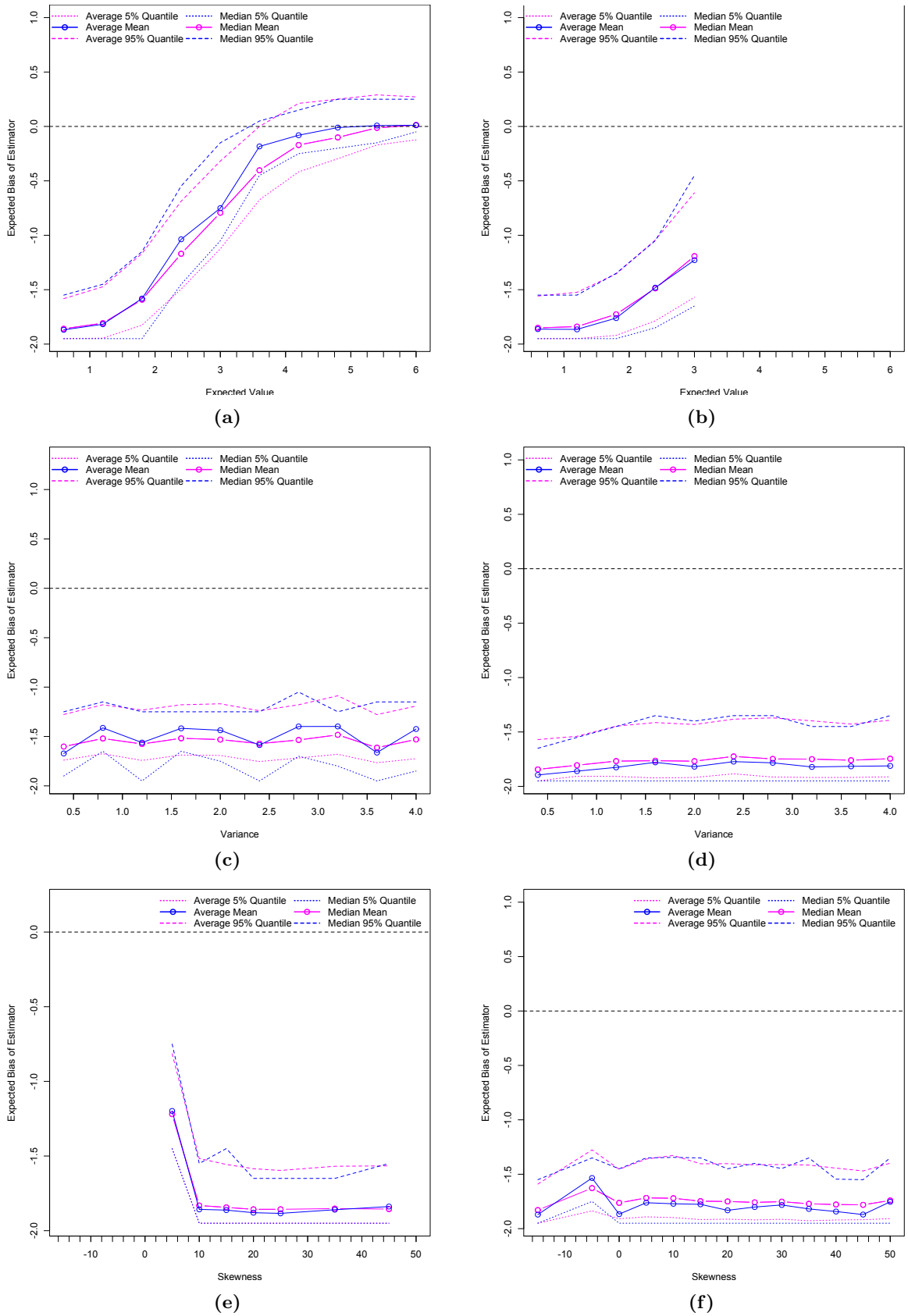


Figure E.9: *Expected bias of sampling distribution and error bounds for GOF estimation method illustrated as a function of expected value (figs. a and b), variance (figs. c and d) and skewness (figs. e and f) of distribution used to model detection probability. Detection probability in Figures a, c and e modelled by a cumulative Log-Normal distribution, while modelled by a cumulative Pareto type II distribution in Figures b, d and f.*

Maximum Curvature estimation method

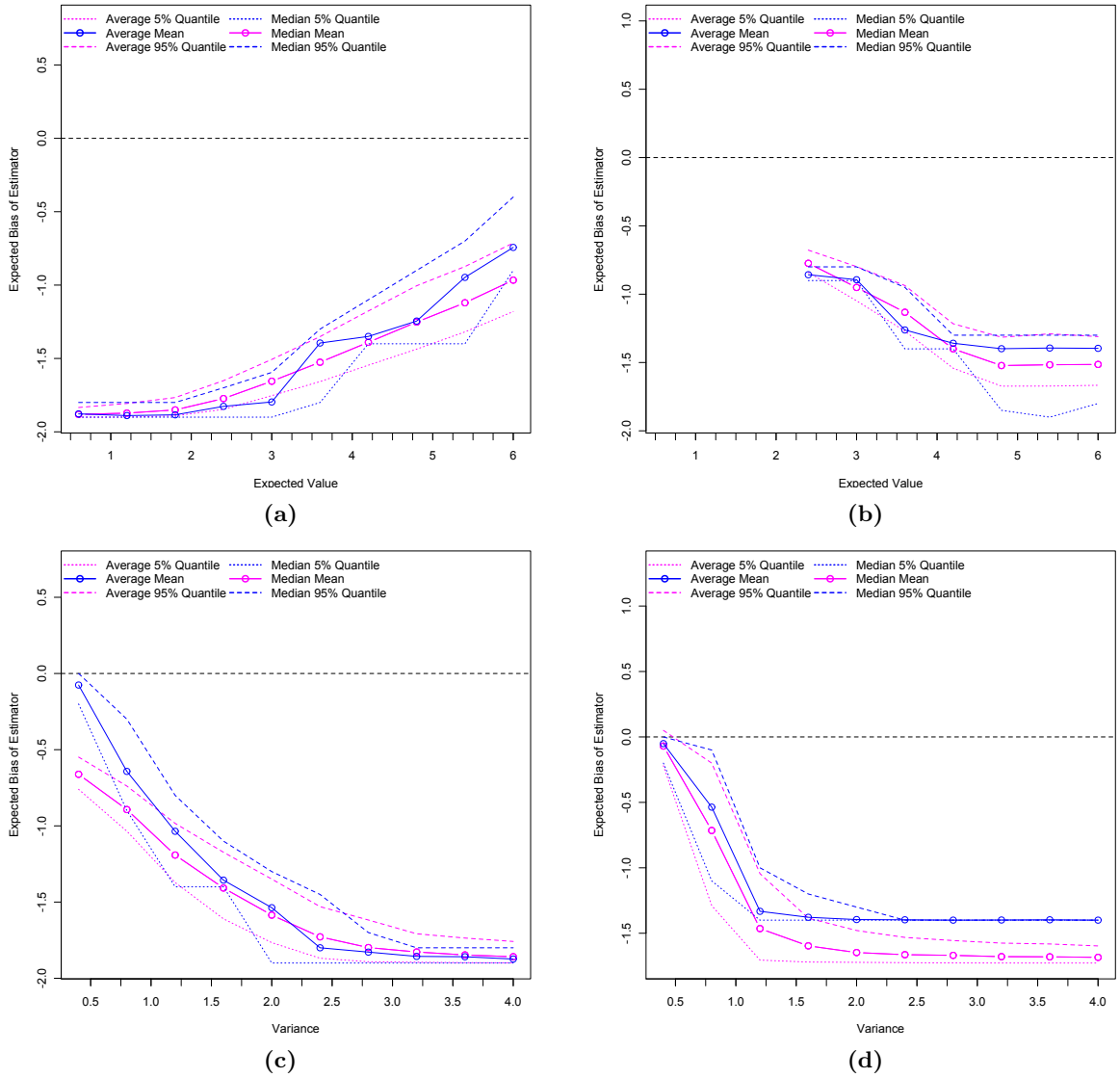


Figure E.10: *Expected bias of sampling distribution together with error bounds for GOF estimation method illustrated as a function of expected value (figs. a and b) and variance (figs. c and d) of distribution used to model detection probability. Detection probability in Figures a and c modelled by a cumulative Normal distribution, while modelled by a cumulative logistic distribution in Figures b and d.*

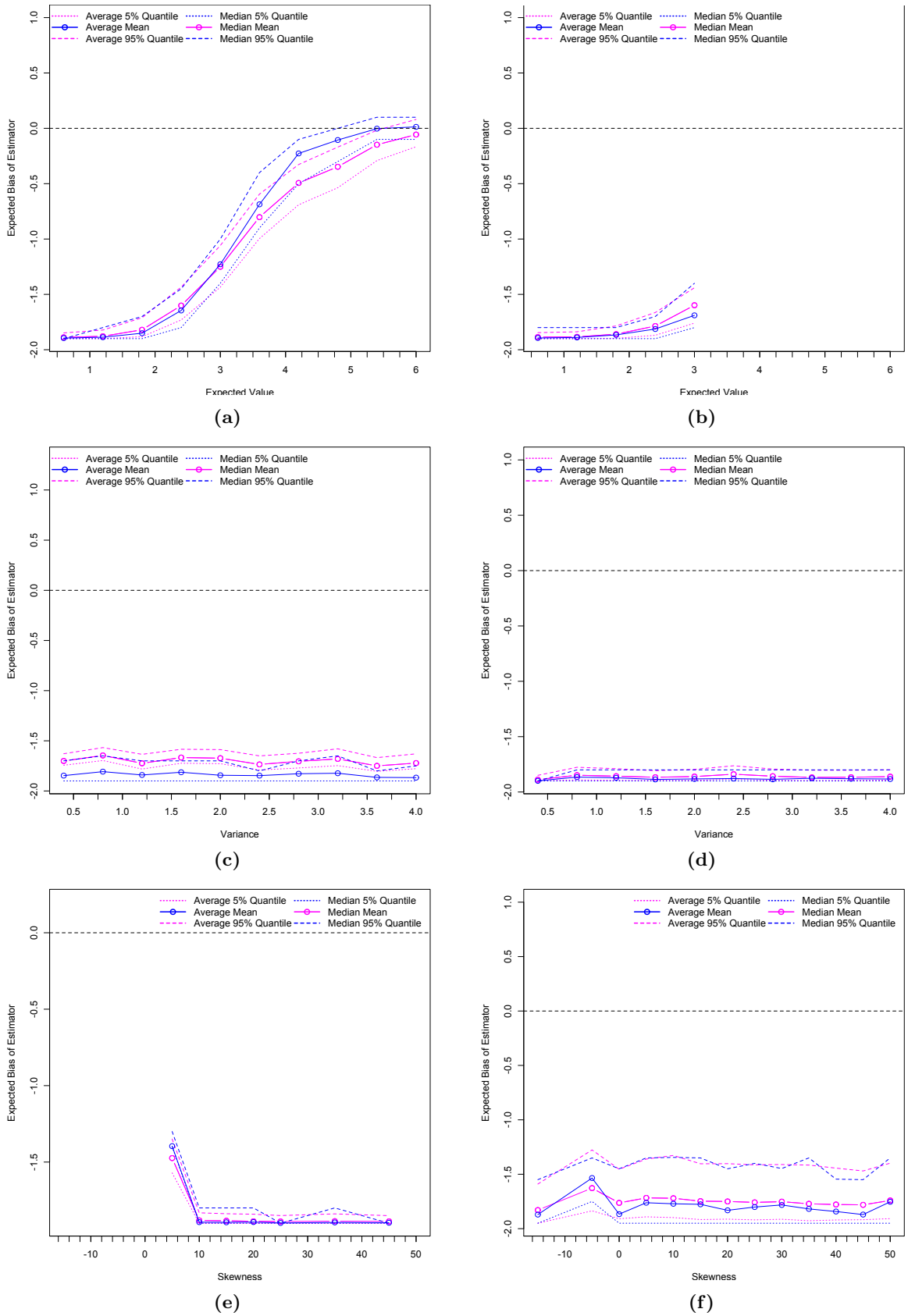


Figure E.11: *Expected bias of sampling distribution and error bounds for GOF estimation method illustrated as a function of expected value (figs. a and b), variance (figs. c and d) and skewness (figs. e and f) of distribution used to model detection probability. Detection probability in Figures a, c and e modelled by a cumulative Log-Normal distribution, while modelled by a cumulative Pareto type II distribution in Figures b, d and f.*

m_c by b-Value stability

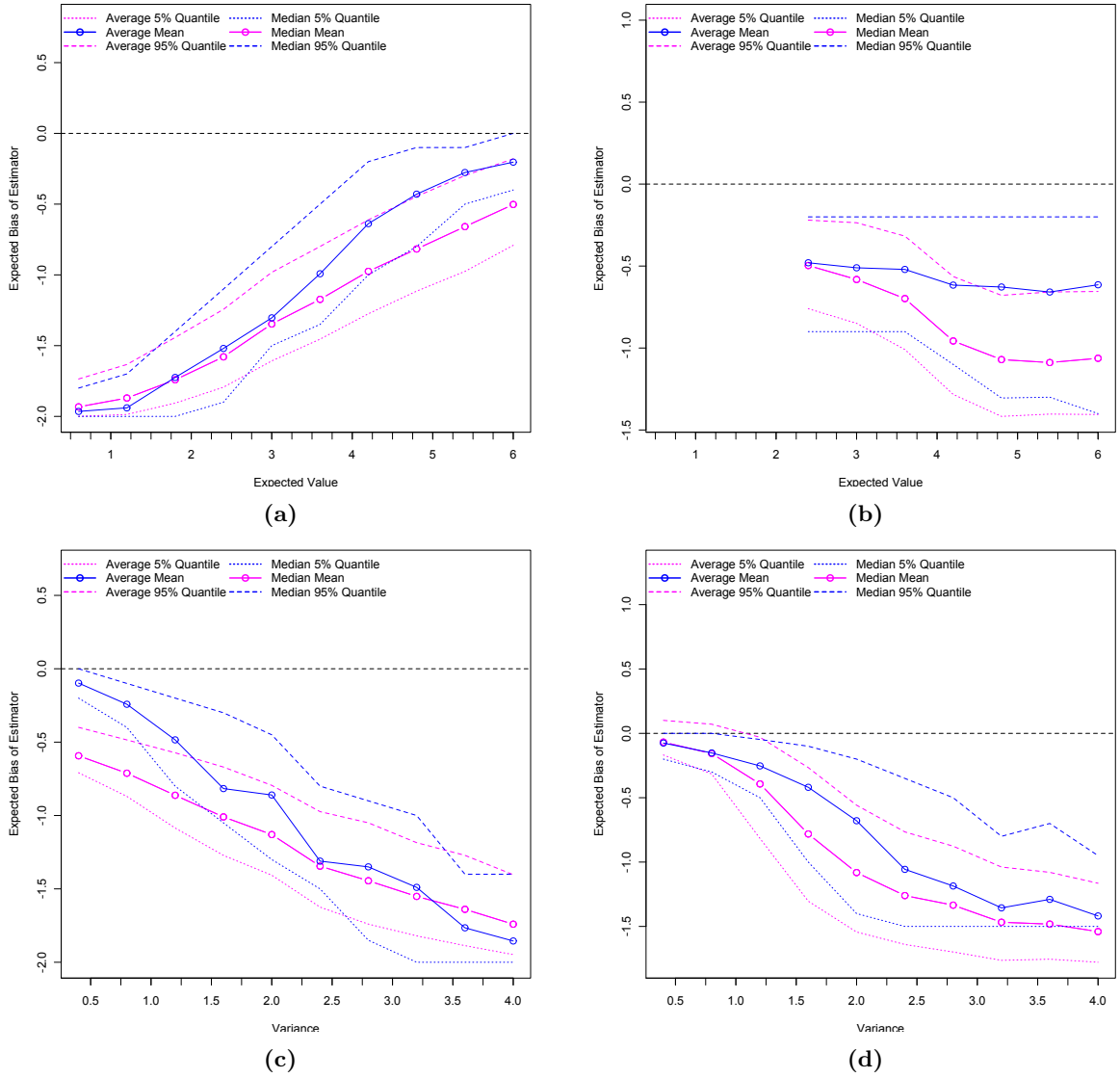


Figure E.12: *Expected bias of sampling distribution together with error bounds for bVS estimation method illustrated as a function of expected value (figs. a and b) and variance (figs. c and d) of distribution used to model detection probability. Detection probability in Figures a and c modelled by a cumulative Normal distribution, while modelled by a cumulative logistic distribution in Figures b and d.*

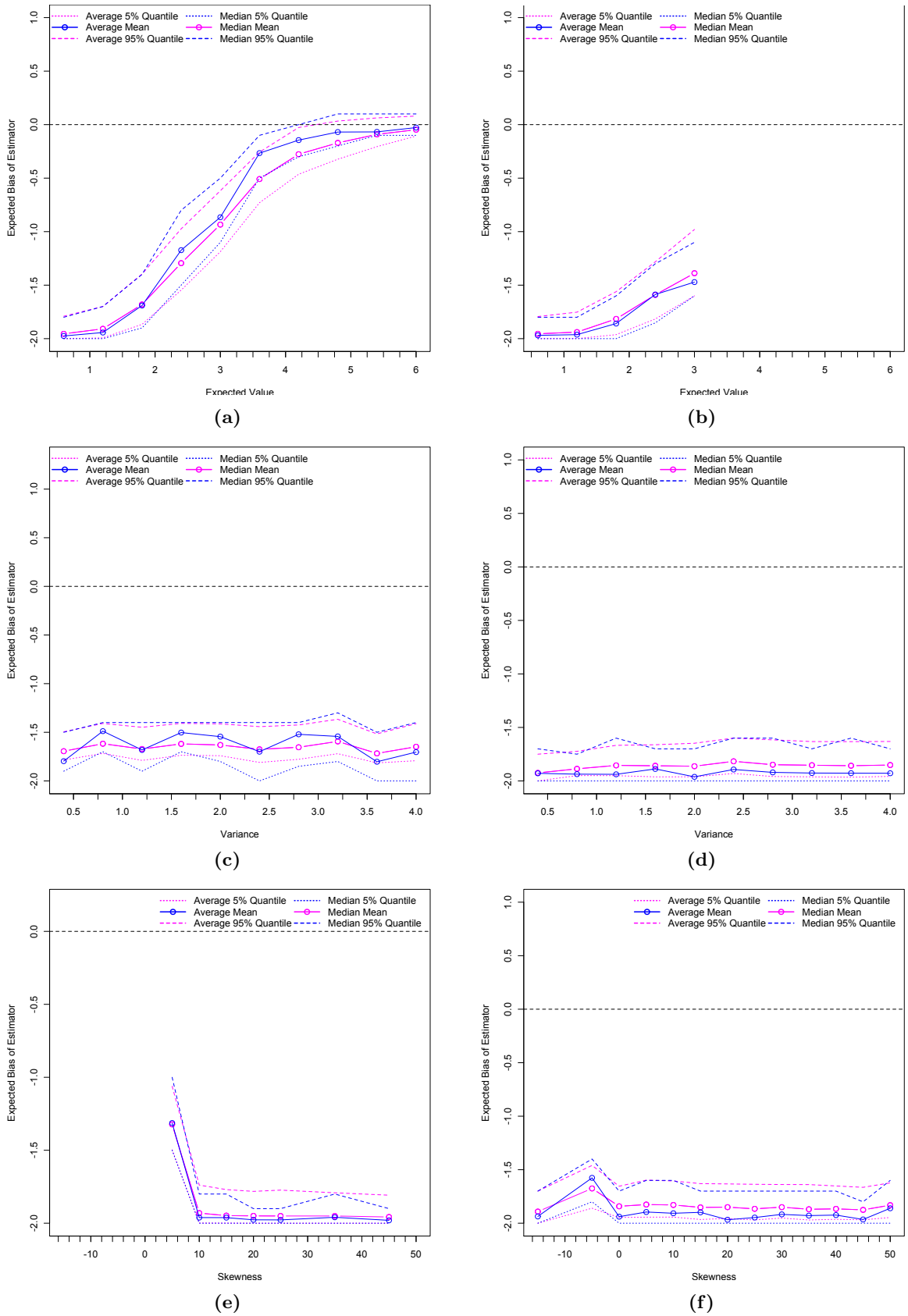


Figure E.13: *Expected bias of sampling distribution and error bounds for bVS estimation method illustrated as a function of expected value (figs. a and b), variance (figs. c and d) and skewness (figs. e and f) of distribution used to model detection probability. Detection probability in Figures a, c and e modelled by a cumulative Log-Normal distribution, while modelled by a cumulative Pareto type II distribution in Figures b, d and f.*

Entire Magnitude Range estimation method

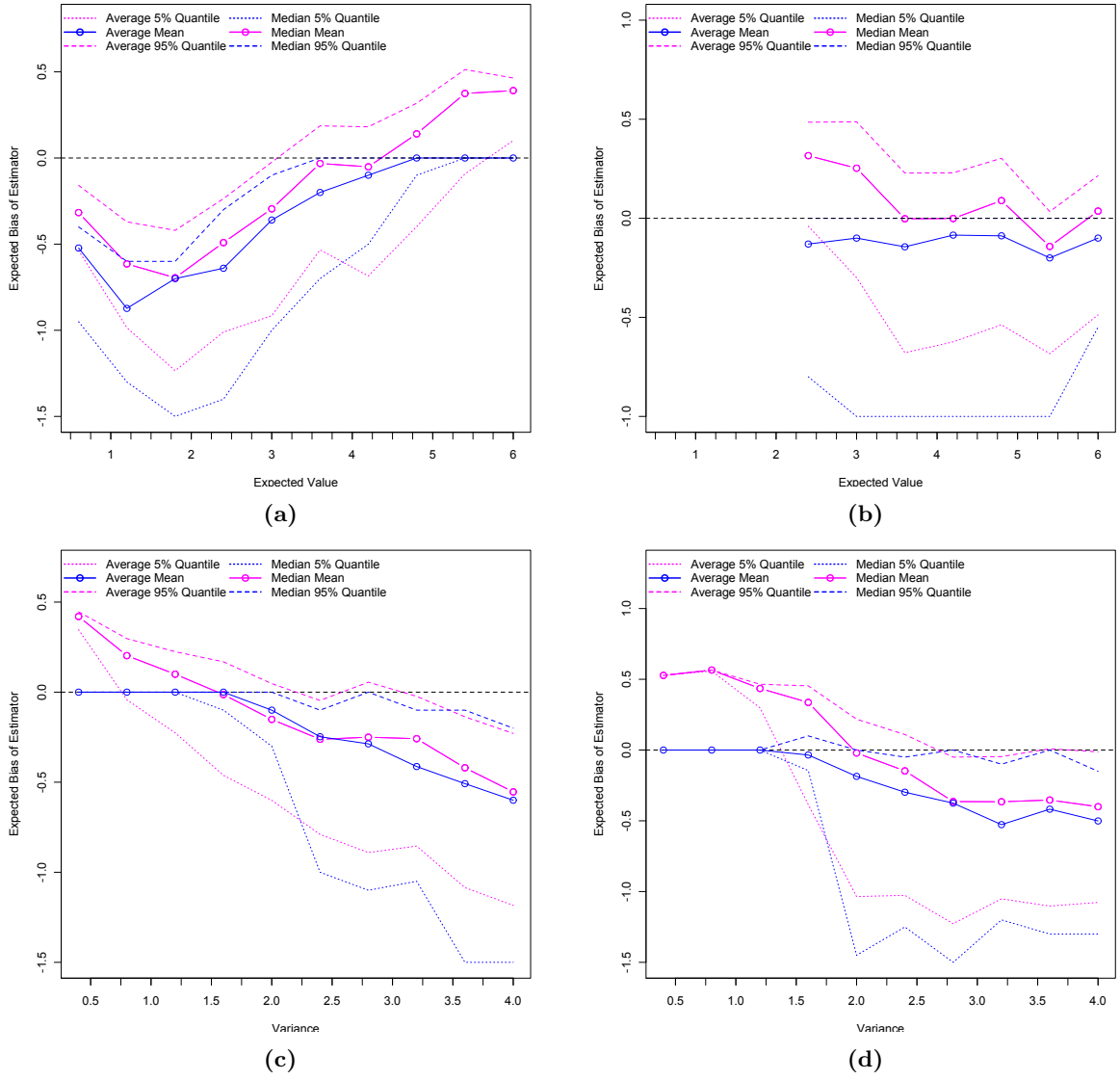


Figure E.14: *Expected bias of sampling distribution together with error bounds for EMR estimation method illustrated as a function of expected value (figs. a and b) and variance (figs. c and d) of distribution used to model detection probability. Detection probability in Figures a and c modelled by a cumulative Normal distribution, while modelled by a cumulative logistic distribution in Figures b and d.*

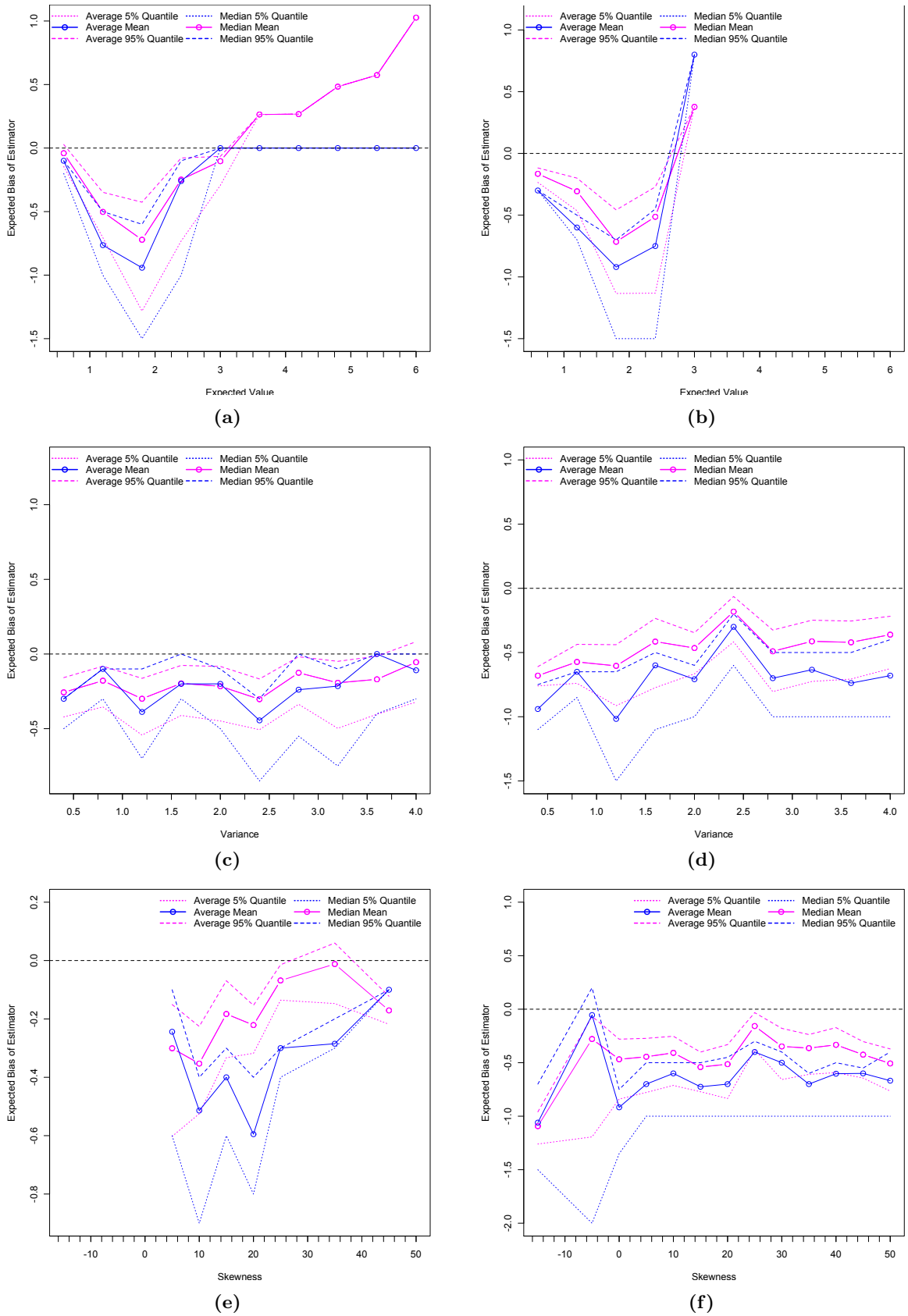


Figure E.15: *Expected bias of sampling distribution and error bounds for EMR estimation method illustrated as a function of expected value (figs. a and b), variance (figs. c and d) and skewness (figs. e and f) of distribution used to model detection probability. Detection probability in Figures a, c and e modelled by a cumulative Log-Normal distribution, while modelled by a cumulative Pareto type II distribution in Figures b, d and f.*

Median based assessment of the segment-slope

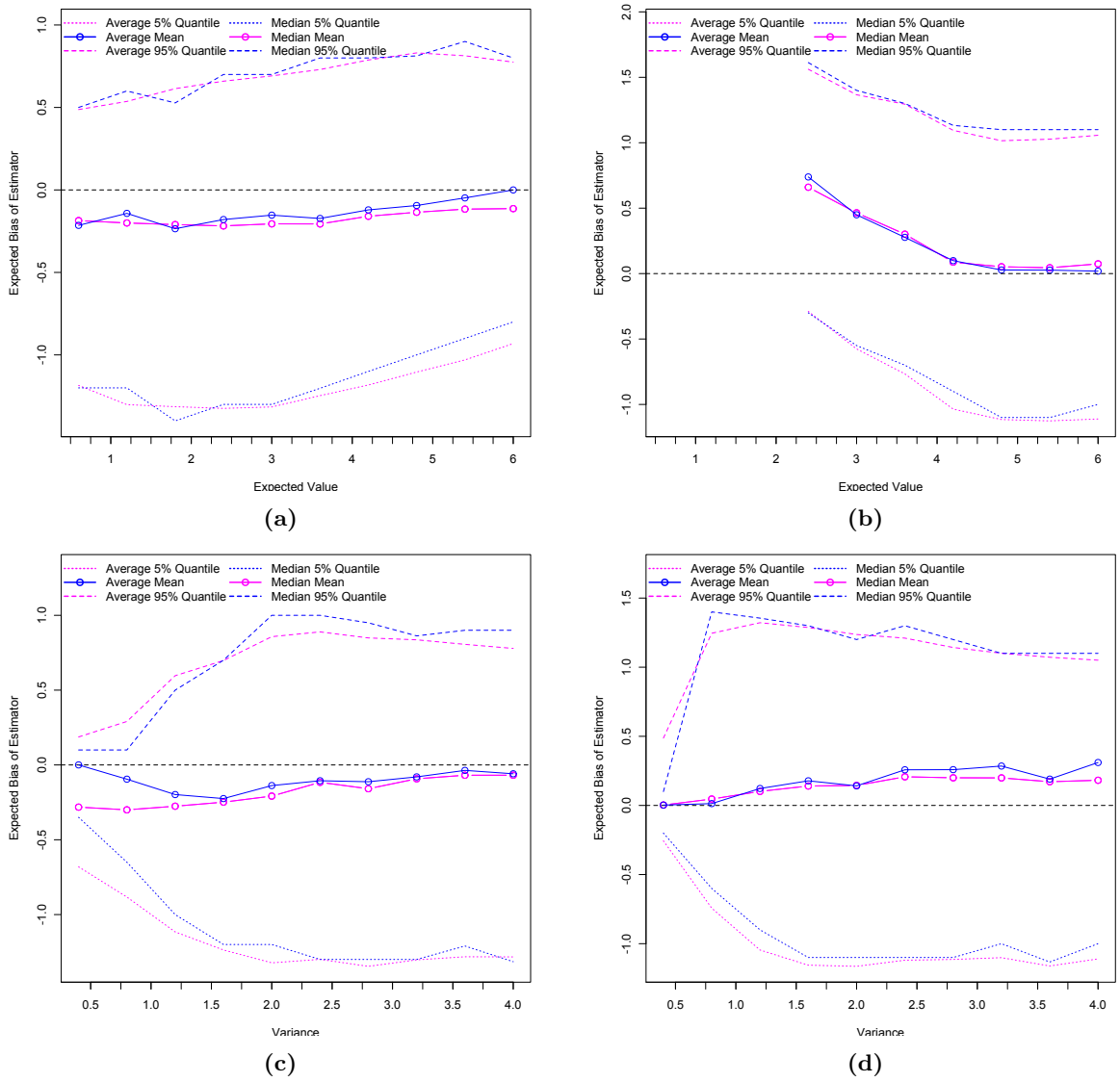


Figure E.16: *Expected bias of sampling distribution together with error bounds for MBASS estimation method illustrated as a function of expected value (figs. a and b) and variance (figs. c and d) of distribution used to model detection probability. Detection probability in Figures a and c modelled by a cumulative Normal distribution, while modelled by a cumulative logistic distribution in Figures b and d.*

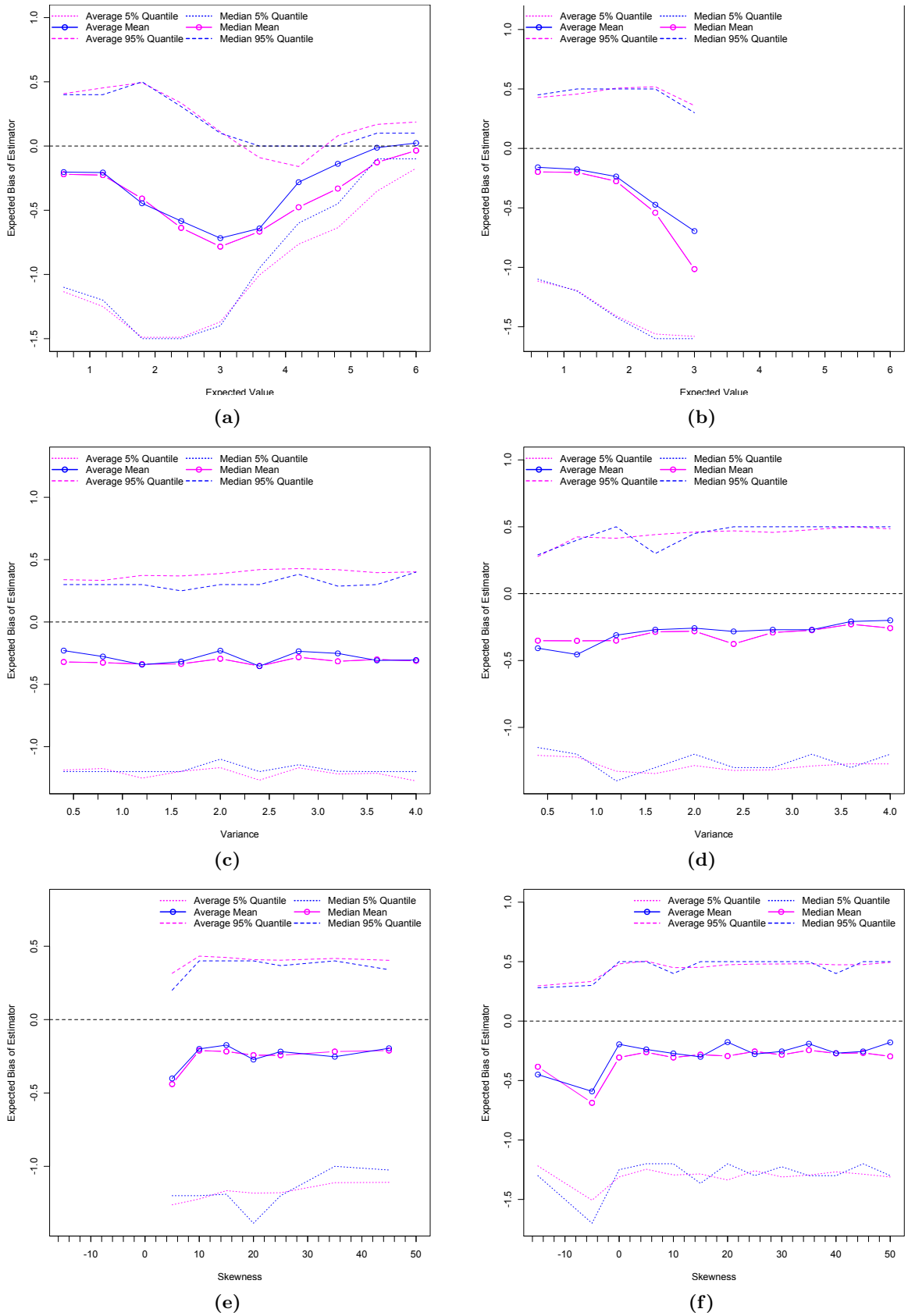


Figure E.17: *Expected bias of sampling distribution and error bounds for GOF estimation method illustrated as a function of expected value (figs. a and b), variance (figs. c and d) and skewness (figs. e and f) of distribution used to model detection probability. Detection probability in Figures a, c and e modelled by a cumulative Log-Normal distribution, while modelled by a cumulative Pareto type II distribution in Figures b, d and f.*

Moment incorporating threshold calculation

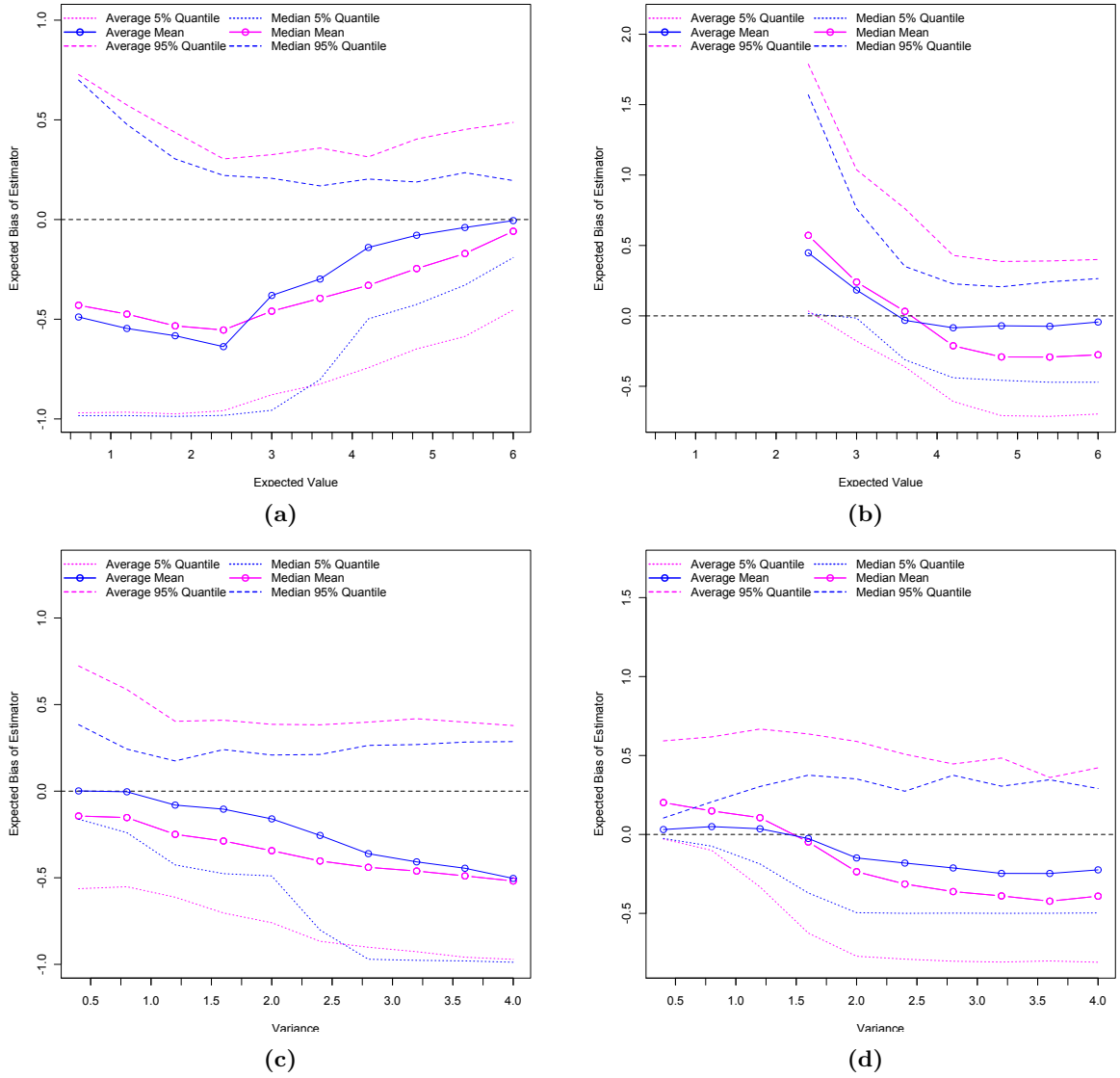


Figure E.18: *Expected bias of sampling distribution together with error bounds for MITC estimation method illustrated as a function of expected value (figs. a and b) and variance (figs. c and d) of distribution used to model detection probability. Detection probability in Figures a and c modelled by a cumulative Normal distribution, while modelled by a cumulative logistic distribution in Figures b and d.*

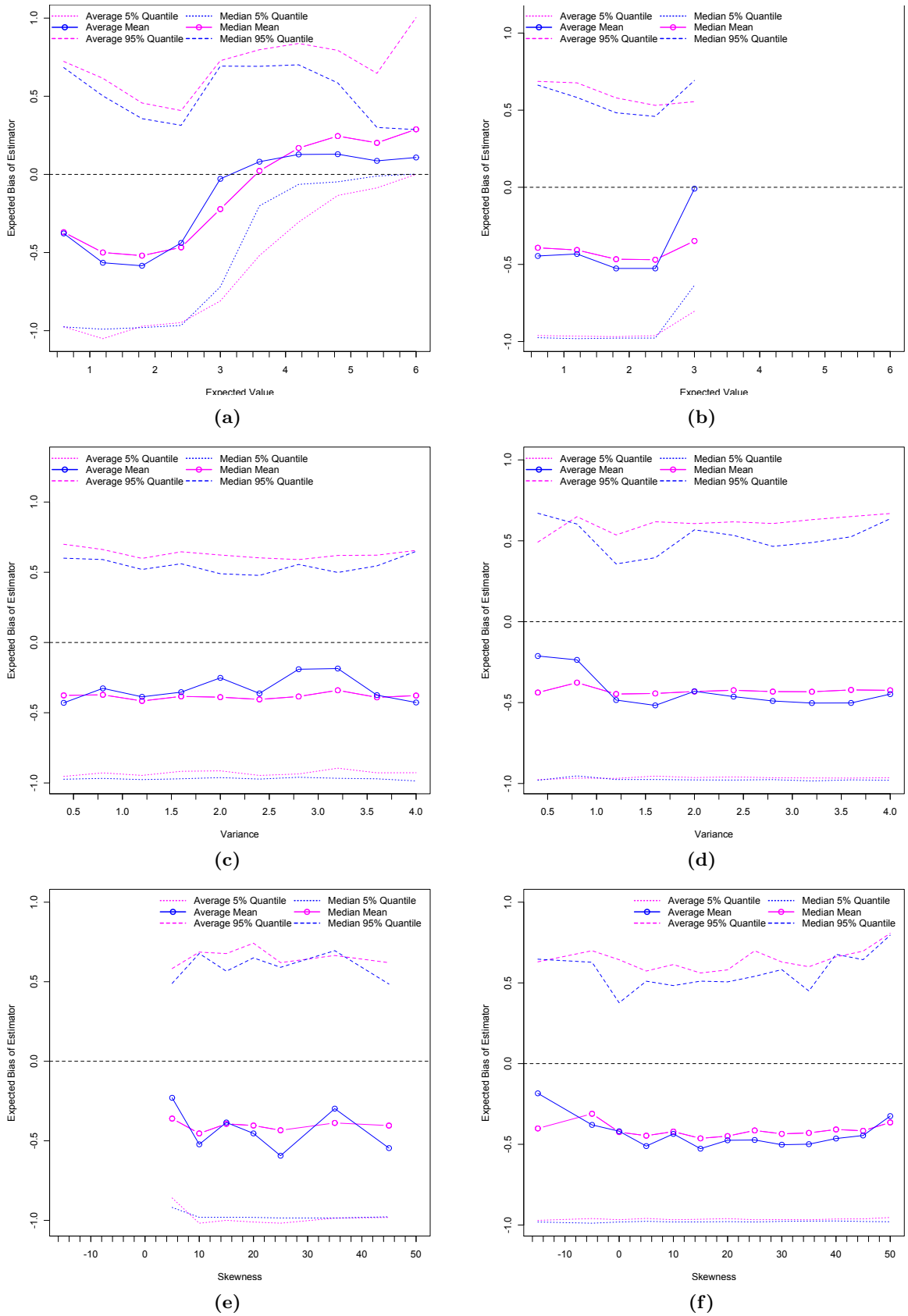


Figure E.19: *Expected bias of sampling distribution and error bounds for MITC estimation method illustrated as a function of expected value (figs. a and b), variance (figs. c and d) and skewness (figs. e and f) of distribution used to model detection probability. Detection probability in Figures a, c and e modelled by a cumulative Log-Normal distribution, while modelled by a cumulative Pareto type II distribution in Figures b, d and f.*

E.1.1.2 Results of threshold estimation methods incorporating threshold as factor

Goodness of fit estimation method

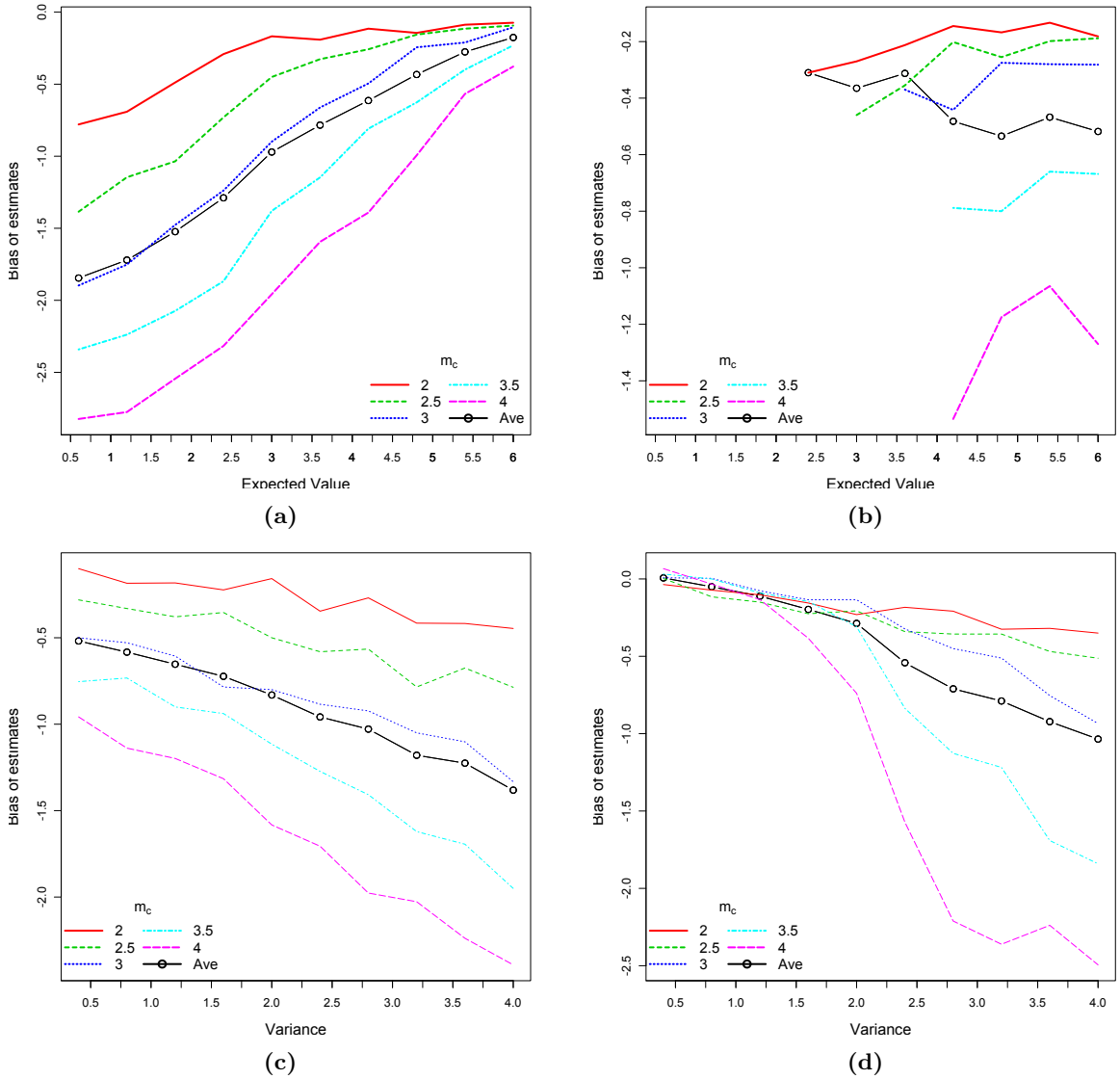


Figure E.20: *Bias of estimates for GOF estimation method illustrated as a function of expected value (figs. a and b) and variance (figs. c and d) of distribution used to model detection probability. Detection probability in Figures a and c modelled by a cumulative Normal distribution, while modelled by a cumulative logistic distribution in Figures b and d.*

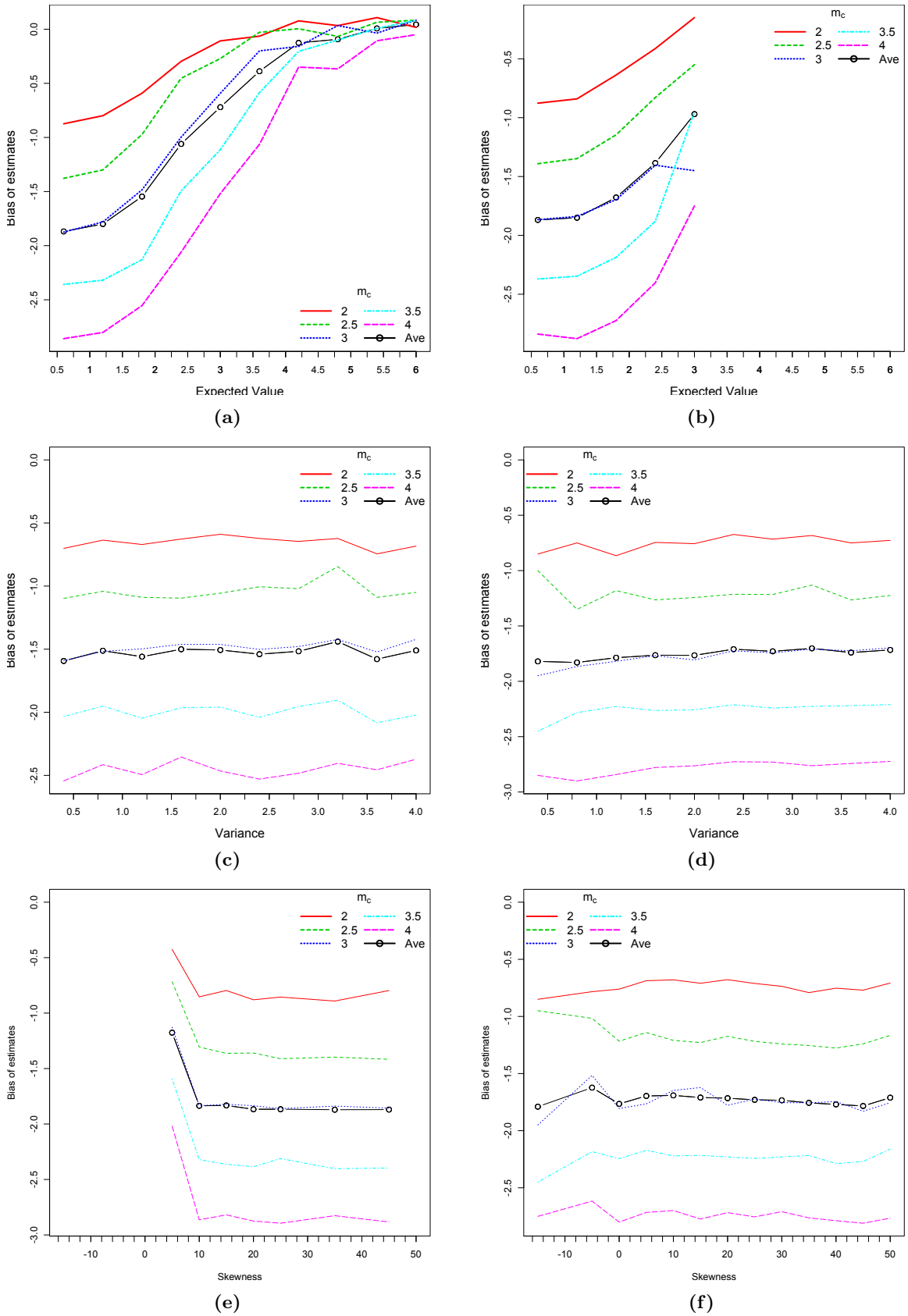


Figure E.21: *Bias of estimates for GOF estimation method illustrated as a function of expected value (figs. a and b), variance (figs. c and d) and skewness (figs. e and f) of distribution used to model detection probability. Detection probability in Figures a, c and e modelled by a cumulative Log-Normal distribution, while modelled by a cumulative Pareto type II distribution in Figures b, d and f.*

Maximum Curvature estimation method

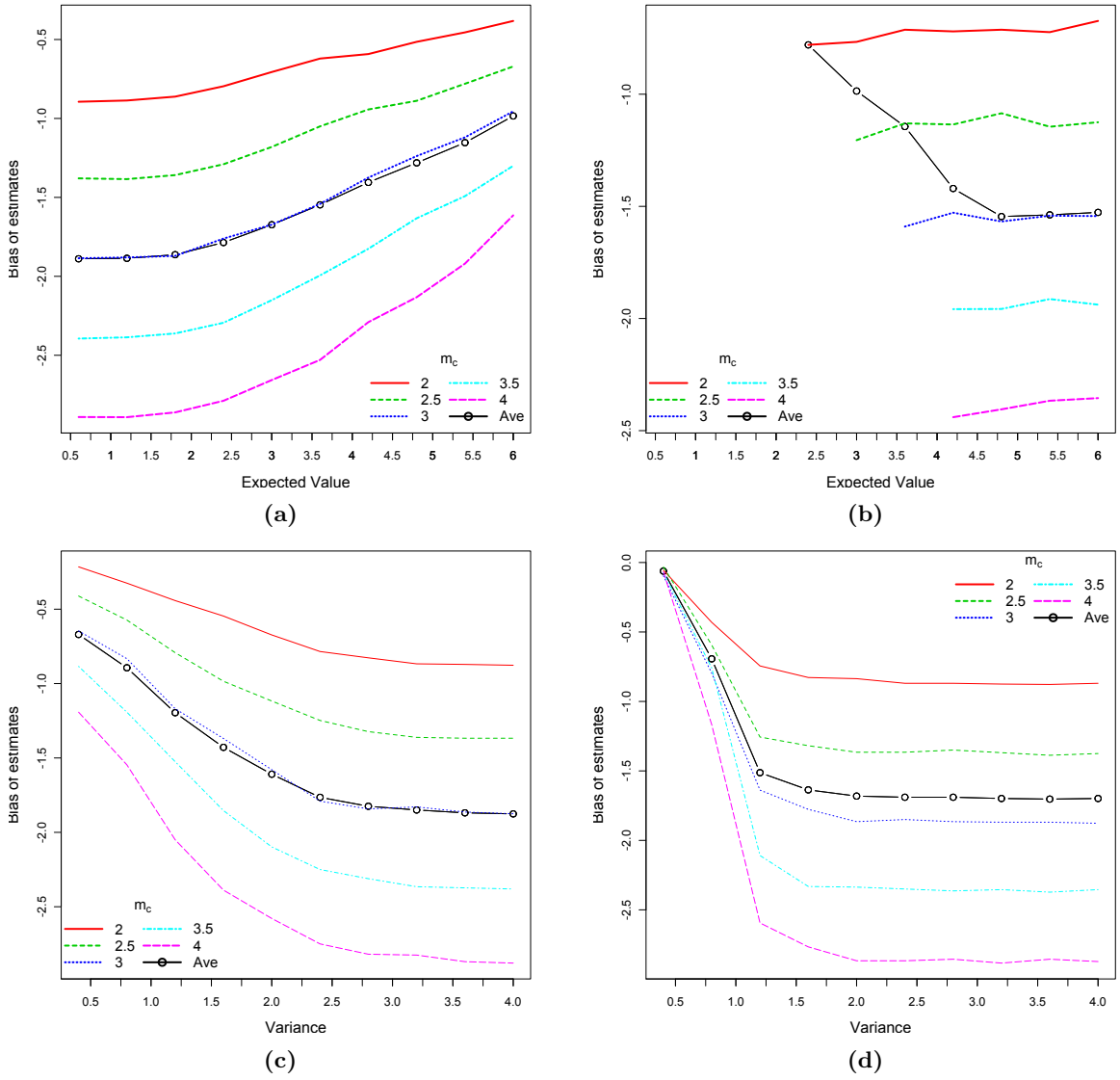


Figure E.22: *Bias of estimates for MAXC estimation method illustrated as a function of expected value (figs. a and b) and variance (figs. c and d) of distribution used to model detection probability. Detection probability in Figures a and c modelled by a cumulative Normal distribution, while modelled by a cumulative logistic distribution in Figures b and d.*

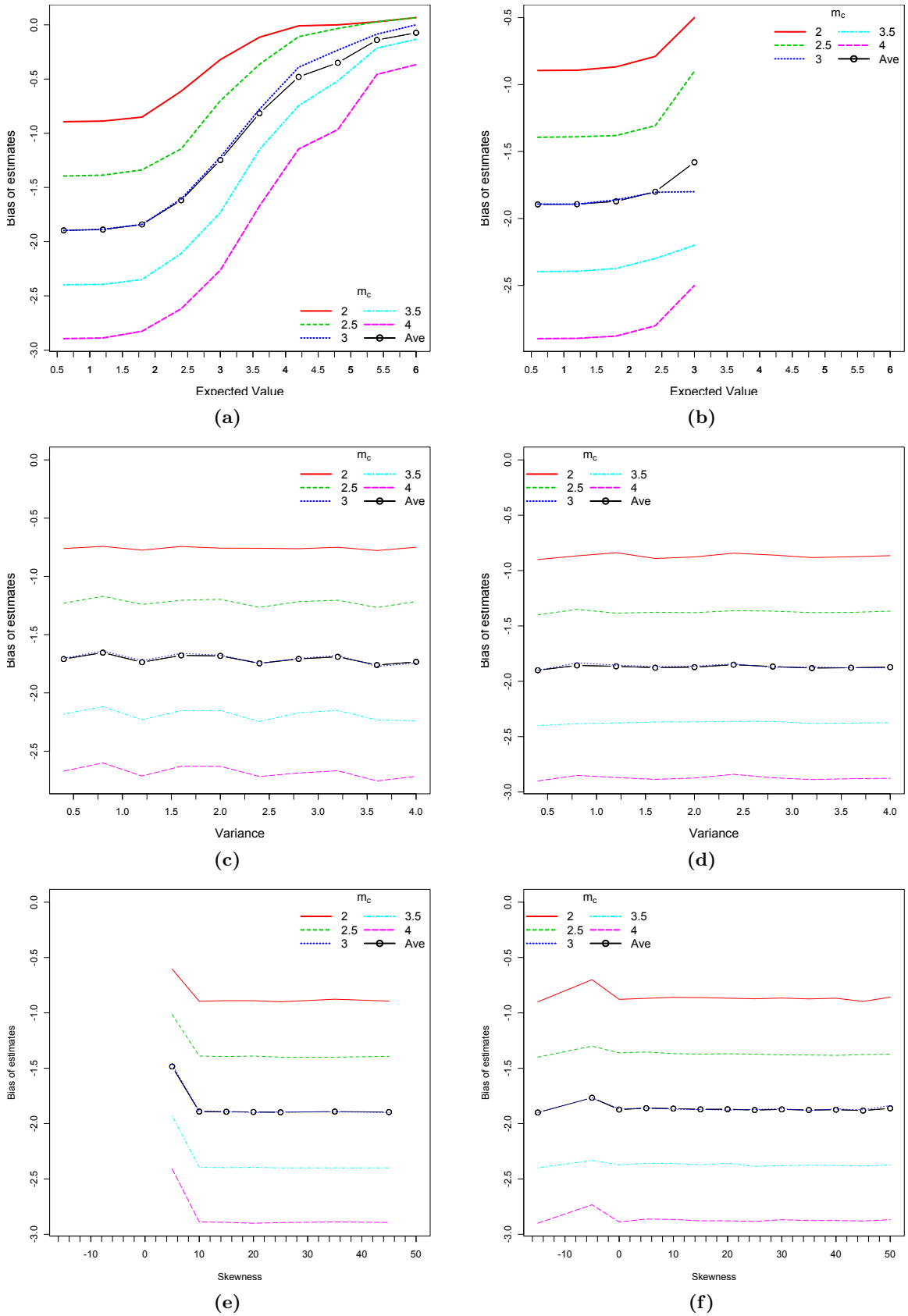


Figure E.23: *Bias of estimates for MAXC estimation method illustrated as a function of expected value (figs. a and b), variance (figs. c and d) and skewness (figs. e and f) of distribution used to model detection probability. Detection probability in Figures a, c and e modelled by a cumulative Log-Normal distribution, while modelled by a cumulative Pareto type II distribution in Figures b, d and f.*

Method of m_c estimation by b-Value stability

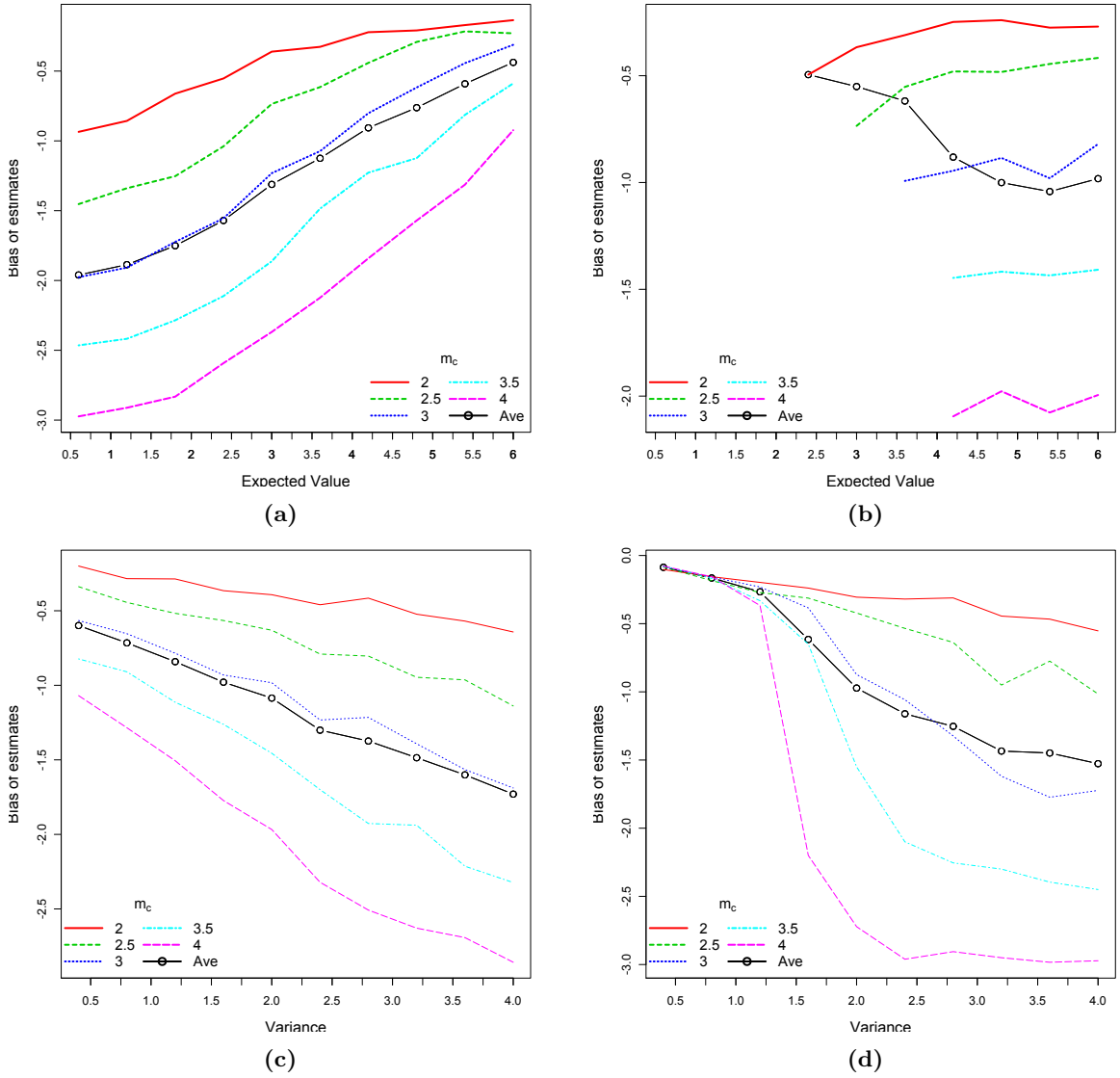


Figure E.24: *Bias of estimates for bVS estimation method illustrated as a function of expected value (figs. a and b) and variance (figs. c and d) of distribution used to model detection probability. Detection probability in Figures a and c modelled by a cumulative Normal distribution, while modelled by a cumulative logistic distribution in Figures b and d.*

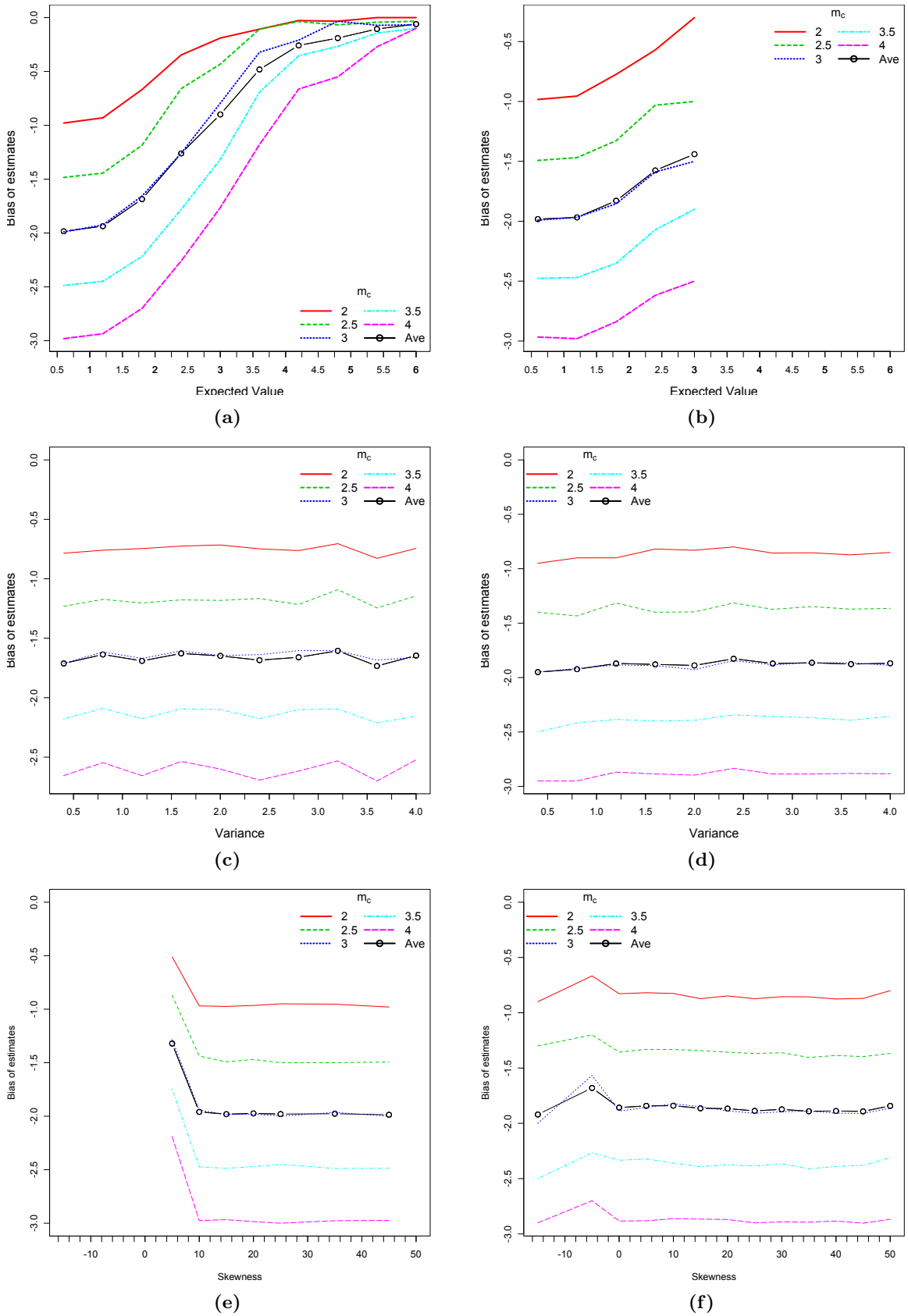


Figure E.25: *Bias of estimates for bVS estimation method illustrated as a function of expected value (figs. a and b), variance (figs. c and d) and skewness (figs. e and f) of distribution used to model detection probability. Detection probability in Figures a, c and e modelled by a cumulative Log-Normal distribution, while modelled by a cumulative Pareto type II distribution in Figures b, d and f.*

Entire magnitude range threshold estimation method

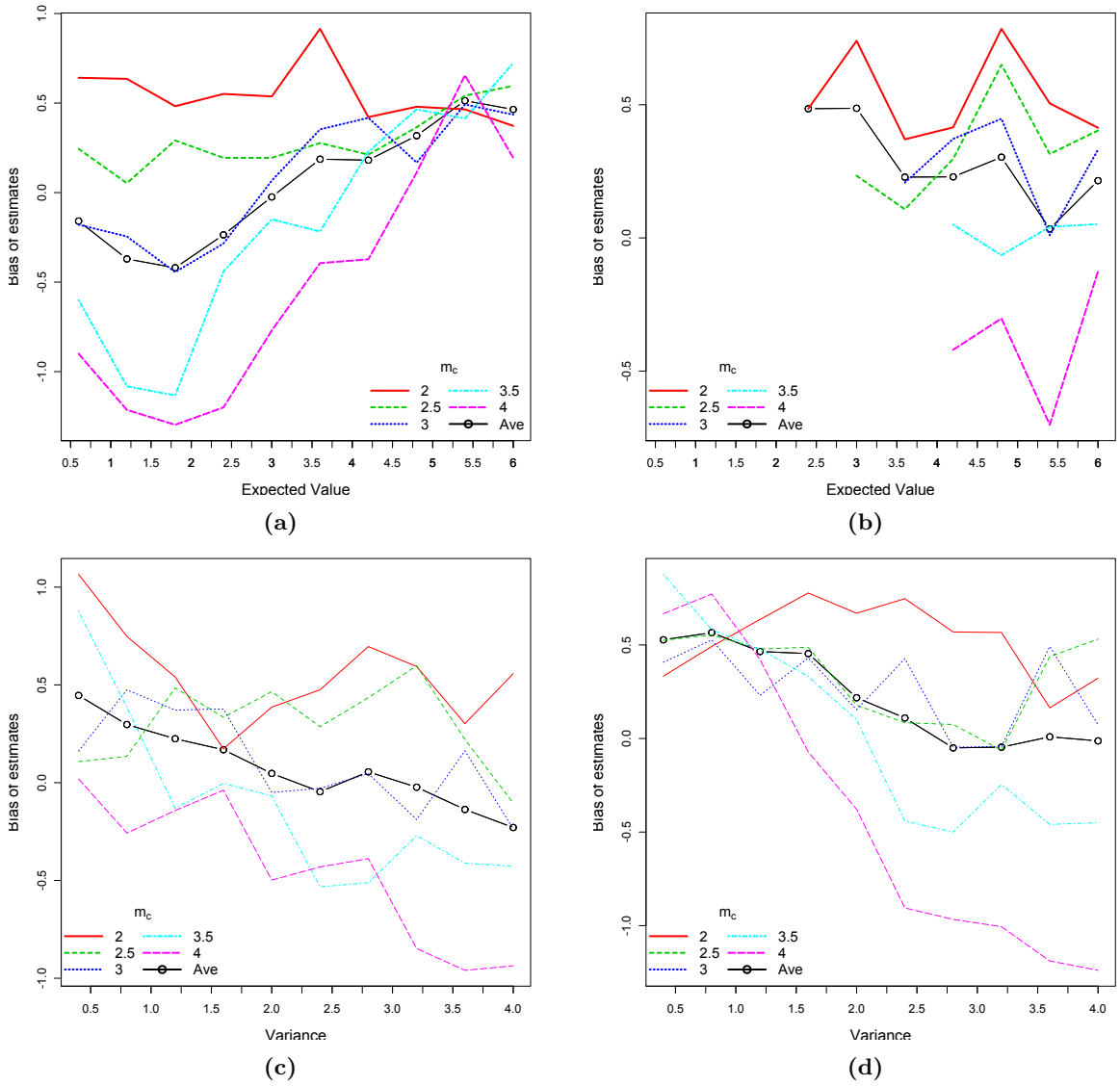


Figure E.26: *Bias of estimates for EMR estimation method illustrated as a function of expected value (figs. a and b) and variance (figs. c and d) of distribution used to model detection probability. Detection probability in Figures a and c modelled by a cumulative Normal distribution, while modelled by a cumulative logistic distribution in Figures b and d.*

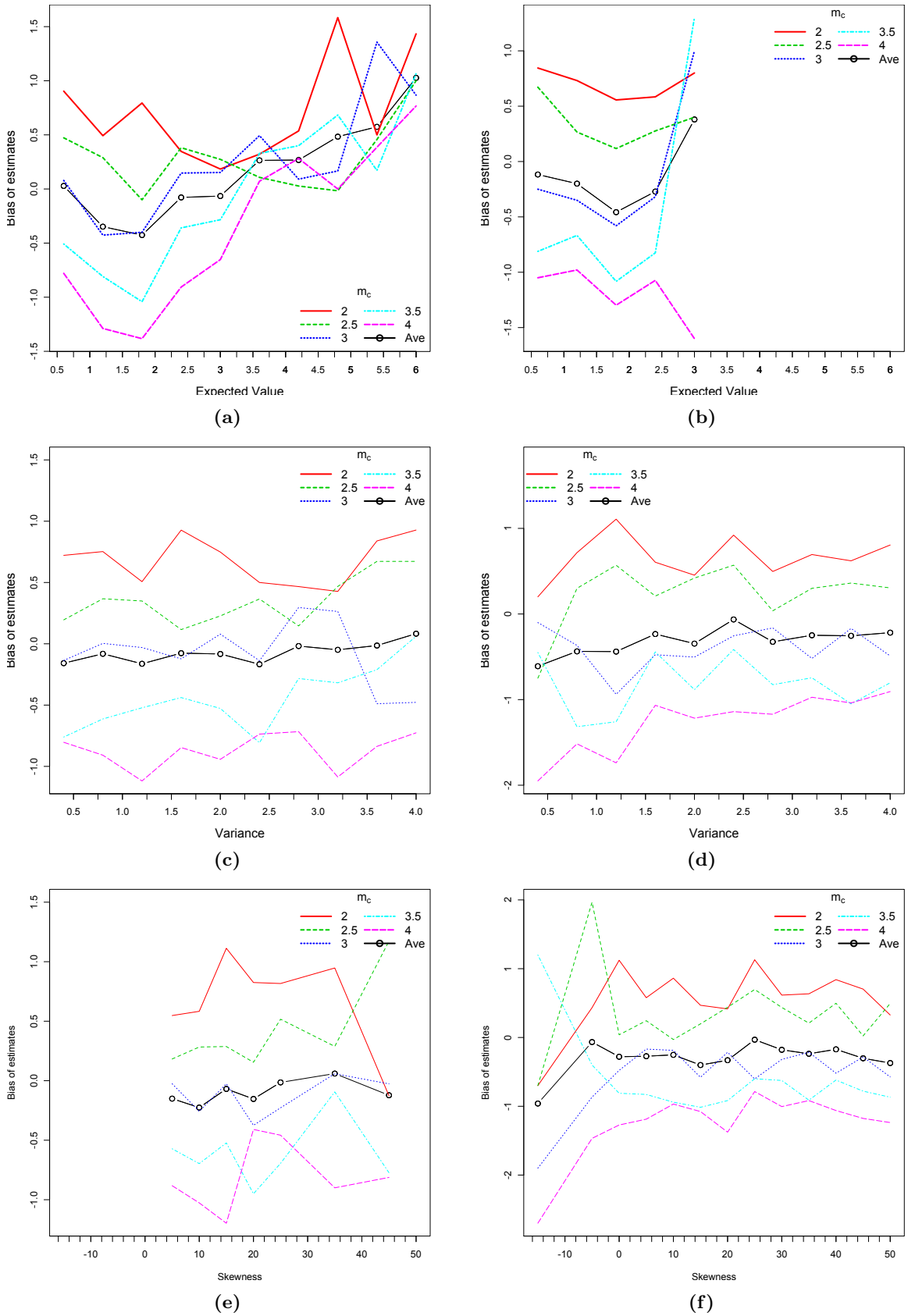


Figure E.27: *Bias of estimates for EMR estimation method illustrated as a function of expected value (figs. a and b), variance (figs. c and d) and skewness (figs. e and f) of distribution used to model detection probability. Detection probability in Figures a, c and e modelled by a cumulative Log-Normal distribution, while modelled by a cumulative Pareto type II distribution in Figures b, d and f.*

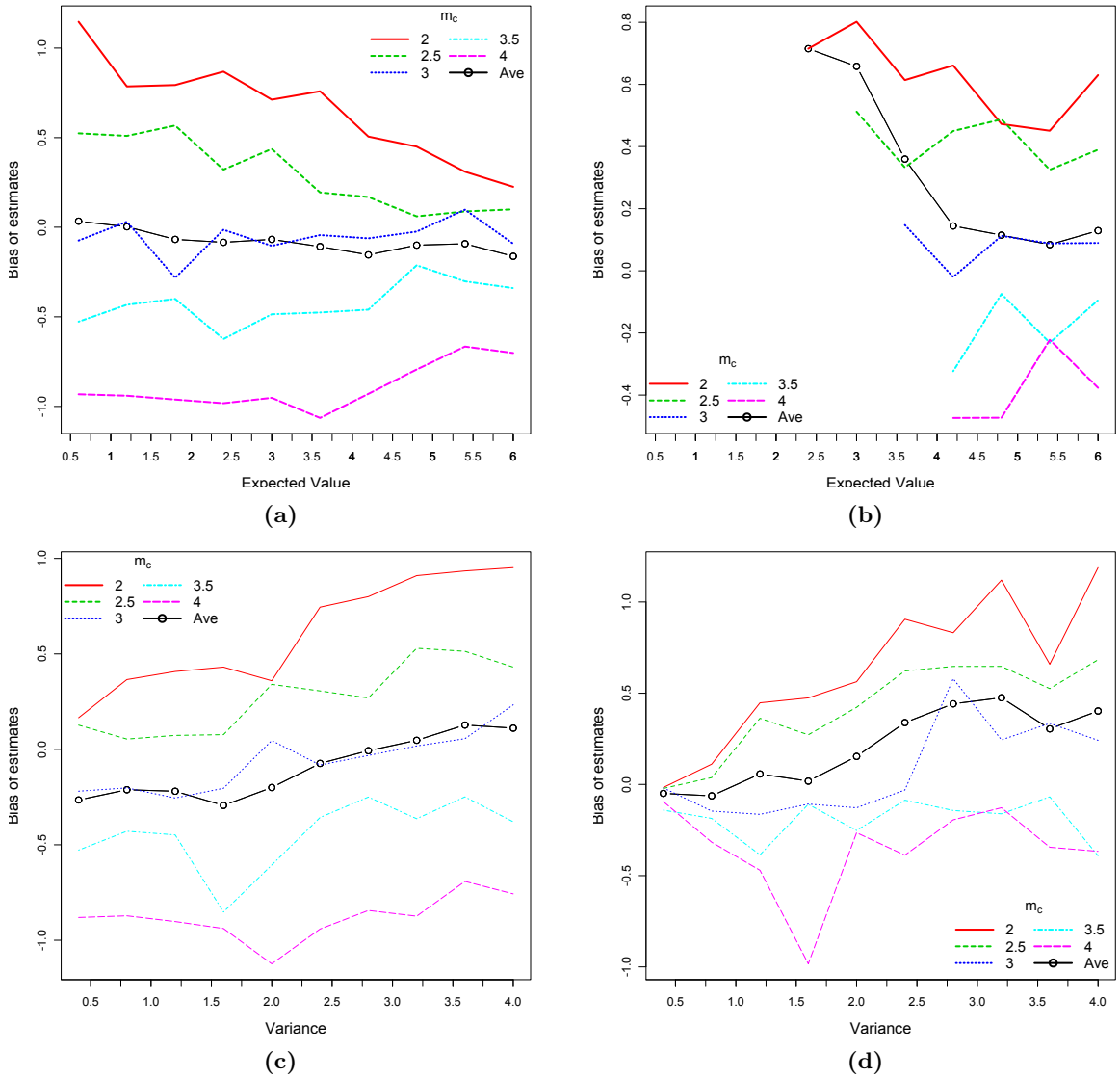


Figure E.28: *Bias of estimates for MBASS estimation method illustrated as a function of expected value (figs. a and b) and variance (figs. c and d) of distribution used to model detection probability. Detection probability in Figures a and c modelled by a cumulative Normal distribution, while modelled by a cumulative logistic distribution in Figures b and d.*

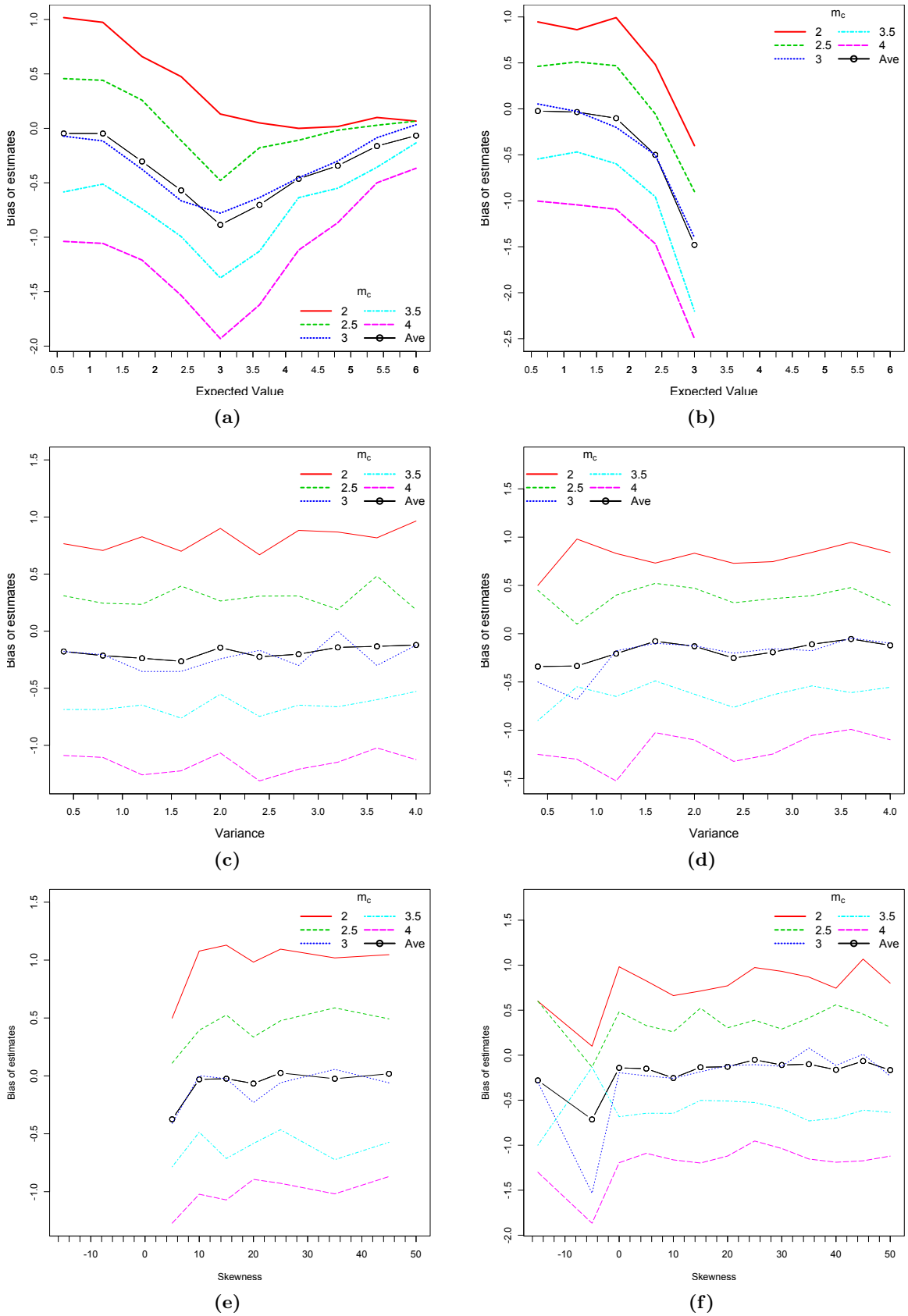


Figure E.29: *Bias of estimates for MBASS estimation method illustrated as a function of expected value (figs. a and b), variance (figs. c and d) and skewness (figs. e and f) of distribution used to model detection probability. Detection probability in Figures a, c and e modelled by a cumulative Log-Normal distribution, while modelled by a cumulative Pareto type II distribution in Figures b, d and f.*

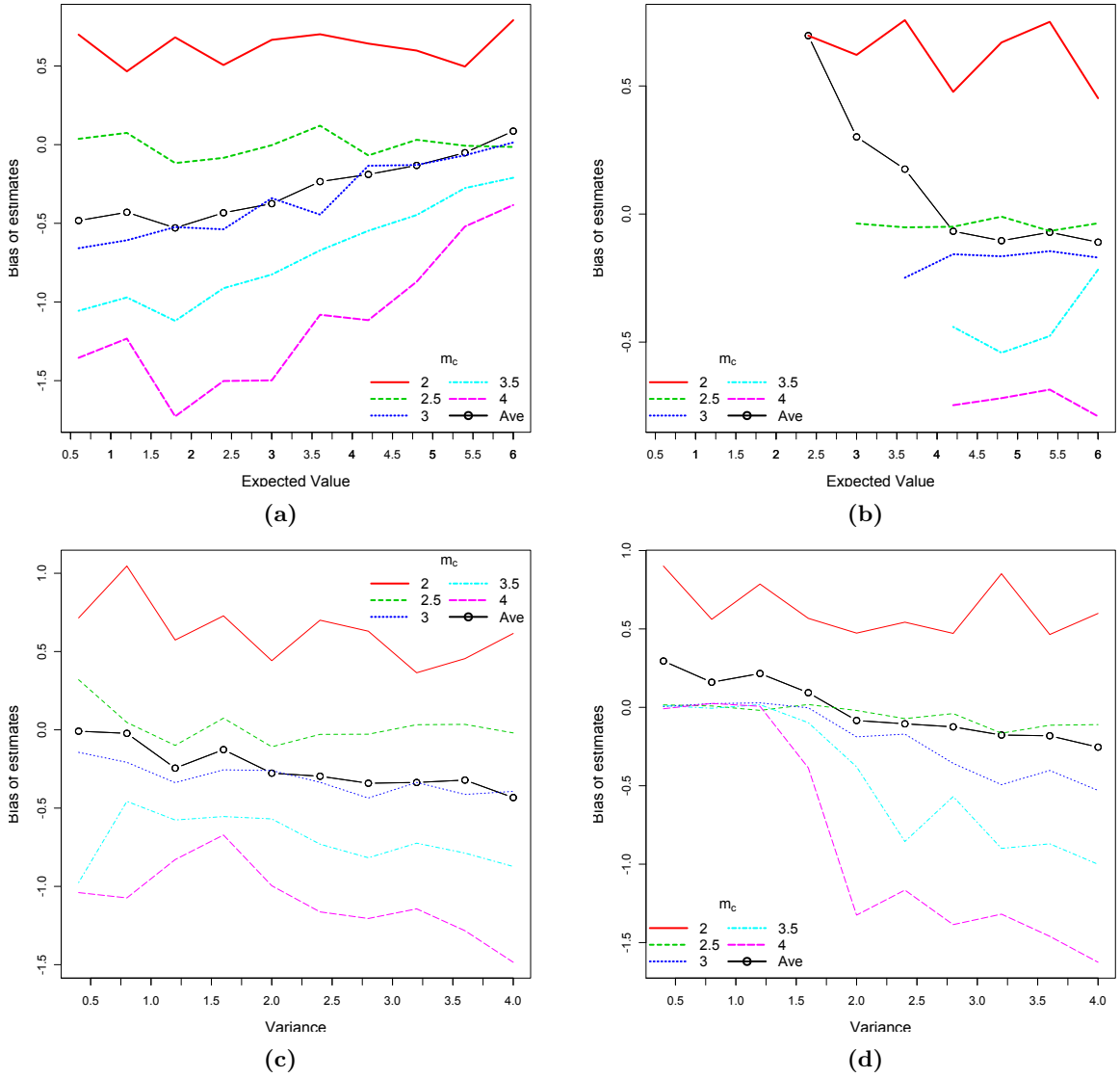


Figure E.30: *Bias of estimates for MITC estimation method illustrated as a function of expected value (figs. a and b) and variance (figs. c and d) of distribution used to model detection probability. Detection probability in Figures a and c modelled by a cumulative Normal distribution, while modelled by a cumulative logistic distribution in Figures b and d.*

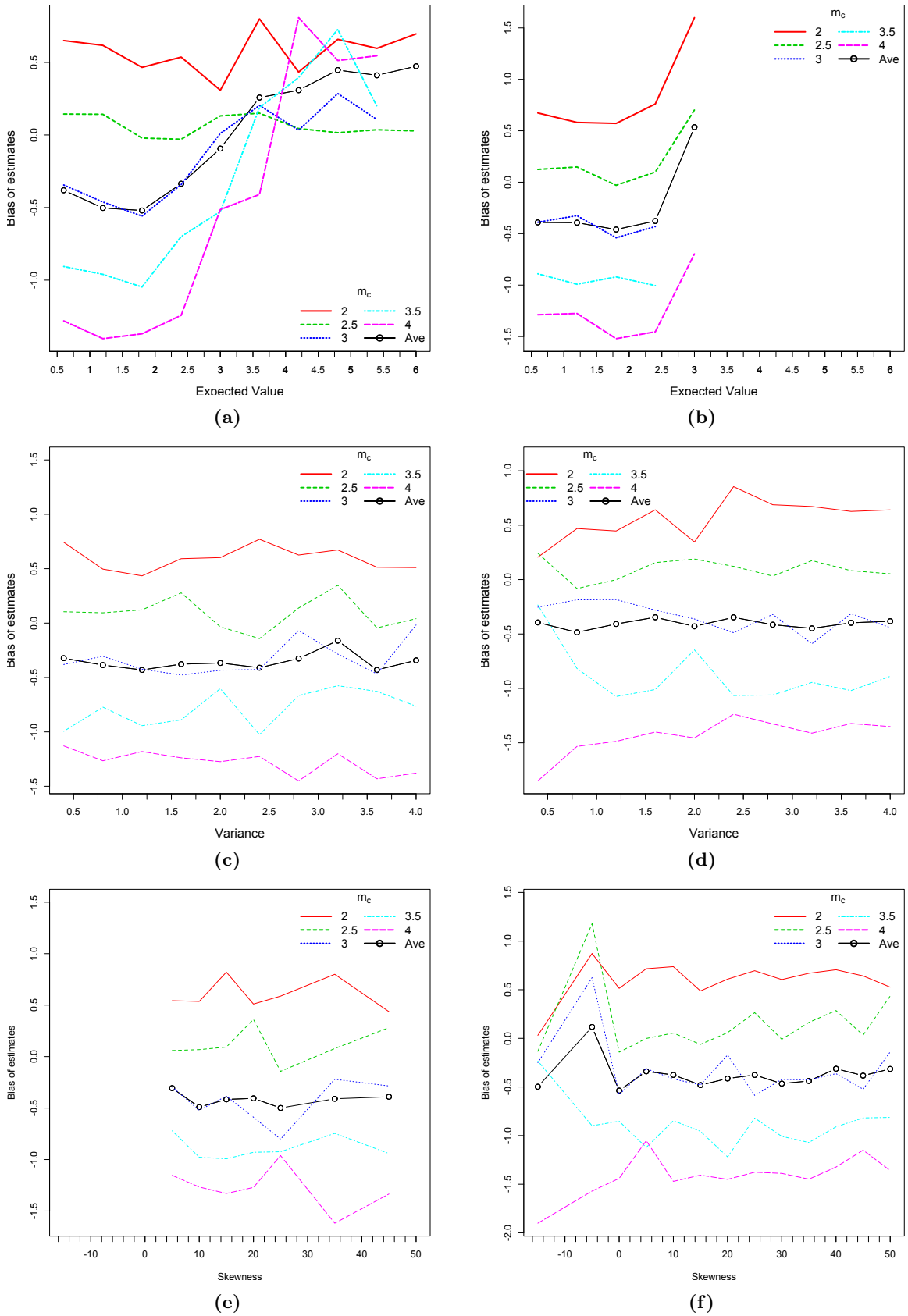


Figure E.31: *Bias of estimates for MITC estimation method illustrated as a function of expected value (figs. a and b), variance (figs. c and d) and skewness (figs. e and f) of distribution used to model detection probability. Detection probability in Figures a, c and e modelled by a cumulative Log-Normal distribution, while modelled by a cumulative Pareto type II distribution in Figures b, d and f.*

E.1.1.3 MSE summary of threshold estimation methods

A summary is given of the MSE values for the various threshold estimation methods. The results are displayed in 5 tables (Tables E.5 to E.10), with respective values of the detection threshold ranging from 2 to 4 in increments of 0.5. A sixth table is included where averaged results over the entire detection threshold spectrum is shown.

In addition to being calculated for different values of the detection threshold and estimation methods, the MSE is tabulated for various values of the 1) Expected Value 2) Variance and 3) Skewness (if applicable) of the random variable whose cumulative distribution function is being used to model the detection probability.

The range of these lower order moments will be, respectively, divided into 3 intervals of equal length. Table E.4 illustrates how these intervals are also labelled as either "Low values" (L), "Mid values" (M) or "High values" (H).

Lower order moments	Categories of values		
	Low (L)	Mid (M)	High (H)
Expected Value	(0,2]	(2,4]	(4,6]
Variance	(0,1.33]	(1.33,2.67]	(2.67,4]
Skewness	(-20,3.33]	(3.33,26.7]	(26.7,50]

Table E.4: *Table illustrating intervals of low (L), mid (M) and high (H) values of lower order moments of random variable whose distribution function is used to model the event detection probability.*

Grouping of event detection distribution by ...		Threshold estimation method	Detection probability modelled by ...											
			Cum. Normal Dist.			Cum. Logistic Dist.			Cum. Log-Normal Dist.			Cum. Pareto type II Dist.		
			L	M	H	L	M	H	L	M	H	L	M	H
Expected Value		GOF	0.595	0.189	0.100	-	0.187	0.137	0.723	0.102	0.053	0.657	0.215	-
		MAXC	0.774	0.578	0.385	-	0.638	0.615	0.780	0.214	0.006	0.771	0.483	-
		bVS	0.722	0.253	0.074	-	0.232	0.130	0.859	0.079	0.017	0.801	0.267	-
		EMR	2.307	2.196	1.283	-	1.806	1.645	2.416	0.698	3.337	2.367	2.402	-
		MBASS	1.416	1.290	0.891	-	1.430	1.153	1.332	0.611	0.129	1.327	0.858	-
		MITC	1.294	1.568	1.491	-	1.607	1.409	1.375	1.424	1.509	1.286	1.844	-
Variance		GOF	0.227	0.261	0.352	0.077	0.153	0.244	0.615	0.565	0.617	0.756	0.616	0.646
		MAXC	0.291	0.608	0.769	0.335	0.734	0.769	0.662	0.665	0.660	0.739	0.757	0.766
		bVS	0.199	0.315	0.468	0.044	0.146	0.332	0.720	0.684	0.703	0.858	0.726	0.800
		EMR	2.467	1.441	1.927	1.532	2.160	1.457	2.320	2.140	1.939	2.555	2.296	2.379
		MBASS	0.697	1.167	1.618	0.646	1.406	1.747	1.123	1.199	1.276	1.245	1.277	1.329
		MITC	2.116	1.462	0.848	1.924	1.146	1.507	1.462	1.352	1.178	0.836	1.327	1.336
Skewness		GOF	NA			NA			0.300	0.760	0.798	0.611	0.617	0.688
		MAXC							0.441	0.799	0.785	0.743	0.759	0.772
		bVS							0.310	0.937	0.954	0.724	0.786	0.799
		EMR							2.085	2.271	2.076	2.954	2.379	2.217
		MBASS							0.882	1.436	1.324	1.381	1.263	1.364
		MITC							1.307	1.423	1.528	1.049	1.406	1.233
Averaged		GOF	0.283			0.162			0.603			0.644		
		MAXC	0.571			0.627			0.663			0.763		
		bVS	0.335			0.181			0.708			0.786		
		EMR	1.914			1.726			2.216			2.368		
		MBASS	1.186			1.290			1.166			1.313		
		MITC	1.464			1.508			1.390			1.306		

Table E.5: *MSE summary of threshold estimation methods where $m_c = 2$. Results grouped by 1) Estimation Method, 2) CDF representing the detection probability and 3) the lower order moments (Expected value, variance and Skewness) of RV attributed to detection probability.*

Grouping of event detection distribution by ...	Threshold estimation method	Detection probability modelled by ...											
		Cum. Normal Dist.			Cum. Logistic Dist.			Cum. Log-Normal Dist.			Cum. Pareto type II Dist.		
		L	M	H	L	M	H	L	M	H	L	M	H
Expected Value	GOF	1.514	0.512	0.154		0.441	0.252	1.729	0.234	0.090	1.675	0.889	-
	MAXC	1.886	1.471	1.004	-	1.561	1.535	1.903	0.740	0.019	1.907	1.342	-
	bVS	1.820	0.823	0.198	-	0.645	0.449	2.021	0.273	0.030	1.958	0.810	-
	EMR	1.829	1.230	1.666	-	1.190	1.612	1.967	0.452	0.933	2.053	0.487	-
	MBASS	1.480	1.518	0.996	-	1.651	1.595	1.275	0.626	0.231	1.363	0.892	-
	MITC	0.593	0.237	0.064	-	0.143	0.100	0.838	0.258	0.028	0.843	0.537	-
Variance	GOF	0.443	0.646	0.957	0.108	0.275	0.588	1.462	1.443	1.350	1.576	1.655	1.656
	MAXC	0.789	1.539	1.875	0.802	1.839	1.892	1.634	1.666	1.680	1.903	1.885	1.891
	bVS	0.496	0.834	1.341	0.072	0.414	1.052	1.692	1.672	1.640	1.893	1.913	1.931
	EMR	1.358	1.555	1.749	1.549	1.174	1.577	1.753	1.421	1.964	1.789	2.117	1.987
	MBASS	0.811	1.415	1.672	0.935	1.932	1.906	1.111	1.123	1.212	1.383	1.366	1.340
	MITC	0.246	0.215	0.366	0.030	0.136	0.194	0.748	0.601	0.828	0.583	0.748	0.879
Skewness	GOF							0.711	1.852	2.016	1.538	1.661	1.668
	MAXC							1.166	1.938	1.952	1.829	1.888	1.908
	bVS							0.880	2.131	2.247	1.820	1.905	1.976
	EMR		NA			NA		1.216	1.990	1.843	2.144	1.852	2.181
	MBASS							0.926	1.281	1.281	1.435	1.324	1.360
	MITC							0.424	0.915	0.881	0.566	0.717	1.041
Averaged	GOF		0.696			0.335			1.440			1.652	
	MAXC		1.437			1.547			1.649			1.890	
	bVS		0.913			0.535			1.679			1.925	
	EMR		1.565			1.427			1.704			2.008	
	MBASS		1.327			1.619			1.130			1.348	
	MITC		0.276			0.119			0.725			0.832	

Table E.6: *MSE summary of threshold estimation methods where $m_c = 2.5$. Results grouped by 1) Estimation Method, 2) CDF representing the detection probability and 3) the lower order moments (Expected value, variance and Skewness) of RV attributed to detection probability.*

Grouping of event detection distribution by ...	Threshold estimation method	Detection probability modelled by ...											
		Cum. Normal Dist.			Cum. Logistic Dist.			Cum. Log-Normal Dist.			Cum. Pareto type II Dist.		
		L	M	H	L	M	H	L	M	H	L	M	H
Expected Value	GOF	3.015	1.180	0.364	-	0.719	0.593	3.225	0.625	0.088	3.187	1.786	-
	MAXC	3.502	2.852	1.952	-	2.861	2.809	3.538	1.627	0.114	3.517	3.029	-
	bVS	3.491	1.952	0.582	-	1.571	1.486	3.655	0.892	0.042	3.621	2.349	-
	EMR	2.150	1.844	1.090	-	1.709	1.382	1.780	0.976	1.931	2.012	1.074	-
	MBASS	1.507	1.781	1.656	-	1.945	2.116	1.208	0.720	0.685	1.385	0.833	-
	MITC	0.855	0.459	0.162	-	0.291	0.311	0.996	0.550	0.115	1.031	0.844	-
Variance	GOF	0.887	1.351	2.052	0.148	0.457	1.221	2.777	2.623	2.579	3.461	3.202	3.105
	MAXC	1.577	2.994	3.457	1.361	3.403	3.502	3.081	3.102	3.203	3.442	3.449	3.525
	bVS	1.041	1.836	2.825	0.076	1.252	3.004	3.114	3.017	3.127	3.641	3.587	3.579
	EMR	1.792	1.547	1.711	1.108	1.395	1.891	1.607	1.663	1.943	2.225	2.043	1.948
	MBASS	1.044	1.826	2.004	1.291	2.158	2.654	1.099	1.083	1.205	1.111	1.378	1.385
	MITC	0.456	0.393	0.567	0.037	0.338	0.600	0.898	0.842	1.091	0.791	1.149	1.007
Skewness	GOF							1.531	3.364	3.487	3.266	3.048	3.247
	MAXC							2.308	3.582	3.599	3.415	3.519	3.501
	bVS							1.769	3.866	3.938	3.463	3.560	3.643
	EMR		NA			NA		1.501	1.720	1.483	1.926	2.045	1.919
	MBASS							0.990	1.257	1.187	1.384	1.337	1.405
	MITC							0.652	0.882	1.157	0.898	1.062	1.007
Averaged	GOF		1.462			0.632			2.708			3.146	
	MAXC		2.741			2.825			3.105			3.503	
	bVS		1.951			1.512			3.092			3.584	
	EMR		1.677			1.482			1.673			1.985	
	MBASS		1.655			2.065			1.111			1.369	
	MITC		0.470			0.305			0.913			1.026	

Table E.7: *MSE summary of threshold estimation methods where $m_c = 3$. Results grouped by 1) Estimation Method, 2) CDF representing the detection probability and 3) the lower order moments (Expected value, variance and Skewness) of RV attributed to detection probability.*

Grouping of event detection distribution by ...		Threshold estimation method	Detection probability modelled by ...											
			Cum. Normal Dist.			Cum. Logistic Dist.			Cum. Log-Normal Dist.			Cum. Pareto type II Dist.		
			L	M	H	L	M	H	L	M	H	L	M	H
Grouping of event detection distribution by ...	Expected Value	GOF	4.976	2.594	0.754	-	1.967	1.608	5.336	1.484	0.105	5.167	2.891	-
		MAXC	5.641	4.720	3.255	-	4.481	4.491	5.680	2.980	0.396	5.651	5.041	-
		bVS	5.703	3.549	1.539	-	3.005	3.253	5.877	1.875	0.100	5.754	3.403	-
		EMR	2.658	2.147	2.067	-	2.111	1.796	2.452	1.332	0.912	2.557	1.709	-
		MBASS	1.552	1.941	2.457	-	2.479	2.774	1.265	0.961	1.365	1.407	1.140	-
		MITC	1.592	1.089	0.587	-	0.797	0.739	1.666	1.208	0.474	1.655	1.550	-
	Variance	GOF	1.586	2.440	3.859	0.149	1.084	3.570	4.574	4.438	4.588	5.231	5.115	5.087
		MAXC	2.673	4.871	5.640	2.023	5.498	5.596	5.027	5.059	5.107	5.674	5.598	5.643
		bVS	1.874	3.375	5.025	0.150	3.272	5.764	5.048	4.988	5.068	5.818	5.681	5.679
		EMR	2.297	2.023	2.513	1.676	1.477	2.381	2.337	2.060	2.033	2.843	2.164	2.634
		MBASS	1.678	2.021	2.268	2.179	2.989	2.921	1.179	1.237	1.385	1.340	1.389	1.406
		MITC	0.892	0.933	1.357	0.063	0.836	1.433	1.613	1.620	1.380	1.183	1.566	1.704
	Skewness	GOF	NA			NA			2.887	5.469	5.810	5.018	5.094	5.129
		MAXC							3.907	5.737	5.761	5.566	5.624	5.662
		bVS							3.224	6.103	6.196	5.495	5.686	5.732
		EMR							1.921	2.633	2.212	2.320	2.603	2.490
		MBASS							1.228	1.271	1.028	1.409	1.419	1.369
		MITC							1.367	1.656	1.791	1.528	1.685	1.635
	Averaged	GOF	2.690			1.673			4.542			5.101		
		MAXC	4.496			4.490			5.048			5.634		
		bVS	3.516			3.208			5.036			5.686		
		EMR	2.277			1.853			2.221			2.533		
		MBASS	2.004			2.719			1.226			1.399		
		MITC	1.060			0.751			1.576			1.652		

Table E.8: *MSE summary of threshold estimation methods where $m_c = 3.5$. Results grouped by 1) Estimation Method, 2) CDF representing the detection probability and 3) the lower order moments (Expected value, variance and Skewness) of RV attributed to detection probability.*

Grouping of event detection distribution by ...		Threshold estimation method	Detection probability modelled by ...											
			Cum. Normal Dist.			Cum. Logistic Dist.			Cum. Log-Normal Dist.			Cum. Pareto type II Dist.		
			L	M	H	L	M	H	L	M	H	L	M	H
Grouping of event detection distribution by ...	Expected Value	GOF	7.395	4.141	1.486	-	-	3.197	7.792	2.590	0.166	7.744	5.556	-
		MAXC	8.260	7.079	4.925	-	-	6.713	8.276	4.876	0.944	8.295	7.330	-
		bVS	8.396	5.774	2.844	-	-	5.836	8.480	3.227	0.273	8.400	6.648	-
		EMR	3.260	2.518	2.298	-	-	2.355	3.005	1.418	0.693	3.225	3.248	-
		MBASS	1.617	1.961	2.919	-	-	3.882	1.323	0.957	1.860	1.370	1.057	-
		MITC	2.923	2.313	1.174	-	-	1.530	2.758	1.908	0.938	2.816	1.812	-
	Variance	GOF	2.607	4.021	5.773	0.167	2.790	6.201	6.737	6.633	6.625	8.236	7.677	7.642
		MAXC	4.200	7.294	8.180	3.354	8.061	8.244	7.409	7.426	7.627	8.228	8.192	8.295
		bVS	3.049	5.546	7.684	0.318	7.561	8.840	7.358	7.355	7.287	8.449	8.262	8.371
		EMR	2.270	2.267	3.406	1.468	2.316	3.155	2.745	2.693	2.387	3.850	3.199	3.190
		MBASS	1.973	2.137	2.449	3.561	4.205	3.850	1.291	1.277	1.344	0.947	1.314	1.405
		MITC	1.701	1.870	2.610	0.115	2.082	2.670	2.591	2.621	2.880	2.811	2.528	2.880
	Skewness	GOF	NA			NA			4.350	8.069	8.213	7.736	7.574	7.808
		MAXC							5.944	8.350	8.357	8.197	8.277	8.270
		bVS							4.929	8.789	8.862	8.216	8.326	8.411
		EMR							2.431	3.094	1.868	3.244	3.216	3.232
		MBASS							1.390	1.279	1.379	1.250	1.456	1.261
		MITC							2.376	2.742	2.838	2.521	2.810	2.822
	Averaged	GOF	4.223			3.197			6.694			7.680		
		MAXC	6.697			6.713			7.447			8.267		
		bVS	5.566			5.836			7.346			8.349		
		EMR	2.670			2.355			2.675			3.225		
		MBASS	2.196			3.882			1.296			1.361		
		MITC	2.078			1.530			2.640			2.790		

Table E.9: *MSE summary of threshold estimation methods where $m_c = 4$. Results grouped by 1) Estimation Method, 2) CDF representing the detection probability and 3) the lower order moments (Expected value, variance and Skewness) of RV attributed to detection probability.*

Grouping of event detection distribution by ...	Threshold estimation method	Detection probability modelled by ... Cum. Normal Dist. L	Cum. Normal Dist. M	Cum. Normal Dist. H	Cum. Logistic Dist. L	Cum. Logistic Dist. M	Cum. Logistic Dist. H	Cum. Log-Normal Dist. L	Cum. Log-Normal Dist. M	Cum. Log-Normal Dist. H	Cum. Pareto type II Dist. L	Cum. Pareto type II Dist. M	Cum. Pareto type II Dist. H
Expected Value	GOF	3.499	1.723	0.572	-	0.526	1.157	3.761	1.007	0.100	3.686	2.267	-
	MAXC	4.013	3.340	2.304	-	1.685	3.233	4.036	2.087	0.296	4.028	3.445	-
	bVS	4.026	2.470	1.047	-	0.859	2.231	4.178	1.269	0.093	4.107	2.695	-
	EMR	2.441	1.987	1.681	-	1.620	1.758	2.327	0.975	1.561	2.443	1.784	-
	MBASS	1.514	1.695	1.781	-	1.690	2.297	1.281	0.776	0.854	1.371	0.956	-
	MITC	1.444	1.131	0.728	-	0.851	0.805	1.590	1.050	0.806	1.534	1.307	-
Variance	GOF	1.150	1.744	2.599	0.122	0.748	1.862	3.233	3.140	3.152	3.852	3.653	3.627
	MAXC	1.906	3.461	3.984	1.328	3.282	3.364	3.563	3.584	3.655	3.997	3.976	4.024
	bVS	1.332	2.381	3.468	0.111	1.934	3.057	3.586	3.543	3.565	4.132	4.034	4.072
	EMR	2.037	1.767	2.261	1.469	1.685	1.949	2.153	1.998	2.055	2.651	2.364	2.428
	MBASS	1.236	1.715	1.996	1.477	2.305	2.441	1.160	1.184	1.285	1.204	1.345	1.373
	MITC	1.153	0.969	1.128	0.584	0.788	1.126	1.514	1.439	1.499	1.257	1.475	1.564
Skewness	GOF	NA			NA			1.956	3.903	4.065	3.634	3.599	3.708
	MAXC							2.753	4.081	4.091	3.950	4.014	4.023
	bVS							2.222	4.365	4.439	3.944	4.053	4.112
	EMR							1.832	2.344	1.896	2.517	2.419	2.408
	MBASS							1.084	1.304	1.236	1.371	1.360	1.352
	MITC							1.223	1.614	1.698	1.295	1.534	1.571
Averaged	GOF	1.871			0.950			3.197			3.645		
	MAXC	3.188			2.725			3.582			4.011		
	bVS	2.456			1.780			3.572			4.066		
	EMR	2.021			1.713			2.099			2.424		
	MBASS	1.670			2.099			1.186			1.358		
	MITC	1.075			0.821			1.493			1.527		

Table E.10: *MSE summary of threshold estimation methods averaged over all values of m_c. Results grouped by 1) Estimation Method, 2) CDF representing the detection probability and 3) lower order moments (Expected value, Variance and Skewness) of RV attributed to detection probability.*

E.1.2 Implicit modelling of detection threshold

E.1.2.1 Estimated proportion of events not detected

Goodness of fit estimation method

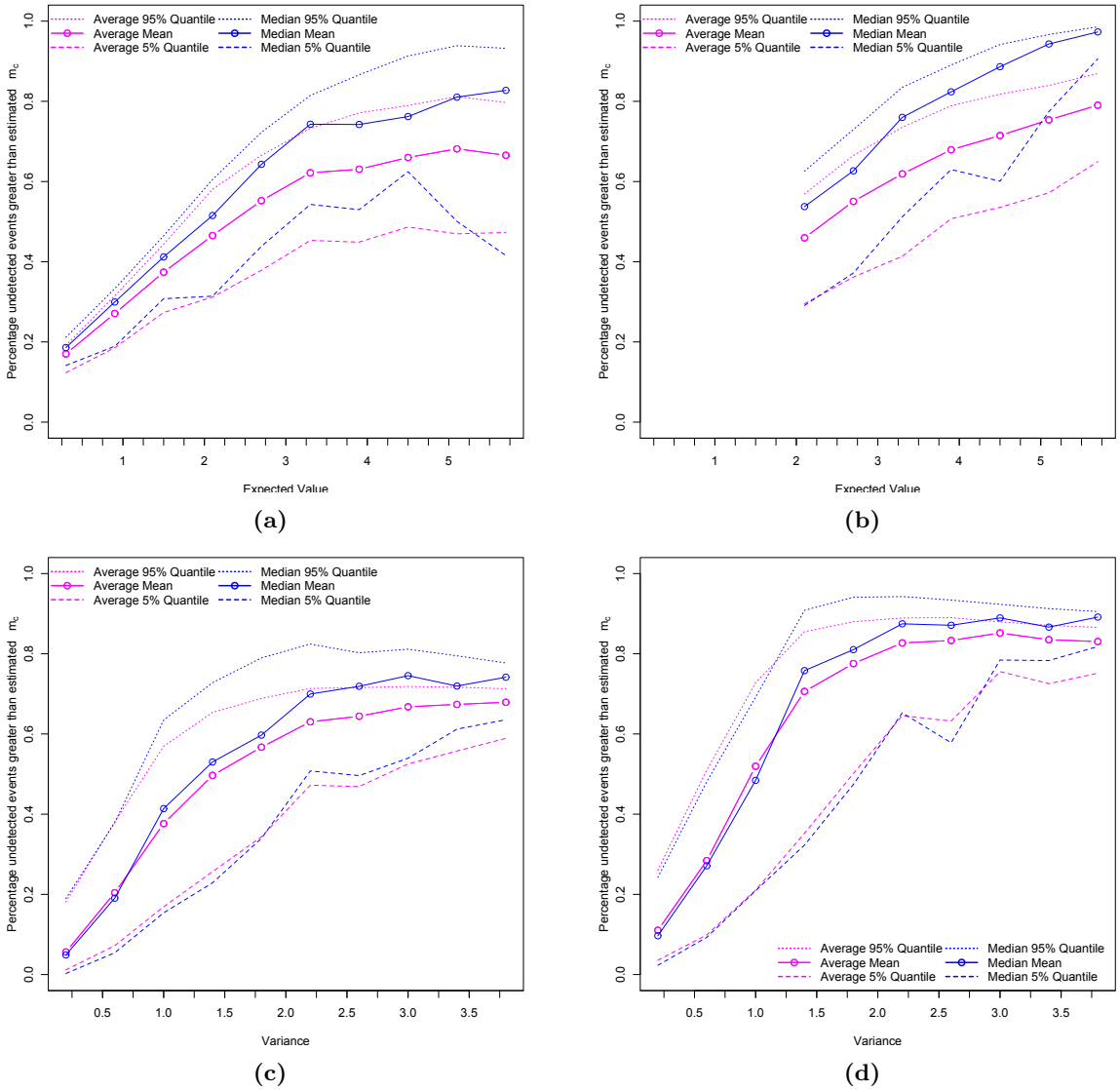


Figure E.32: *Expected number of events not detected as a percentage of expected number of events greater than $\hat{m}_c$ together with error bounds for GOF estimation method illustrated as a function of expected value (figs. a and b) and variance (figs. c and d) of distribution used to model detection probability. Detection probability in Figures a and c modelled by a cumulative Normal distribution, while modelled by a cumulative logistic distribution in Figures b and d.*

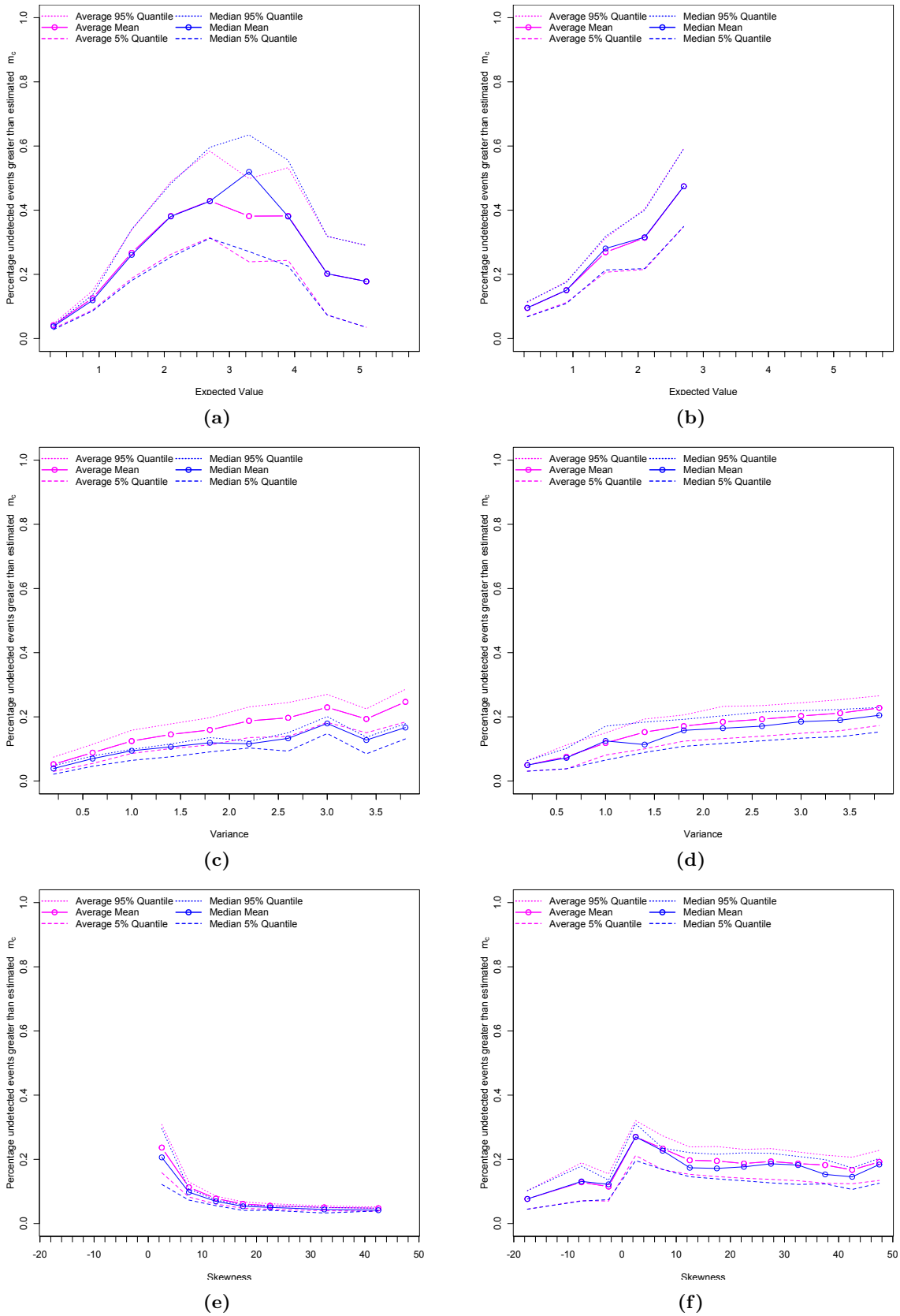


Figure E.33: *Expected number of events not detected as a percentage of expected number of events greater than $\hat{m}_c$ together with error bounds for GOF estimation method illustrated as a function of expected value (figs. a and b), variance (figs. c and d) and skewness (figs. e and f) of distribution used to model detection probability. Detection probability in Figures a, c and e modelled by a cumulative Log-Normal distribution, while modelled by a cumulative Pareto type II distribution in Figures b, d and f.*

Maximum curvature estimation method

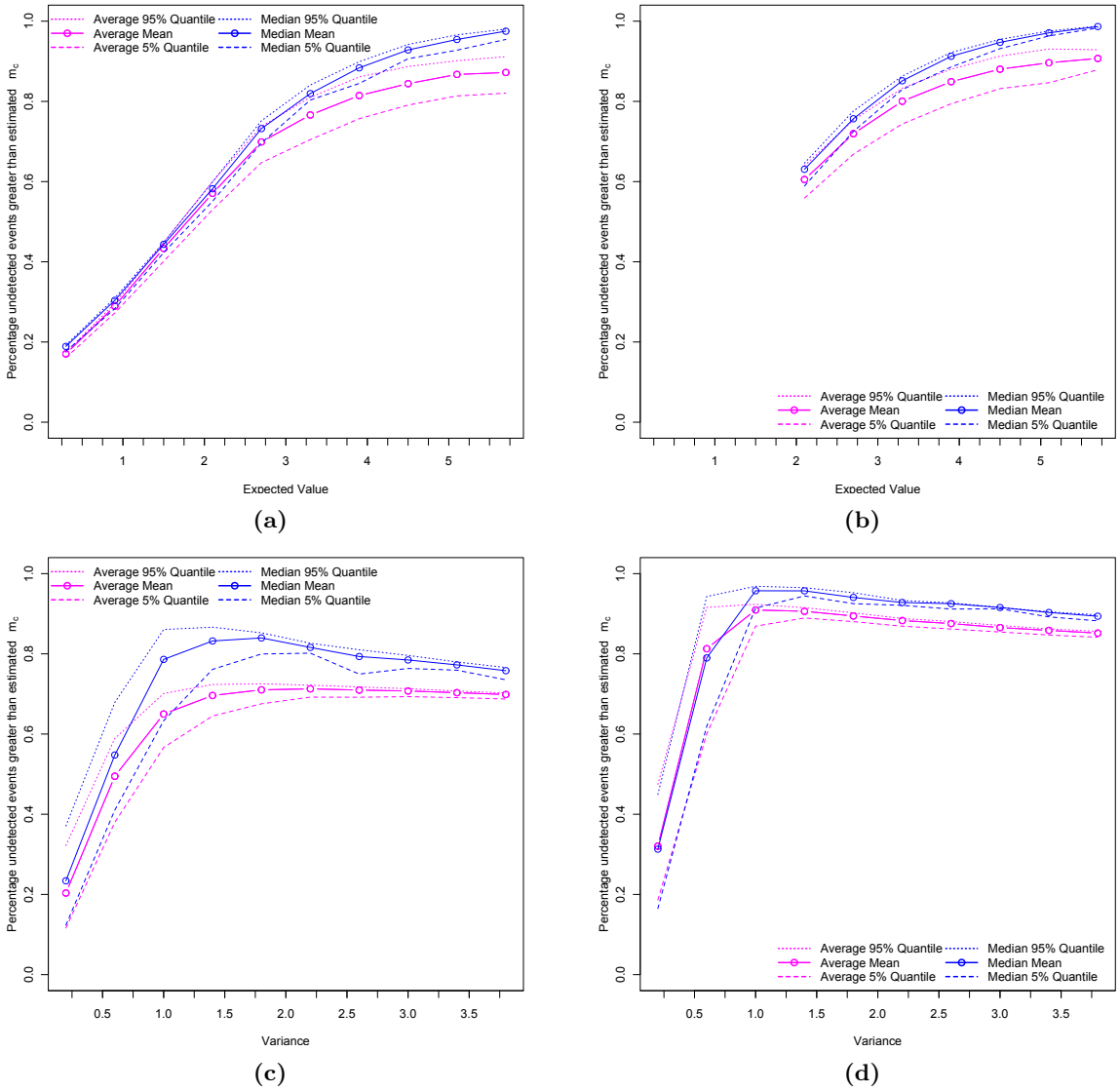


Figure E.34: *Expected number of events not detected as a percentage of expected number of events greater than $\hat{m}_c$ together with error bounds for MAXC estimation method illustrated as a function of expected value (figs. a and b) and variance (figs. c and d) of distribution used to model detection probability. Detection probability in Figures a and c modelled by a cumulative Normal distribution, while modelled by a cumulative logistic distribution in Figures b and d.*

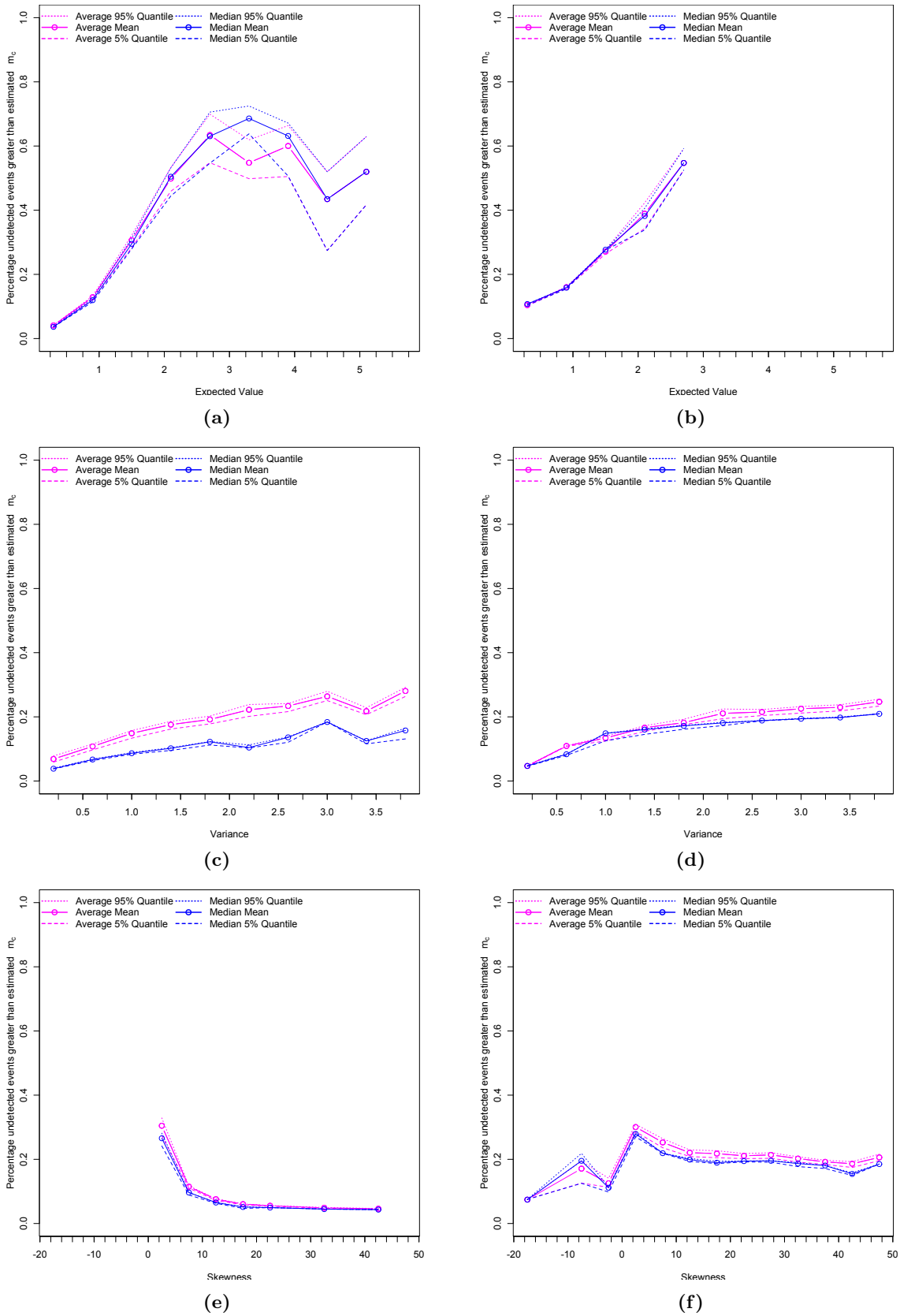


Figure E.35: *Expected number of events not detected as a percentage of expected number of events greater than $\hat{m}_c$ together with error bounds for MAXC estimation method illustrated as a function of expected value (figs. a and b), variance (figs. c and d) and skewness (figs. e and f) of distribution used to model detection probability. Detection probability in Figures a, c and e modelled by a cumulative Log-Normal distribution, while modelled by a cumulative Pareto type II distribution in Figures b, d and f.*

m_c estimation by b-Value stability

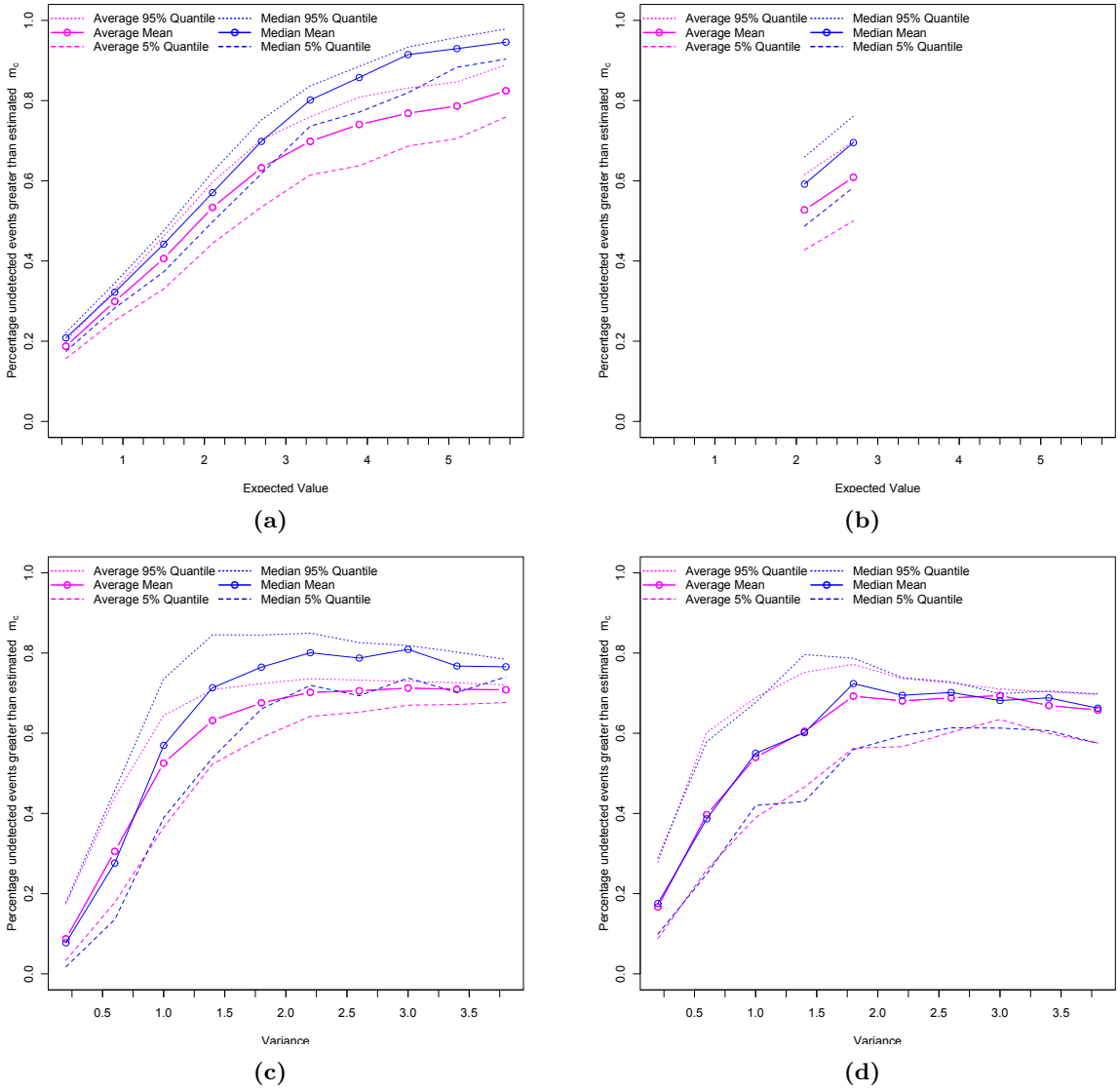


Figure E.36: *Expected number of events not detected as a percentage of expected number of events greater than $\hat{m}_c$ together with error bounds for bVS estimation method illustrated as a function of expected value (figs. a and b) and variance (figs. c and d) of distribution used to model detection probability. Detection probability in Figures a and c modelled by a cumulative Normal distribution, while modelled by a cumulative logistic distribution in Figures b and d.*

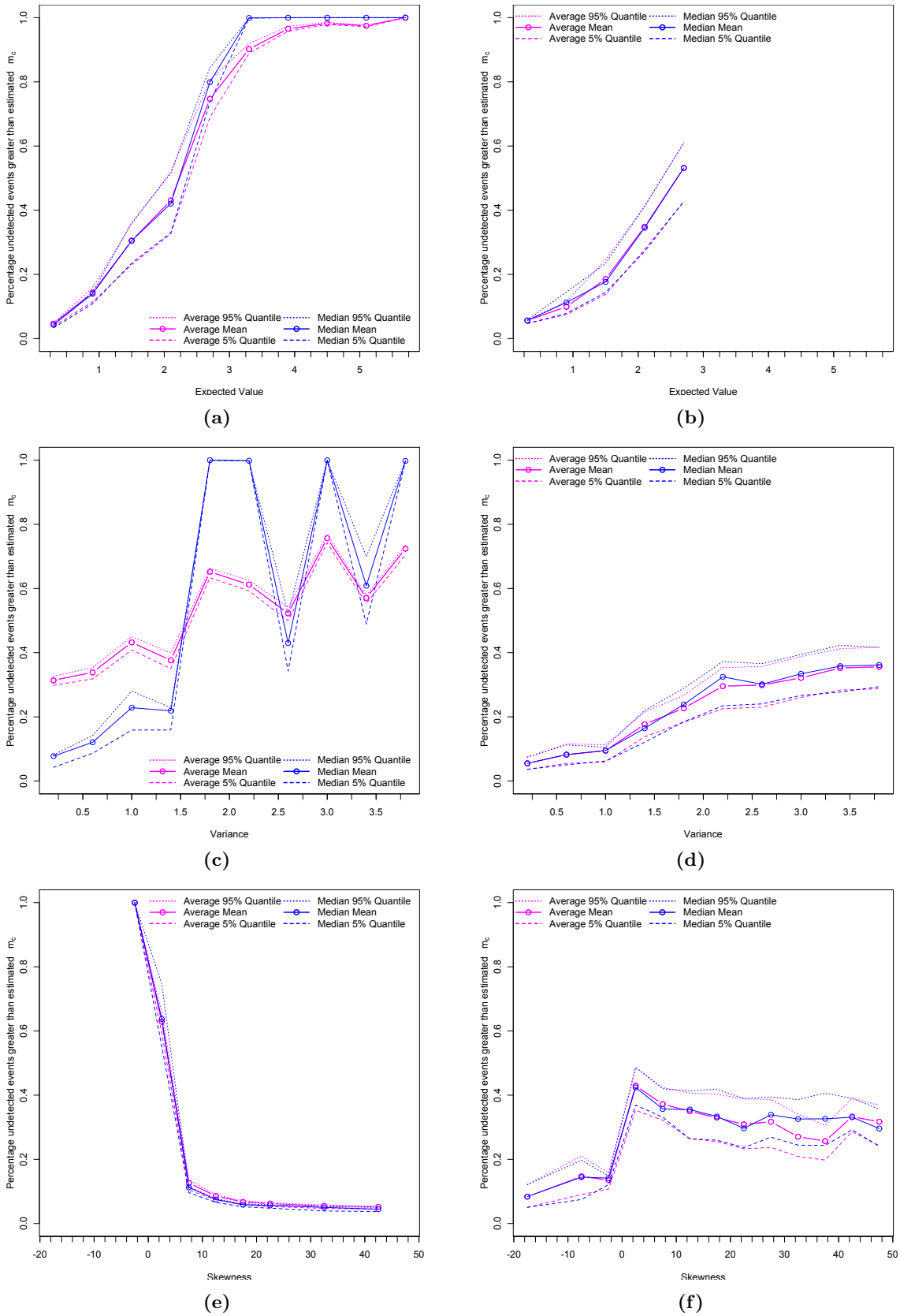


Figure E.37: *Expected number of events not detected as a percentage of expected number of events greater than $\hat{m}_c$ together with error bounds for bVS estimation method illustrated as a function of expected value (figs. a and b), variance (figs. c and d) and skewness (figs. e and f) of distribution used to model detection probability. Detection probability in Figures a, c and e modelled by a cumulative Log-Normal distribution, while modelled by a cumulative Pareto type II distribution in Figures b, d and f.*

Entire magnitude range estimation method

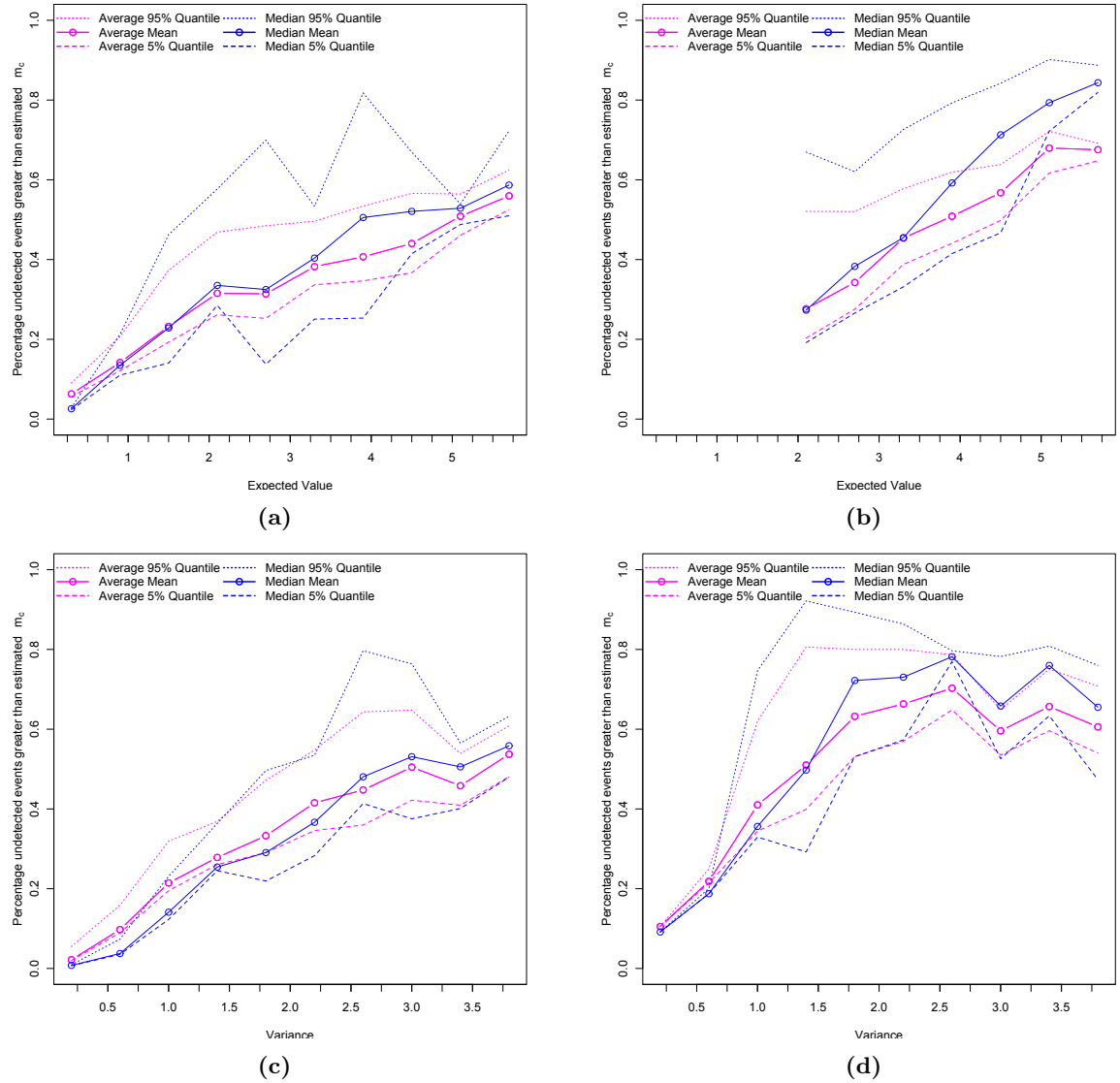


Figure E.38: *Expected number of events not detected as a percentage of expected number of events greater than $\hat{m}_c$ together with error bounds for EMR estimation method illustrated as a function of expected value (figs. a and b) and variance (figs. c and d) of distribution used to model detection probability. Detection probability in Figures a and c modelled by a cumulative Normal distribution, while modelled by a cumulative logistic distribution in Figures b and d.*

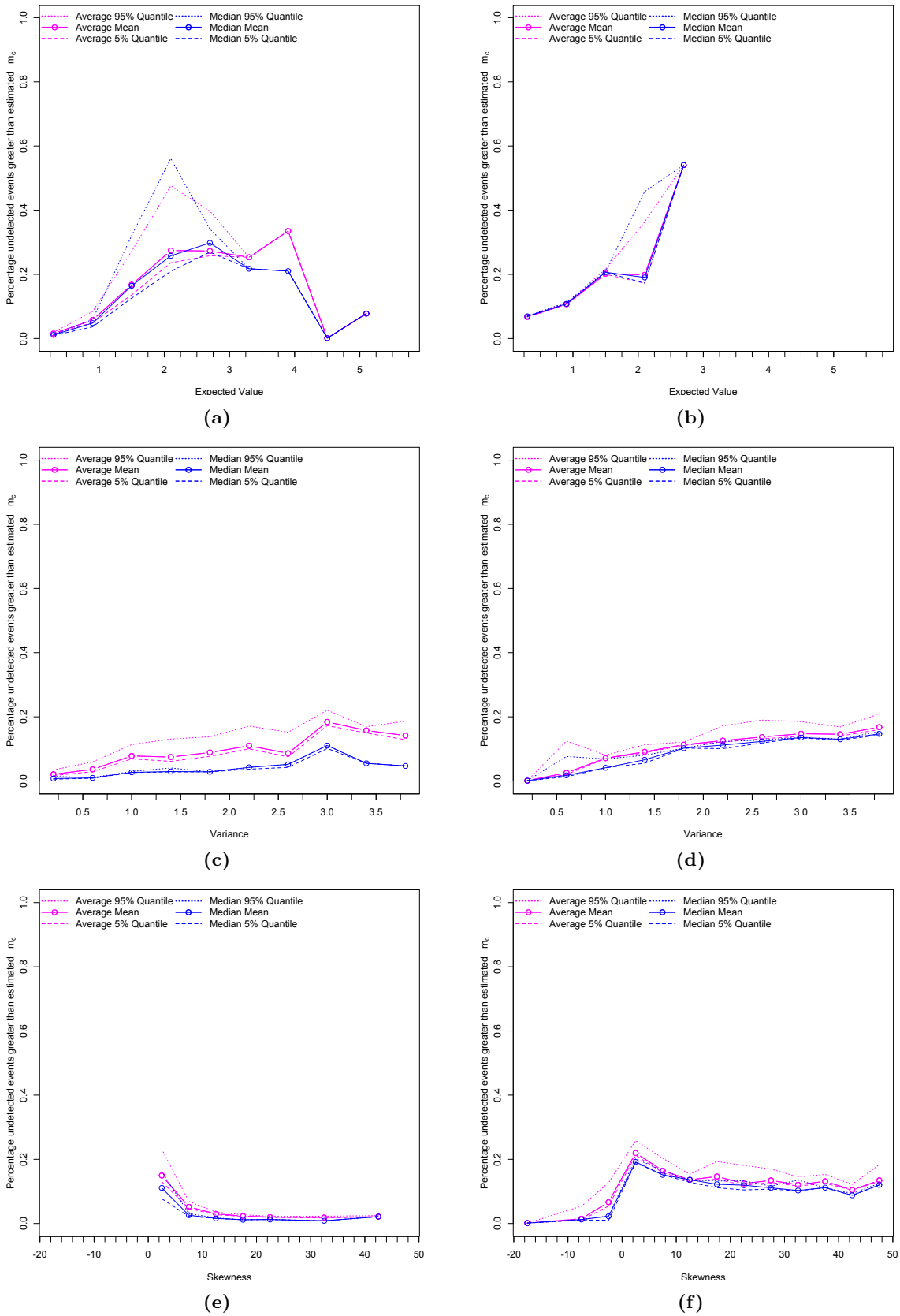


Figure E.39: *Expected number of events not detected as a percentage of expected number of events greater than $\hat{m}_c$ together with error bounds for EMR estimation method illustrated as a function of expected value (figs. a and b), variance (figs. c and d) and skewness (figs. e and f) of distribution used to model detection probability. Detection probability in Figures a, c and e modelled by a cumulative Log-Normal distribution, while modelled by a cumulative Pareto type II distribution in Figures b, d and f.*

Median based assessment of the segment-slope

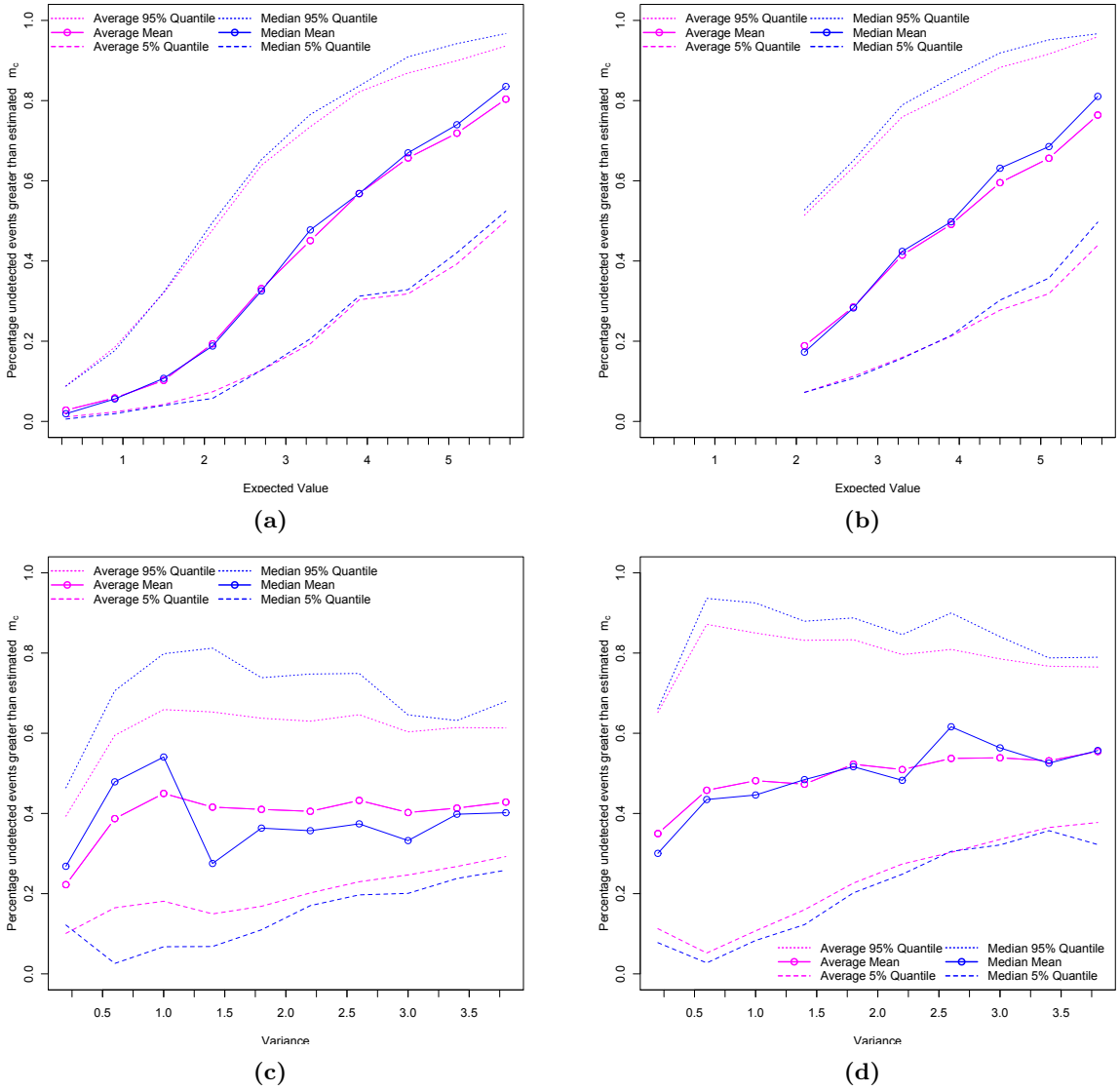


Figure E.40: *Expected number of events not detected as a percentage of expected number of events greater than $\hat{m}_c$ together with error bounds for MBASS estimation method illustrated as a function of expected value (figs. a and b) and variance (figs. c and d) of distribution used to model detection probability. Detection probability in Figures a and c modelled by a cumulative Normal distribution, while modelled by a cumulative logistic distribution in Figures b and d.*

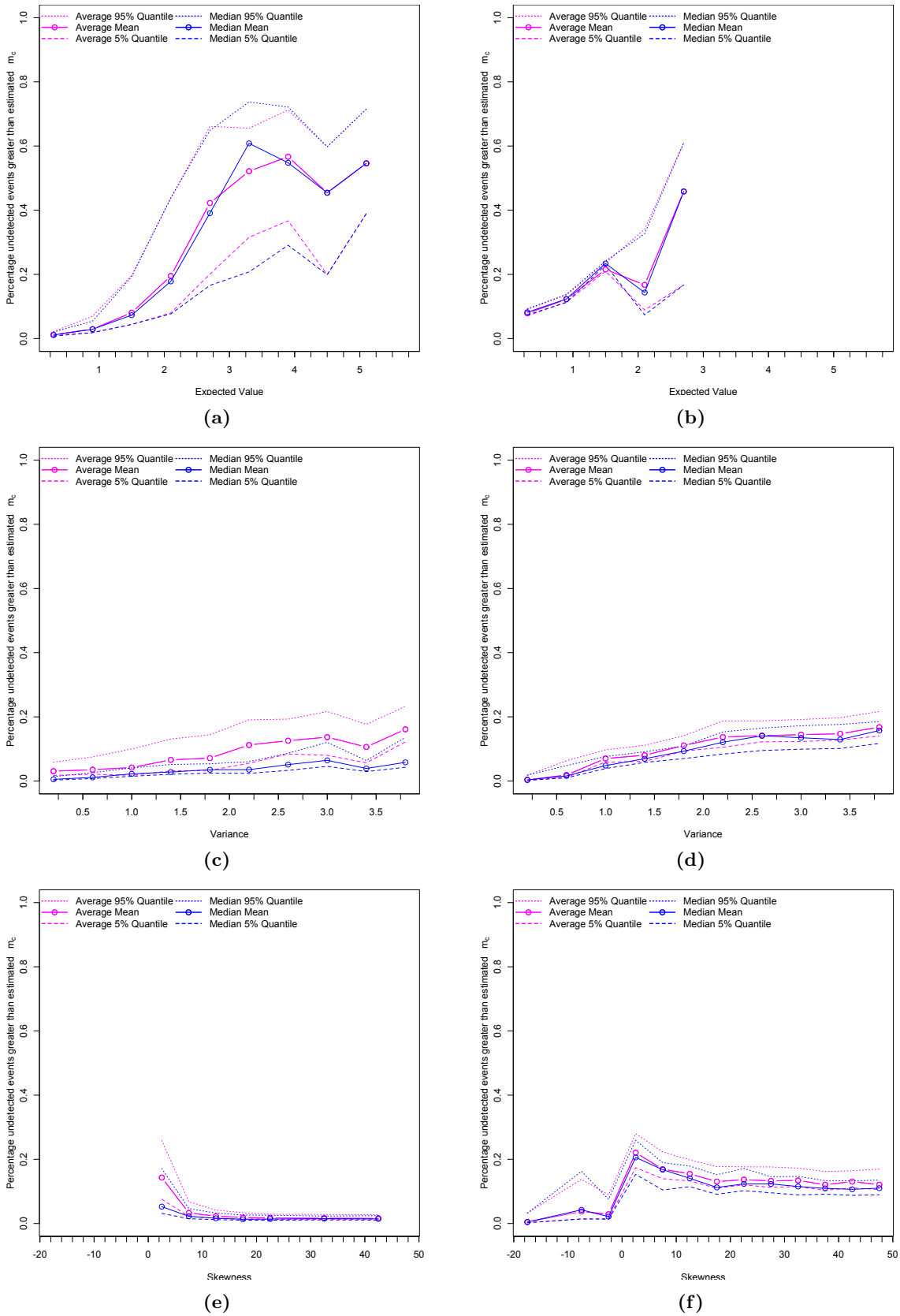


Figure E.41: *Expected number of events not detected as a percentage of expected number of events greater than $\hat{m}_c$ together with error bounds for MBASS estimation method illustrated as a function of expected value (figs. a and b), variance (figs. c and d) and skewness (figs. e and f) of distribution used to model detection probability. Detection probability in Figures a, c and e modelled by a cumulative Log-Normal distribution, while modelled by a cumulative Pareto type II distribution in Figures b, d and f.*

Moment incorporating threshold computation

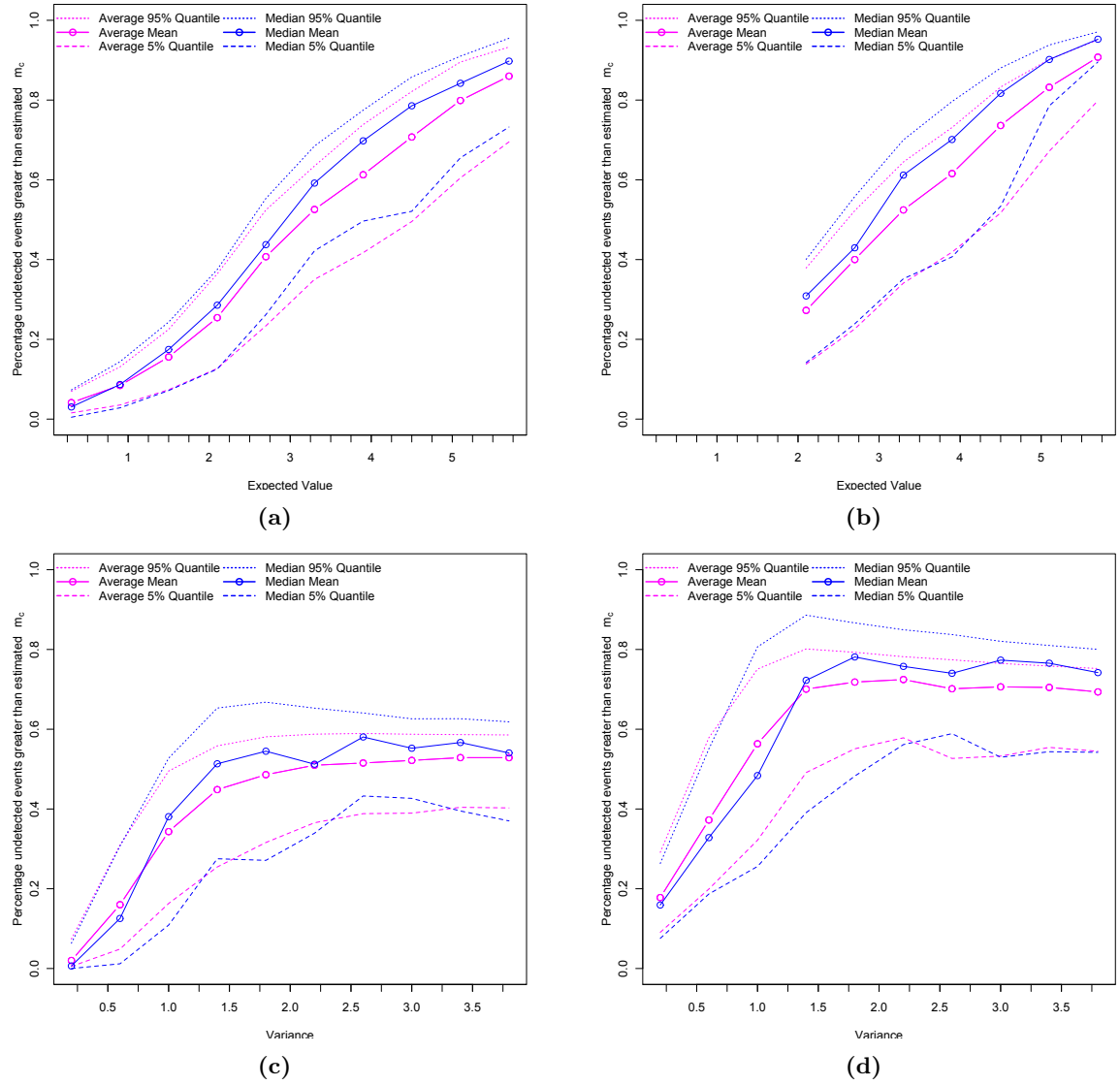


Figure E.42: *Expected number of events not detected as a percentage of expected number of events greater than $\hat{m}_c$ together with error bounds for MITC estimation method illustrated as a function of expected value (figs. a and b) and variance (figs. c and d) of distribution used to model detection probability. Detection probability in Figures a and c modelled by a cumulative Normal distribution, while modelled by a cumulative logistic distribution in Figures b and d.*

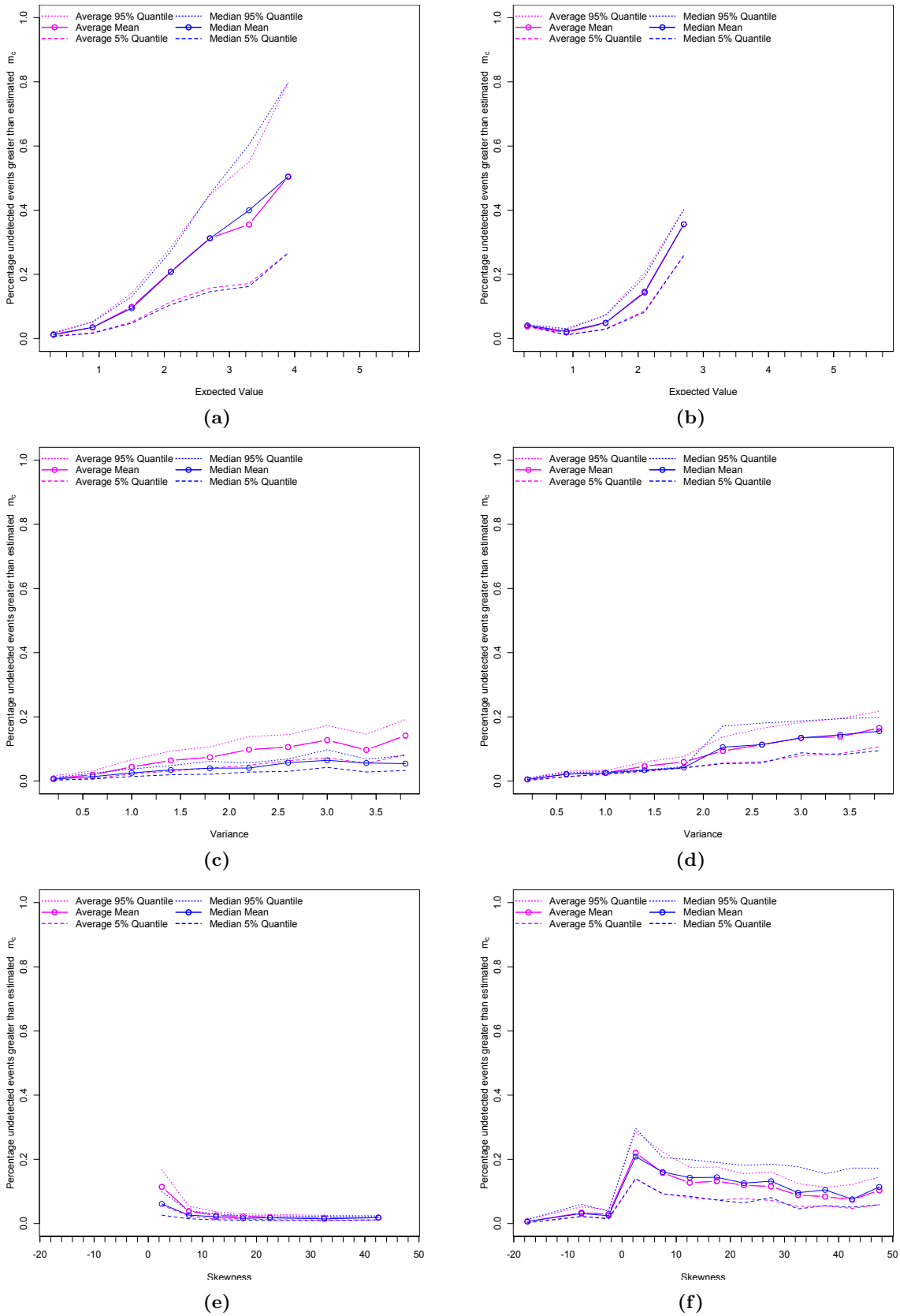


Figure E.43: *Expected number of events not detected as a percentage of expected number of events greater than $\hat{m}_c$ together with error bounds for MITC estimation method illustrated as a function of expected value (figs. a and b), variance (figs. c and d) and skewness (figs. e and f) of distribution used to model detection probability. Detection probability in Figures a, c and e modelled by a cumulative Log-Normal distribution, while modelled by a cumulative Pareto type II distribution in Figures b, d and f.*

E.1.2.2 Four Figure summary of threshold estimation methods

In this section a brief overview is given pertaining to the four-figure summary of the threshold estimation methods when being subjected to a soft detection threshold.

Since no explicit threshold is defined when constructing the distribution a different metric must be chosen when evaluating the efficiency of threshold estimation methods. The specified measure is the percentage of non-detected events greater than the estimated threshold.

This percentage can be seen to vary for different estimation methods as as well as for differing values of the 1) Expected Value 2) Variance and 3) Skewness (if applicable) of the random variable whose cumulative distribution function is being used to model the detection probability.

It is due to this variation that the following three figures (based on the respective lower order moments as stated above) are included in the four-figure summary :

1. Value of the lower order moment at which 25% of events are non-detected.
2. Value of the lower order moment at which 50% of events are non-detected.
3. Value of the lower order moment at which 75% of events are non-detected.

In situations where the exact magnitude of these non-detection percentages is not available, linear interpolation can be used to approximate these figures. In this study data has been grouped by intervals relating to the lower order moments of the random variable whose distribution function is used to model the detection probability and therefore linear interpolation has been used to approximate these exact percentages.

Knowledge of the above three values can be indicative of the following:

1. What characteristics the detected magnitude distribution must exhibit (based on the lower order moments) for a reasonable amount (25%), half of the events (50%) and an extreme amount (75%) of events to be non-detected.
2. The rate at which the percentage of non-detected events increases.

The final figure to be included in the summary is the maximum percentage of events not detected over the entire range considered. This allows for the consideration of a “worst case scenario” over the considered range. This figure also holds key information in the event that the estimation method does not attain at least one of the event non-detection figures. This figure is also useful to consider maximal loss when the rate of event non-detection increases rapidly and continues above the 75% mark.

In this investigation the four-figure summary (Table E.11) is provided for 1) the various threshold estimation methods, 2) differing forms of the detection probability and 3) for differing values of the lower order moments of the random variable attributed to each respective detection probability.

Grouping of event detection distribution by ...		Threshold estimation method	Detection probability modelled by ...															
			Cum. Normal Dist.				Cum. Logistic Dist.				Cum. Log-Normal Dist.				Cum. Pareto type II Dist.			
			% of events lost			Max	% of events lost			Max	% of events lost			Max	% of events lost			Max
			0.25	0.5	0.75	% lost	0.25	0.5	0.75	% lost	0.25	0.5	0.75	% lost	0.25	0.5	0.75	% lost
Grouping of event detection distribution by ...	Expected Value	GOF	0.77	2.33	-	0.68	-	2.36	5.04	0.79	2.29	-	-	0.42	1.4	-	-	0.47
		bVS	0.63	1.94	4.1	0.82	-	-	-	0.6	1.29	2.23	2.71	0.99	1.73	2.59	-	0.53
		MAXC	0.7	1.79	3.15	0.87	-	-	2.92	0.9	1.3	2.3	-	0.63	1.38	2.52	-	0.54
		EMR	1.75	5.02	-	0.55	-	3.8	-	0.67	3.23	-	-	0.33	1.66	2.55	-	0.54
		MBASS	2.34	3.55	5.32	0.8	2.48	3.94	5.62	0.76	2.24	3.68	-	0.56	1.67	-	-	0.45
		MITC	2.07	3.16	4.77	0.85	-	3.18	4.58	0.9	2.33	3.88	-	0.5	2.39	-	-	0.35
	Variance	GOF	0.7	1.41	-	0.67	0.52	0.96	1.65	0.85	-	-	-	0.24	-	-	-	0.22
		bVS	0.49	0.95	-	0.71	0.34	0.88	-	0.69	-	2.19	3.16	0.75	1.93	-	-	0.35
		MAXC	0.26	0.61	-	0.71	-	0.34	0.54	0.9	2.81	-	-	0.28	-	-	-	0.24
		EMR	1.22	3.04	-	0.53	0.66	1.35	-	0.7	-	-	-	0.18	-	-	-	0.16
		MBASS	0.26	-	-	0.44	-	1.79	-	0.55	-	-	-	0.16	-	-	-	0.16
		MITC	0.79	2.03	-	0.52	0.34	0.86	-	0.72	-	-	-	0.14	-	-	-	0.16
	Skewness	GOF	NA				NA				-	-1.81	-	0.55	-	-1.81	-	0.55
		bVS									-	-1.81	-	0.55	-	-1.81	-	0.55
		MAXC									-	-1.81	-	0.55	-	-1.81	-	0.55
		EMR									-	-1.81	-	0.55	-	-1.81	-	0.55
		MBASS									-	-1.81	-	0.55	-	-1.81	-	0.55
		MITC									-	-1.81	-	0.55	-	-1.81	-	0.55

Table E.11: *Four-figure summary relating to percentage of events lost as a result of each threshold estimation method. Displayed figures are based on results grouped by 1) Estimation Method, 2) detection distribution and 3) the lower order moments (Expected value, Variance and Skewness) of RV attributed to detection probability.*

Appendix F :
Maximum likelihood parameter estimation of detected magnitude distribution

In this appendix the maximum likelihood equations for the detected magnitude distribution (f_{M_D}) is derived. This is done for the case where a sharp detection threshold is found in the distribution and the detection probability is modelled by the cumulative Normal distribution.

As seen in (8.25), on p. 29, the detected magnitude distribution takes on the following form :

$$f_{M_D}(m) = C_{Norm} \begin{cases} f_{M_G}(m)\Phi^T(z_m) & \text{if } m_{min} \leq m < m_c \\ f_{M_G}(m) & \text{if } m_c < m \leq m_{max} \\ 0 & \text{otherwise} \end{cases} \tag{F.79}$$

where $z_m = \frac{m-\mu}{\sigma}$ and $\Phi^T(z_m) = \frac{\Phi(z_m)}{\Phi(z_{m_c})}$. Furthermore $\underline{\boldsymbol{m}}$ represents the data vector $(m_1, m_2, \ldots, m_n)$ of event magnitudes related to the n observations. From here the likelihood and subsequent log-likelihood can be found to be :

$$\begin{aligned} L(\beta, \mu, \sigma|\underline{\boldsymbol{m}}) &= \prod_{i=1}^{n} f_{M_D}(m) = \left(\prod_{m_i < m_c} f_{M_D}(m_i) \right) \left(\prod_{m_i \geq m_c} f_{M_D}(m_i) \right) \\ l(\beta, \mu, \sigma|\underline{\boldsymbol{m}}) &= \sum_{m_i < m_c} \ln f_{M_D}(m_i) + \sum_{m_i \geq m_c} \ln f_{M_D}(m_i) \\ &= \sum_{m_i < m_c} \ln f_{M_G}(m_i) \frac{\Phi(z_m)}{\Phi(z_{m_c})} + \sum_{m_i \geq m_c} \ln f_{M_G}(m_i) + n \ln C_{Norm} \\ &= \sum_{m_i < m_c} (\ln \Phi(z_m) - \ln \Phi(z_{m_c})) + \sum_{i=1}^{n} \ln f_{M_G}(m_i) + n \ln C_{Norm} \end{aligned} \tag{F.80}$$

The following general result holds where x is one of the parameters β, μ or σ of the likelihood function

$$\begin{aligned} \frac{\partial}{\partial x} \ln C_{Norm} &= \frac{\frac{\partial}{\partial x} C_{Norm}}{C_{Norm}} \\ &= (C_{Norm})^{-1} \frac{\partial}{\partial x} \left(\int_{m_{min}}^{m_c} f_{M_G}(m)\Phi^T(z_m)\, dm + \int_{m_c}^{m_{max}} f_{M_G}(m)\, dm \right)^{-1} \\ &= (C_{Norm})^{-1}(-1)(C_{Norm})^2 \frac{\partial}{\partial x} \left(\int_{m_{min}}^{m_c} f_{M_G}(m)\Phi^T(z_m)\, dm + \int_{m_c}^{m_{max}} f_{M_G}(m)\, dm \right) \\ &= (-1)(C_{Norm}) \left(\int_{m_{min}}^{m_c} \frac{\partial}{\partial x} f_{M_G}(m)\Phi^T(z_m)\, dm + \int_{m_c}^{m_{max}} \frac{\partial}{\partial x} f_{M_G}(m)\, dm \right) \end{aligned}$$

By equation 0.410 from the Table of Inegrals, Series and Products [20].

(F.81)

Three cases are now individually considered.

Case 1 : $\boldsymbol{x} = \boldsymbol{\beta}$

From (9.2) the following can be seen, together with the subsequent restatement

$$\frac{\partial}{\partial\beta}\ln f_{M_G}(m) = \frac{1}{\beta} - m + \frac{m_{min}e^{-\beta m_{min}} - m_{max}e^{-\beta m_{max}}}{e^{-\beta m_{min}} - e^{-\beta m_{max}}} \tag{F.82}$$

$$\frac{\partial}{\partial\beta}f_{M_G}(m) = f_{M_G}(m)\left(E[M_G] - m\right) \tag{F.83}$$

Therefore,

$$\begin{aligned}
&\frac{\partial}{\partial\beta}\ln C_{Norm} \\
&= (-1)(C_{Norm})\left(\int_{m_{min}}^{m_c}\frac{\partial}{\partial\beta}f_{M_G}(m)\Phi^T(z_m)\,dm + \int_{m_c}^{m_{max}}\frac{\partial}{\partial\beta}f_{M_G}(m)\,dm\right) \\
&= (-1)(C_{Norm})(E[M]\cdot C_{Norm}^{-1} - \left(\int_{m_{min}}^{m_c} m f_{M_G}(m)\Phi^T(z_m)\,dm + \int_{m_c}^{m_{max}} m f_{M_G}(m)\,dm\right) \\
&= (-1)E[M] + C_{Norm}^{-1}\Bigg(- \frac{e^{-\beta m}}{e^{-\beta m_{min}} - e^{-\beta m_{max}}}\left(m + \frac{1}{\beta}\right)\Phi^T(z_m)\Big|_{m_{min}}^{m_c} \\
&\quad - \frac{1}{\Phi(z_{m_c})}\left(\int_{m_{min}}^{m_c}\left(-f_{M_G}(m)\frac{1}{\beta}\left(m + \frac{1}{\beta}\right)\right)\frac{1}{\sigma}\phi(z_m)\,dm + \int_{m_c}^{m_{max}} m f_{M_G}(m)\,dm\right)\Bigg)
\end{aligned}$$

Integration by parts and utelizing (B.31). (F.84)

When considering the first of the two integrals from above, the following simplification can be made

$$\begin{aligned}
&-\frac{1}{\beta}\int_{m_{min}}^{m_c} f_{M_G}(m)\left(m + \frac{1}{\beta}\right)\frac{1}{\sigma}\phi(z_m)\,dm \\
&= -\frac{1}{\beta}\frac{\beta e^{-\beta\left(\mu - \frac{1}{2}\sigma^2\beta\right)}}{e^{-\beta m_{min}} - e^{-\beta m_{max}}}\left(\int_{m_{min}}^{m_c} m\frac{1}{\sigma}\phi(z_m^*)\,dm + \frac{1}{\beta}\int_{m_{min}}^{m_c}\frac{1}{\sigma}\phi(z_m^*)\,dm\right)
\end{aligned}$$

By combining the expressions as in (B.33)

$$= -\frac{e^{-\beta\left(\mu - \frac{1}{2}\sigma^2\beta\right)}}{e^{-\beta m_{min}} - e^{-\beta m_{max}}}\left(\left(\mu - \sigma^2\beta + \frac{1}{\beta}\right)\left(\Phi(z_{m_c}^*) - \Phi(z_{m_{min}}^*)\right) - \sigma\left(\phi(z_{m_c}^*) - \phi(z_{m_{min}}^*)\right)\right)$$

By incorporating (B.36) and thereafter simplifying. (F.85)

The second integral in (F.84), is of the form represented by (B.30). Therefore, the partial derivative of the normalizing constant can be found to be

$$\begin{aligned}
&\frac{\partial}{\partial\beta}\ln C_{Norm} \\
&\quad = (-1)E[M_G] \\
&\qquad + \frac{1}{C_{Norm}\cdot\left(e^{-\beta m_{min}} - e^{-\beta m_{max}}\right)}\Bigg(- \left(e^{-\beta m_c}\left(m_c + \frac{1}{\beta}\right)\Phi^T(z_{m_c}) - e^{-\beta m_{min}}\left(m_{min} + \frac{1}{\beta}\right)\Phi^T(z_{m_{min}})\right) \\
&\qquad\quad - \frac{1}{\Phi(z_{m_c})}\Bigg(- e^{-\beta\left(\mu - \frac{1}{2}\sigma^2\beta\right)}\left(\left(\mu - \sigma^2\beta + \frac{1}{\beta}\right)\left(\Phi(z^*_{m_c}) - \Phi(z^*_{m_{min}})\right) - \sigma\left(\phi(z^*_{m_c}) - \phi(z^*_{m_{min}})\right)\right) \\
&\qquad\qquad + \frac{1}{\beta}\left(e^{-\beta m_{min}} - e^{-\beta m_c}\right) + m_{min}e^{-\beta m_{min}} - m_c e^{-\beta m_c}\Bigg)\Bigg)
\end{aligned} \tag{F.86}$$

Before considering the remaining two cases ($x = \mu$ and $x = \sigma$), it is worth noting that the following can be derived by using the chain rule of differentiation :

$$\frac{\partial}{\partial\mu}\Phi(z_m) = \phi(z_m)\ \frac{\partial}{\partial\mu}\left(\frac{m-\mu}{\sigma}\right) = -\frac{1}{\sigma}\ \phi(z_m) \tag{F.87}$$

$$\frac{\partial}{\partial\sigma}\Phi(z_m) = \phi(z_m)\ \frac{\partial}{\partial\sigma}\left(\frac{m-\mu}{\sigma}\right) = -\frac{m-\mu}{\sigma^2}\ \phi(z_m) \tag{F.88}$$

Case 2 : $\boldsymbol{x = \mu}$

$$\begin{aligned}
\frac{\partial}{\partial\mu}\ln C_{Norm} &= (-1)C_{Norm}\left(\int_{m_{min}}^{m_c}\frac{\partial}{\partial\mu}f_{M_G}(m)\Phi^T(z_m)\,dm + \int_{m_c}^{m_{max}} f_{M_G}(m)\,dm\right)\\
&= (-1)C_{Norm}\int_{m_{min}}^{m_c} f_{M_G}(m)\frac{\partial}{\partial\mu}\frac{\Phi(z_m)}{\Phi(z_{m_c})}\,dm\\
&= (-1)C_{Norm}\int_{m_{min}}^{m_c} f_{M_G}(m)\frac{\left(\frac{\partial}{\partial\mu}\Phi(z_m)\right)\Phi(z_{m_c}) - \Phi(z_m)\frac{\partial}{\partial\mu}\Phi(z_{m_c})}{(\Phi(z_{m_c}))^2}\,dm\\
&= -\frac{C_{Norm}}{(\Phi(z_{m_c}))^2}\left(\Phi(z_{m_c})\int_{m_{min}}^{m_c} f_{M_G}(m)\left(\frac{\partial}{\partial\mu}\Phi(z_m)\right)dm - \left(\frac{\partial}{\partial\mu}\Phi(z_{m_c})\right)\int_{m_{min}}^{m_c} f_{M_G}(m)\Phi(z_m)\,dm\right)
\end{aligned} \tag{F.89}$$

$$= -\frac{C_{Norm}}{(\Phi(z_{m_c}))^2}\left(\Phi(z_{m_c})\int_{m_{min}}^{m_c} f_{M_G}(m)\left(-\frac{1}{\sigma}\phi(z_m)\right)dm - \left(-\frac{1}{\sigma}\phi(z_{m_c})\right)\int_{m_{min}}^{m_c} f_{M_G}(m)\Phi(z_m)\,dm\right)$$

By (F.87).

$$= \frac{C_{Norm}}{(\Phi(z_{m_c}))^2}\left(\Phi(z_{m_c})\int_{m_{min}}^{m_c} f_{M_G}(m)\frac{1}{\sigma}\phi(z_m)\,dm - \frac{1}{\sigma}\phi(z_{m_c})\int_{m_{min}}^{m_c} f_{M_G}(m)\Phi(z_m)\,dm\right)$$

$$\begin{aligned}
= \frac{C_{Norm}}{(\Phi(z_{m_c}))^2}\Bigg(&\Phi(z_{m_c})\frac{\beta e^{-\beta\left(\mu-\frac{1}{2}\sigma^2\beta\right)}}{e^{-\beta m_{min}} - e^{-\beta m_{max}}}\left(\Phi(z^*_{m_c}) - \Phi(z^*_{m_{min}})\right)\\
&-\frac{1}{\sigma}\phi(z_{m_c})\Bigg(F_{M_G}(m_c)\Phi(z_{m_c}) - \frac{1}{1-\exp\left(-\beta(m_{max}-m_{min})\right)}\Big(\Phi(z_{m_c}) - \Phi(z_{m_{min}})\\
&-e^{-\beta\left(\mu-m_{min}-\frac{1}{2}\sigma^2\beta\right)}\left(\Phi(z^*_{m_c}) - \Phi(z^*_{m_{min}})\right)\Big)\Bigg)\Bigg)
\end{aligned}$$

By (B.33) and (B.35), where $z^*_m = \dfrac{m-(\mu-\sigma^2\beta)}{\sigma}$ and since $F_{M_G}(m_{min}) = 0$

Case 3 : $\boldsymbol{x = \sigma}$

$$\frac{\partial}{\partial\sigma}\ln C_{Norm} = -\frac{C_{Norm}}{(\Phi(z_{m_c}))^2}\left(\Phi(z_{m_c})\int_{m_{min}}^{m_c} f_{M_G}(m)\left(\frac{\partial}{\partial\sigma}\Phi(z_m)\right)dm - \left(\frac{\partial}{\partial\sigma}\Phi(z_{m_c})\right)\int_{m_{min}}^{m_c} f_{M_G}(m)\Phi(z_m)\,dm\right)$$

Simmilarly as for (F.89).

$$\begin{aligned}
= -\frac{C_{Norm}}{(\Phi(z_{m_c}))^2}\Bigg(&\Phi(z_{m_c})\int_{m_{min}}^{m_c} f_{M_G}(m)\left(-\frac{m-\mu}{\sigma^2}\phi(z_m)\right)dm\\
&-\left(-\frac{m_c-\mu}{\sigma^2}\phi(z_{m_c})\right)\int_{m_{min}}^{m_c} f_{M_G}(m)\Phi(z_m)\,dm\Bigg)
\end{aligned}$$

By (F.88).

(F.90)

when considering the first integral in isolation the following simplifications can be made

$$\int_{m_{min}}^{m_c} f_{M_G}(m)\left(-\frac{m-\mu}{\sigma^2}\phi(z_m)\right)dm$$
$$= -\frac{1}{\sigma}\int_{m_{min}}^{m_c} f_{M_G}(m)(m-\mu)\frac{1}{\sigma}\phi(z_m)\,dm$$
$$= -\frac{1}{\sigma}\frac{\beta e^{-\beta\left(\mu-\frac{1}{2}\sigma^2\beta\right)}}{e^{-\beta m_{min}}-e^{-\beta m_{max}}}\int_{m_{min}}^{m_c}(m-\mu)\frac{1}{\sigma}\phi(z_m^*)\,dm$$

By combining the expressions $f_{M_G}(m)$ and $\phi(m)$ as in (B.32) where $z_m^* = \frac{m-(\mu-\sigma^2\beta)}{\sigma}$.

$$= \frac{\beta e^{-\beta\left(\mu-\frac{1}{2}\sigma^2\beta\right)}}{e^{-\beta m_{min}}-e^{-\beta m_{max}}}\left(\sigma\beta(\Phi(z_{m_c}^*)-\Phi(z_{m_{min}}^*))+\phi(z_{m_c}^*)-\phi(z_{m_{min}}^*)\right)$$

By (B.36) and thereafter rearranging. (F.91)

The second integral can be simplified according to (B.35). Therefore it holds that

$$\frac{\partial}{\partial\sigma}\ln C_{Norm} = -\frac{C_{Norm}}{(\Phi(z_{m_c}))^2}\Bigg(\Phi(z_{m_c})\frac{\beta e^{-\beta\left(\mu-\frac{1}{2}\sigma^2\beta\right)}}{e^{-\beta m_{min}}-e^{-\beta m_{max}}}(\sigma\beta(\Phi(z_{m_c}^*)-\Phi(z_{m_{min}}^*))+\phi(z_{m_c}^*)-\phi(z_{m_{min}}^*))$$
$$+\frac{m_c-\mu}{\sigma^2}\phi(z_{m_c})\Bigg(F_{M_G}(m_c)\Phi(z_{m_c})-\frac{1}{1-e^{-\beta(m_{max}-m_{min})}}\Big(\Phi(z_{m_c})-\Phi(z_{m_{min}})-$$
$$e^{-\beta\left(\mu-m_{min}-\frac{1}{2}\sigma^2\beta\right)}\left(\Phi\left(z_{m_c}^*\right)-\Phi\left(z_{m_{min}}^*\right)\right)\Big)\Bigg)\Bigg) \tag{F.92}$$

Therefore, from the log-likelihood function in (F.80) and utilizing the above results, the partial derivatives of the function are as follows

$$\frac{\partial}{\partial\beta}l(\beta,\mu,\sigma|\underline{\boldsymbol{m}}) = \sum_{i=1}^{n}\frac{\partial}{\partial\beta}\ln f_{M_G}(m_i)+n\frac{\partial}{\partial\beta}\ln C_{Norm}$$
$$= n\left(\frac{1}{\beta}+\frac{m_{min}e^{-\beta m_{min}}-m_{max}e^{-\beta m_{max}}}{e^{-\beta m_{min}}-e^{-\beta m_{max}}}\right)-\sum_{i=1}^{n}m_i+n\Bigg((-1)\left(\frac{1}{\beta}+\frac{m_{min}e^{-\beta m_{min}}-m_{max}e^{-\beta m_{max}}}{e^{-\beta m_{min}}-e^{-\beta m_{max}}}\right)$$
$$+\frac{1}{C_{Norm}\cdot(e^{-\beta m_{min}}-e^{-\beta m_{max}})}\Bigg(-\left(e^{-\beta m_c}\left(m_c+\frac{1}{\beta}\right)\Phi^T(z_{m_c})-e^{-\beta m_{min}}\left(m_{min}+\frac{1}{\beta}\right)\Phi^T(z_{m_{min}})\right)$$
$$-\frac{1}{\Phi(z_{m_c})}\Bigg(-e^{-\beta\left(\mu-\frac{1}{2}\sigma^2\beta\right)}\left(\left(\mu-\sigma^2\beta+\frac{1}{\beta}\right)\left(\Phi(z_{m_c}^*)-\Phi(z_{m_{min}}^*)\right)-\sigma\left(\phi(z_{m_c}^*)-\phi(z_{m_{min}}^*)\right)\right)$$
$$+\frac{1}{\beta}\left(e^{-\beta m_{min}}-e^{-\beta m_c}\right)+m_{min}e^{-\beta m_{min}}-m_ce^{-\beta m_c}\Bigg)\Bigg)\Bigg) \tag{F.93}$$

$$\frac{\partial}{\partial\mu}l(\beta,\mu,\sigma|\underline{\boldsymbol{m}}) = \sum_{m_i<m_c}\left(\frac{\frac{\partial}{\partial\mu}\Phi(z_{m_i})}{\Phi(z_{m_i})} - \frac{\frac{\partial}{\partial\mu}\Phi(z_{m_c})}{\Phi(z_{m_c})}\right) + n\frac{\partial}{\partial\mu}\ln C_{Norm}$$

$$= -\frac{1}{\sigma}\sum_{m_i<m_c}\left(\frac{\phi(z_{m_i})}{\Phi(z_{m_i})} - \frac{\phi(z_{m_c})}{\Phi(z_{m_c})}\right)$$

$$+n\frac{C_{Norm}}{(\Phi(z_{m_c}))^2}\Bigg(\Phi(z_{m_c})\frac{\beta e^{-\beta\left(\mu-\frac{1}{2}\sigma^2\beta\right)}}{e^{-\beta m_{min}}-e^{-\beta m_{max}}}\left(\Phi(z^*_{m_c})-\Phi(z^*_{m_{min}})\right)$$

$$-\frac{1}{\sigma}\phi(z_{m_c})\Bigg(F_{M_G}(m_c)\Phi(z_{m_c}) - \frac{1}{1-\exp\left(-\beta(m_{max}-m_{min})\right)}\Big(\Phi(z_{m_c})-\Phi(z_{m_{min}})$$

$$-e^{-\beta\left(\mu-m_{min}-\frac{1}{2}\sigma^2\beta\right)}\left(\Phi(z^*_{m_c})-\Phi(z^*_{m_{min}})\right)\Big)\Bigg)\Bigg) \tag{F.94}$$

$$\frac{\partial}{\partial\sigma}l(\beta,\mu,\sigma|\underline{\boldsymbol{m}}) = \sum_{m_i<m_c}\left(\frac{\frac{\partial}{\partial\sigma}\Phi(z_{m_i})}{\Phi(z_{m_i})} - \frac{\frac{\partial}{\partial\sigma}\Phi(z_{m_c})}{\Phi(z_{m_c})}\right) + n\frac{\partial}{\partial\sigma}\ln C_{Norm}$$

$$= -\frac{1}{\sigma^2}\sum_{m_i<m_c}(m_i-\mu)\left(\frac{\phi(z_{m_i})}{\Phi(z_{m_i})} - \frac{\phi(z_{m_c})}{\Phi(z_{m_c})}\right)$$

$$-n\frac{C_{Norm}}{(\Phi(z_{m_c}))^2}\Bigg(\Phi(z_{m_c})\frac{\beta e^{-\beta\left(\mu-\frac{1}{2}\sigma^2\beta\right)}}{e^{-\beta m_{min}}-e^{-\beta m_{max}}}(\sigma\beta(\Phi(z^*_{m_c})-\Phi(z^*_{m_{min}}))+\phi(z^*_{m_c})-\phi(z^*_{m_{min}}))$$

$$+\frac{m_c-\mu}{\sigma^2}\phi(z_{m_c})\Bigg(F_{M_G}(m_c)\Phi(z_{m_c}) - \frac{1}{1-e^{-\beta(m_{max}-m_{min})}}\Big(\Phi(z_{m_c})-\Phi(z_{m_{min}})-$$

$$e^{-\beta\left(\mu-m_{min}-\frac{1}{2}\sigma^2\beta\right)}\left(\Phi\left(z^*_{m_c}\right)-\Phi\left(z^*_{m_{min}}\right)\right)\Big)\Bigg)\Bigg) \tag{F.95}$$

Appendix G :
List of Acronyms

CPP Change point problem

FMD Frequency-Magnitude Distribution

PDF Probability density function

CDF Cumulative distribution function

bVS b-Value Stability threshold estimation method as described by Cao and Gao [8]

EMR Entire Magnitude Range threshold estimation method as described by Woessner and Wiemer [60]

GOF Goodness of fit threshold estimation method as described by Wiemer and Wyss [59]

MAXC Maximum Curvature threshold estimation method as described by Wiemer and Wyss [59]

MBASS Median based assessment of the segment slope threshold estimation method as described by D. Amorèse [3]

MITC Moment incorporating threshold computation, as developed in this investigation.

MSE Mean-squared error

Appendix H :
List of Programming Aids

The R Project for Statistical Computing [45]

Package "coin" [29], [30]

Package "multcomp" [28]

Package "pgirmess" [19]

Package "ismev" [16]

Package "evd" [54]

Package "Rmpfr" [37]

Appendix I : Programming

I.1.1 Catalogue functions

```
#~~~~~~~~~~~~~~~~~~ Earthquake Catalog Functions & Operations~~~~~~~~~~~~~~~~~~#
#~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~#
#~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~#
################################################################################

#Binning (Incremental or Cumulative)
  EQBinning<-function(dataEQ, binWidth = 0.1, bType = "inc"){
      # bType = "inc" (Incremental) or "cum" (Cumulative)

      #Get Min Mag
          MagMin <- floor(min(dataEQ)/binWidth) * binWidth
      #Get Max Mag
          MagMax <- ceiling(max(dataEQ)/binWidth) * binWidth

      #Construct Binning partition
        bins <- base::matrix(seq(MagMin,MagMax, by = binWidth))
        numBins <- nrow(bins)

      #Binning
        lvDataEQ <- base::matrix(dataEQ)
        RetArray <- base::matrix(,numBins,2)
        RetArray[,1] <- bins

        for(i in 1:numBins){
          RetArray[i,2] <- NROW(lvDataEQ[(abs(lvDataEQ - c(bins[i])) <= binWidth / 2)])
        }

      if(tolower(bType) == "cum"){
        BinnedEQ <- RetArray
        for(i in 1:numBins){
        RetArray[i,2] <- sum(BinnedEQ[i:numBins,2])
        }
      }

      return(RetArray)
  }
```

./RCode/EQFunc_Catalog.r

I.1.2 Distributional functions

```
#Density - Distribution - Quantile - Simulation


#Pure Continuous GR ~~~~~~ mMin <= M <= mMax (can be Inf)
  #(d)Density Function
    dGR<-function(Mag, bValue, mMin = 0, mMax = Inf){
      betaValue <- log(10)*bValue
      ret <- betaValue * exp(-betaValue * Mag) / (exp(-betaValue * mMin) - exp(-betaValue * mMax))
      return(ret)
    }

  #(p)Probability - Cumulative Distribution Function
    pGR<-function(Mag, bValue, mMin = 0, mMax = Inf){
      betaValue <- log(10)*bValue
      ret <- (1 - exp(-betaValue * (Mag - mMin))) / (1 - exp(-betaValue * (mMax - mMin)))
      return(ret)
    }

  #(q)Quantile Function
    qGR<-function(x, bValue, mMin = 0, mMax = Inf){
      betaValue <- log(10)*bValue
      ret <- mMin - (1 / betaValue) * log(1 - x * (1 - exp(-betaValue * (mMax - mMin))))
      return(ret)
    }

  #Simulate (r)Random GR Deviate
    rGR <- function(n, bValue, mMin = 0, mMax = Inf, obsErr = 0){
      U <- runif(n)
```

```
      ret <- qGR(U, bValue, mMin, mMax)
          #Include random observation error
      if(obsErr > 0){ret <- ret + rnorm(n, sd = obsErr)}
      return(ret)
    }
   #Log-Likelihood Function
     likl.GR <- function(mVec, bValue, mMin = 0, mMax = Inf){
      betaValue <- log(10) * bValue
      n <- NROW(mVec)
      sumM <- sum(mVec)

      ret <- (betaValue ^ n) * (exp(-betaValue * sumM)) / (exp(-betaValue * mMin) - exp(-betaValue *
          mMax)) ^ n
      ret <- log(ret)

      return(ret)
     }


#Binned GR Distribution
  #Probability (m)Mass Function of Binned GR
    mBinGR<-function(mCenter, bValue, mMin, mMax = Inf, binWidth =0.1){
    #See Note on summing these bins! sum {Mmin - (binWidth / 2) TO Mmax + (binWidth / 2)} = 1
    M2 <- mCenter + binWidth / 2
    M1 <- mCenter - binWidth / 2

    ret <- pGR(M2, bValue, mMin, mMax) - pGR(M1, bValue, mMin, mMax)
    return(ret)
  }

  #Log-Likelihood Functions
    #Multinomial Formulation
      likl.MultBinGR<-function(lvEQData, bValue, mMin = -1, mMax = -1){
        if(mMin == -1){MagMin <- min(lvEQData[,1])}
        if(mMax == 1){MagMax <- max(lvEQData[,1])}

        numEQs <- sum(lvEQData[,2])

        binWidth <- abs(lvEQData[2,1] - lvEQData[1,1])
        betaValue <- bValue * log(10)

        #Function
          #NOTE : log(N! / product(xi!)) has been left out since factorials introduce large number
              scaling
          #       This term is however constant for all likelihood values and can therefore be left
              out

        tempA <-  betaValue* sum(matrix(lvEQData[,1] * lvEQData[,2]))
        tempB <-  numEQs * log(sinh(betaValue * binWidth / 2))
        tempC <-  numEQs * log(exp(-betaValue * MagMin) - exp(-betaValue * MagMax))

        likl <- -tempA + tempB - tempC
        return(likl)
      }
```

./RCode/EQFunc_DDQS.r

I.1.3 Implementation of b-Value estimation algorithm

```
  #GR Distribution (Continious) - TRuncated on Left & Right - Page MaxLik Relation
    betaValue.mle.PageRelation<-function(lvEQData, epsilon = 0.0000001, maxIt = 10000, mmin = -1,
        mmax = -1){
      #Page (1968) - Data in Perfect World
      #Starting value for beta in Newton-Raphson Scheme will be obtained from the Aki(1965) - Utsu
          (1965) formula.

      f_beta <- function(beta, mbar, mmin, mmax){1 / beta - mbar + mmin - (mmax - mmin) / (exp(beta*(
          mmax - mmin)) - 1)}

      fPrime_beta <- function(beta, mmin, mmax){-1/ (beta^2) + ((mmax - mmin) / (exp(beta*(mmax -
          mmin)) - 1))^2 * exp(beta*(mmax - mmin))}

      mbar = mean(lvEQData)
      if(mmin == -1){mmin = min(lvEQData)}
      if(mmax == -1){mmax = max(lvEQData)}

      if(mbar == mmin){return(NaN)}

      betaStar = 1 / (mbar - mmin)     #Start Value
```

```
      beta = betaStar + 2 * epsilon
      itCount = 0

      betaReset <- 0

      #Note : If only one value passed - mmin = mmax = mbar & exception will be thrown
      while((abs(betaStar-beta) >= epsilon) && (itCount <= maxIt)){
        itCount = itCount + 1
        beta = betaStar
        betaStar = beta - f_beta(beta, mbar, mmin, mmax) / fPrime_beta(beta, mmin, mmax)
        #print(c(betaStar, beta, betaReset))
        if((abs(betaStar) == Inf) && (betaReset < 10)){
          betaReset = betaReset + 1
          if(betaReset == 1){betaStar = 7}
          else if(betaReset == 2){betaStar = 10}
        }
      }
      return(betaStar)
}
```

./RCode/EQFunc_bValue_Est.r

I.1.4 Implementation of GOF threshold estimation algorithm

```
findMc.WiemerWyss2000 <-function(dImpEQ, binWidth=0.1, blnPrintSteps=F, lngBootStrapResamples=-1,
    signif=-1, mmin=-1, mmax=-1){

    ffDetectFunc <- function(dImpEQ, binWidth, blnPrintSteps, signif, mmin, mmax){
      #Full Magnitudge Range
        dImpBinIncEQ = EQBinning(dImpEQ, binWidth)
        dImpBinCumEQ = EQBinning(dImpEQ, binWidth, "cum")     #plot(dImpBinCumEQ[,1], log10(
            dImpBinCumEQ[,2]))

        if(mmin==-1){minMag_ImpData = max(min(dImpBinIncEQ[,1]) - binWidth / 2,0)}else{minMag_ImpData
             = mmin}
        if(mmax==-1){maxMag_ImpData = max(dImpBinIncEQ[,1]) + binWidth / 2}else{maxMag_ImpData = mmax
            }

        numBins = NROW(dImpBinIncEQ)

        Rstat = cbind(dImpBinIncEQ[-numBins,1],matrix(0,(numBins-1),1),matrix(0,(numBins-1),1))
        ABstat = cbind(dImpBinIncEQ[-numBins,1],matrix(0,(numBins-1),3))

      for(i in 1:(numBins - 1)){
        #print(paste(1," : ", i, " of ", numBins, "=", numBins))
        #Setting up cencored data - for detection procedure
          dCenBinIncEQ = dImpBinIncEQ[i:NROW(dImpBinIncEQ),]
          minMag_CenData = min(dCenBinIncEQ[,1]) - binWidth / 2

          #Binning
            dCenBinCumEQ = dCenBinIncEQ
            for(j in 1:NROW(dCenBinIncEQ)){dCenBinCumEQ[j,2] = sum(dCenBinIncEQ[j:NROW(dCenBinIncEQ)
                ,2])}

        #print(paste(2," : ", i, " of ", numBins, "=", numBins))

          dCenEQ = dImpEQ[which(dImpEQ >= dImpBinIncEQ[i,1])]
          betaHat = betaValue.mle.PageRelation(dCenEQ)
          estBValue = betaHat / log(10)
          estAValue = log10(dCenBinCumEQ[1,2]) + estBValue * minMag_CenData

        #print(paste(3," : ", i, " of ", numBins, "=", numBins))

        #Estimating Expected numbers (Entire Range)
            ExTotalEQ = 10 ^ (estAValue - estBValue * minMag_ImpData)

          #Incremental
            dER_ExBinIncEQ = dImpBinIncEQ
            dER_ExBinIncEQ[,2] = mBinGR(dER_ExBinIncEQ[,1], estBValue, minMag_ImpData, maxMag_ImpData
                ) * ExTotalEQ

        #print(paste(4," : ", i, " of ", numBins, "=", numBins))

          #Binning
            dER_ExBinCumEQ = dER_ExBinIncEQ
            for(j in 1:NROW(dER_ExBinIncEQ)){dER_ExBinCumEQ[j,2] = sum(dER_ExBinIncEQ[j:NROW(dER_
                ExBinIncEQ),2])}

        #print(paste(5," : ", i, " of ", numBins, "=", numBins))
```

```
49
50          #Estimating GOF Statistic
51            Rstat[i,2] = 100 * (1 - sum(abs(dER_ExBinCumEQ[i:NROW(dER_ExBinCumEQ),2] - dCenBinCumEQ
                  [,2])) / dCenBinCumEQ[1,2] )
52            tempData <- dImpEQ[which(dImpEQ > minMag_CenData)]
53
54            #x <- try(ks.test(tempData, "dGR", bValue=betaHat/log(10), mMin=minMag_CenData, mMax=mmax),
                   silent=!blnPrintSteps)
55            if(FALSE){Rstat[i,3] = x$p.value} # !inherits(x, "try-error") #Error will be printed even
                  though is will continue with rest of program
56            if(blnPrintSteps){print(paste("mmin = ", minMag_CenData," - numObs = ", NROW(dImpEQ[which(
                  dImpEQ > minMag_CenData)]), " - Rstat = ", Rstat[i,3], sep=""))}
57
58          #print(paste(6," : ", i, " of ", numBins, "=", numBins))
59          #Additional Output data
60            ABstat[i,2] = estAValue
61            ABstat[i,3] = estBValue
62            ABstat[i,4] = ExTotalEQ
63
64      }
65
66        ret <- data.frame(BinCenterMag = ABstat[,1]) #which is also Rstat[,1]
67        ret$estValue_a <- ABstat[,2]
68        ret$estValue_b <- ABstat[,3]
69        ret$estValue_TotalNumEQs <- ABstat[,4]
70
71        ret$Stat_ModifiedGOF <- Rstat[,2]
72        ret$Stat_KS.pVal <- Rstat[,3]
73
74        #Find Mc
75          mc_GOF <- ret[complete.cases(ret),]
76          mc_GOF <- mc_GOF[which(mc_GOF$Stat_ModifiedGOF==max(mc_GOF$Stat_ModifiedGOF)),]$BinCenterMag
                - binWidth/2
77
78          if(signif != -1){
79            mc_KS <- ret[complete.cases(ret),]
80            mc_KS <- min(mc_KS[mc_KS$Stat_KS.pVal >= signif,]$BinCenterMag) - binWidth / 2
81          }else{
82            mc_KS = 0
83          }
84
85        ret <- list(SampleStats=ret, mc_GOF=mc_GOF, mc_KS=mc_KS)
86
87        return(ret)
88    }
89
90      ret <- ffDetectFunc(dImpEQ, binWidth, blnPrintSteps, signif, mmin, mmax)
91
92      if(lngBootStrapResamples != -1){
93        stats <- matrix(0, lngBootStrapResamples,2)
94        for(i in 1:lngBootStrapResamples){
95          rData <- sample(dImpEQ,size=NROW(dImpEQ),replace=T)
96          temp <- ffDetectFunc(rData, binWidth, blnPrintSteps=F, signif, mmin, mmax)
97          stats[i,1] <- temp$mc_GOF
98          stats[i,2] <- temp$mc_KS
99          if(blnPrintSteps){print(paste("Done : ",i, " of ", lngBootStrapResamples, sep=""))}
100       }
101
102       ret <- list(SampleStats=ret$SampleStats, mc_GOF=ret$mc_GOF, mc_KS=ret$mc_KS ,SamplingDist_GOF=
                stats[,1], SamplingDist_KS=stats[,2])
103     }
104
105   return(ret)
106 }
```

./RCode/ThresholdEstimation_GOF_Exp.r

I.1.5 Implementation of MAXC threshold estimation algorithm

```
1  findMc.WiemerWyss2000.MAXC <- function(dEQ, Indices=-1, binWidth=0.1, blnPrintSteps=T,
       lngBootStrapResamples=-1){
2
3    if(NROW(Indices) != 1){
4        d <- dEQ[Indices]
5    }else{
6        if(Indices != -1){
7          d <- dEQ[Indices]
8        }else{
9          d <- dEQ
```

```
      }
  }

  ffDetectFunc <- function(dEQ, binWidth, blnPrintSteps){
      dBinEQ_Cum <- EQBinning(dEQ, binWidth, "inc")

      max.count <- max(dBinEQ_Cum[,2])

      mc <- as.numeric(dBinEQ_Cum[which(dBinEQ_Cum[,2] == max.count),1])
      if(length(mc) == 0){
        mc <- NA}
      else{
        if(NROW(mc) > 1){
          mc <- mc[NROW(mc)]
        }
      }

      mc <- mc + binWidth / 2

       Ret <- data.frame(mc = mc)
    return(Ret)
  }

  Ret <- ffDetectFunc(dEQ=d, binWidth, blnPrintSteps)

   if(lngBootStrapResamples != -1){
     stats <- vector("numeric",lngBootStrapResamples)
     for(i in 1:lngBootStrapResamples){
       rData <- sample(dEQ,size=NROW(dEQ),replace=T)
       stats[i] <- ffDetectFunc(rData, binWidth, blnPrintSteps=F)$mc
       if(blnPrintSteps==T){print(paste("Done : ",i, " of ", lngBootStrapResamples, sep=""))}
     }
     Ret <- list(mc = Ret$mc, SamplingDist=stats)
   }

  return(Ret)
}
```

./RCode/ThresholdEstimation_MAXC.r

I.1.6 Implementation of bVS threshold estimation algorithm

```
findMc.CaoGao2002 <- function(dEQ, binWidth=0.1, numBinsToAve=5, blnPrintSteps=T,
    lngBootStrapResamples=-1, numBinsLeftOfAve = 0){

  ffDetectFunc <- function(dEQ, binWidth, numBinsToAve, blnPrintSteps, numBinsLeftOfAve){
      dBinEQ_Cum <- EQBinning(dEQ, binWidth, "cum")

        tempRet <- base::matrix(1,(NROW(dBinEQ_Cum)-1), 6)

        if(NROW(tempRet) != 0){
          #Set up subsamples & Statistics
          for(i in 1:(NROW(dBinEQ_Cum)-1)){
            mco <- dBinEQ_Cum[i,1]
            #print(paste(i, " ~~~ ", mco, sep=""))
            dSampleEQ <- dEQ[which(dEQ >= mco)]
            MinMag_Sample <- min(dSampleEQ)
            bVal1 <- try(betaValue.mle.PageRelation(dSampleEQ, mmin=(mco - binWidth/2)) / log(10),
                 silent=T)
            if(inherits(bVal1, "try-error")){bVal1 <- NA}

            bVal2 <- mle.bValue.Cont.trunc.left(dSampleEQ, MagMin=(mco - binWidth/2))


            temp <- dSampleEQ - mean(dSampleEQ)
            temp <- t(temp) %*% temp
            numObs <- NROW(dSampleEQ)
            deltab <- 2.3 * (bVal2 ^ 2) * sqrt(temp / (numObs * (numObs - 1)))

            #Insert Data
            #print(paste(i, " ~~~ ", mco, " ~~~ ", NROW(dBinEQ_Cum), sep=""))
            #print(dSampleEQ)

            tempRet[i,1] <- mco
            tempRet[i,2] <- numObs
            tempRet[i,3] <- MinMag_Sample
            tempRet[i,4] <- bVal1
            tempRet[i,5] <- bVal2
            tempRet[i,6] <- deltab
```

```
36
37            if(blnPrintSteps){print(paste("(",i,"/", NROW(dBinEQ_Cum) ,")", " m_co=", format(mco,
                  nsmall=2), " || MinMag= ", format(MinMag_Sample,nsmall=8), "|| numObs=", format(
                  numObs, nsmall=3) ," || delta.b=", format(deltab, nsmall=4) ," || bVal1=", format(
                  bVal1,nsmall=8), " || bVal2=", format(bVal2,nsmall=8), sep=""))}
38          }
39          ##### REMEMBER !!! sTATISTICS ARE ONLY CALCULATED FOR THE FIRST (N-1) MAGNITUDE BINS !!!!
40
41        #Search for Mc
42  #       bDiff <- matrix(0, (NROW(dBinEQ_Cum) - numBinsToAve - 1), 2)
43  #       for(i in 1:NROW(bDiff)){
44  #         #print(paste(i, " of ", NROW(bDiff), " and MaxRow= ", NROW(tempRet) , sep="" ))
45  #         bDiff[i,1] <- mean(tempRet[i:(i+numBinsToAve - 1),4])
46  #         bDiff[i,2] <- mean(tempRet[i:(i+numBinsToAve - 1),5])
47  #       }
48
49        #Search for Mc
50        bDiff <- base::matrix(NA, (NROW(dBinEQ_Cum) - 1), 2)
51        for(i in (numBinsLeftOfAve+1):(NROW(bDiff) - numBinsToAve + numBinsLeftOfAve + 1)){
52          #print(paste(i, " from ",(numBinsLeftOfAve+1)," to ",(NROW(bDiff) - numBinsToAve +
                numBinsLeftOfAve + 1), " and MaxRow= ", NROW(tempRet) , sep="" ))
53          #print(tempRet[(i-numBinsLeftOfAve):(i - 1 + (numBinsToAve - numBinsLeftOfAve)),4])
54          bDiff[i,1] <- mean(tempRet[(i-numBinsLeftOfAve):(i - 1 + (numBinsToAve - numBinsLeftOfAve))
                ,4])
55          bDiff[i,2] <- mean(tempRet[(i-numBinsLeftOfAve):(i - 1 + (numBinsToAve - numBinsLeftOfAve))
                ,5])
56        }
57
58        #print(cbind(bDiff,tempRet[,4:5]))
59
60        bDiff[(numBinsLeftOfAve+1):(NROW(bDiff) - numBinsLeftOfAve),] <- abs(bDiff[(numBinsLeftOfAve
              +1):(NROW(bDiff) - numBinsLeftOfAve),] - tempRet[(numBinsLeftOfAve+1):(NROW(bDiff) -
              numBinsLeftOfAve),4:5])
61
62        #print(bDiff[1,])
63
64        Ret <- data.frame(mco = tempRet[1:NROW(bDiff),1])
65        Ret$Sample_numObs             <- tempRet[1:NROW(bDiff),2]
66        Ret$Sample_MinMag             <- tempRet[1:NROW(bDiff),3]
67        Ret$Sample_bVal.Page          <- tempRet[1:NROW(bDiff),4]
68        Ret$Sample_bVal.AkiUtsu       <- tempRet[1:NROW(bDiff),5]
69        Ret$Sample_bValVar.ShiBolt    <- tempRet[1:NROW(bDiff),6]
70
71        Ret$Sample_bDiff.Page         <- bDiff[,1]
72        Ret$Sample_bDiff.AkiUtsu      <- bDiff[,2]
73
74        mc <- Ret[complete.cases(Ret),]
75        #mc <- mc[which(mc$Sample_bDiff.Page <= mc$Sample_bValVar.ShiBolt),]
76        mc <- mc[which(mc$Sample_bDiff.AkiUtsu <= mc$Sample_bValVar.ShiBolt),]
77        mc <- mc[which(mc$mco == min(mc$mco)),]
78        mc <- as.numeric(mc$mco)
79
80        if(length(mc) == 0){mc <- NA}
81      }else{
82        Ret <- data.frame(mco = NA)
83        Ret$Sample_numObs             <- NA
84        Ret$Sample_MinMag             <- NA
85        Ret$Sample_bVal.Page          <- NA
86        Ret$Sample_bVal.AkiUtsu       <- NA
87        Ret$Sample_bValVar.ShiBolt    <- NA
88
89        Ret$Sample_bDiff.Page         <- NA
90        Ret$Sample_bDiff.AkiUtsu      <- NA
91        mc <- NA
92      }
93
94      Ret <- list(SampleStats=Ret, mc=mc)
95      return(Ret)
96    }
97
98    Ret <- ffDetectFunc(dEQ, binWidth, numBinsToAve, blnPrintSteps,numBinsLeftOfAve)
99
100   if(lngBootStrapResamples != -1){
101     stats <- vector("numeric",lngBootStrapResamples)
102     for(i in 1:lngBootStrapResamples){
103       rData <- sample(dEQ,size=NROW(dEQ),replace=T)
104       stats[i] <- ffDetectFunc(rData, binWidth, numBinsToAve, blnPrintSteps=F, numBinsLeftOfAve)$mc
105       if(blnPrintSteps==T){print(paste("Done : ",i, " of ", lngBootStrapResamples, sep=""))}
106     }
107     Ret <- list(SampleStats=Ret$SampleStats, mc = Ret$mc, SamplingDist=stats)
108   }
109
110   return(Ret)
```

```
}
```

./RCode/ThresholdEstimation_bVS_Exp.r

I.1.7 Implementation of MEMR threshold estimation algorithm

```
findMc.EMR2005 <- function(dEQ, mmin=-1, mmax=-1, blnPrintSteps=F, lngBootStrapResamples=-1){

  ffDetectFunc <- function(obsEQ, mmin=-1, mmax=-1, binWidth=0.1){
    ##   Specify Log likelihood functions   #~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    #Detection probability moddled by Normal distribution with sharp m_c
    ffCNorm <- function(mmin, mc, mmax, beta, mu, sigma){
      C1 <- (1 - exp(-beta*(mmax - mmin)))^(-1)
      mu2 <- mu - beta * sigma^2

      ret <- 1 - (C1 / pnorm(mc, mu, sigma)) * (pnorm(mc, mu, sigma) - pnorm(mmin, mu, sigma) - exp(-
          beta*(mu - mmin - 0.5*beta*sigma^2)) * (pnorm(mc, mu2, sigma) - pnorm(mmin, mu2, sigma)))
      ret <- 1 / ret
      return(ret)
    }

    loglikl.EMR <- function(param, obs, mmin, mmax, mc) {
      #Can't pass normalising constant to here <- is determined by following params :
      mu    <- param[1]
      sigma <- param[2]
      beta  <- param[3]

      C_Norm <- C_Norm <- ffCNorm(mmin, mc, mmax, beta, mu, sigma)

      n <- NROW(obs)

      q <- rep(1,NROW(obs))
      q[which(obs < mc)] <- pnorm(obs[which(obs < mc)], mean=mu,sd=sigma) / pnorm(mc, mean=mu,sd=
          sigma)

      -(n * log(C_Norm) + n * log(beta) - beta * sum(obs - mmin) - n*log(1 - exp(-beta*(mmax - mmin))
          ) + sum(log(q)))
    }

    mle.est <- function(mc, obsEQ, mmin, mmax){
        #Starting Values
        sVal_mu = 0.1 #mean(m_belowMC)
        sVal_sigma = 0.2 #sd(m_belowMC)
        sVal_beta = log(10) #1/(mean(m_aboveMC)-mc)
        sVal_Vec <- c(sVal_mu, sVal_sigma, sVal_beta)

      ##   Optimization Routine #~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
        # Not usuing nlm - not always as robust
        #O1=try(nlm(f = loglikl.EMR, p = sVal_Vec, obs=obsEQ, mmin=mmin, mmax=mmax, mc=mc, hessian=F)
            , silent=T)
        # Not using nlminb - emperically found to return answers slightly off
        O3 <- try(nlminb(sVal_Vec, objective=loglikl.EMR, obs=obsEQ,mmin=mmin,mmax=mmax,mc=mc,hessian
            =TRUE), silent=T)

        loglikl.EMR.GlobalParams <- function(param) {loglikl.EMR(param, obsEQ, mmin, mmax, mc)}
        O2 <- try(optim(par=sVal_Vec, fn = loglikl.EMR.GlobalParams, hessian=TRUE), silent=T)

        indOptimRoutine = 0
        if(!inherits(O2, "try-error")){
          ret.routine <- list(estimate=O2$par, minimum=O2$value, hessian=O2$hessian)
          indOptimRoutine <- 2
        }else if(!inherits(O3, "try-error")){
          ret.routine <- list(estimate=O3$par, minimum=O3$objective, hessian=O3$hessian)
          indOptimRoutine <- 3
        }

        SE <- rep(-1, 3) #Initially
        if(!is.null(ret.routine$hessian)){
          if(det(ret.routine$hessian) != 0){SE <- sqrt(diag(solve(ret.routine$hessian)))}
        }

        params <- data.frame(mmin=mmin, mmax=mmax, mc=mc,
                             est.mu=ret.routine$estimate[1], est.sigma=ret.routine$estimate[2], est.
                                 beta=ret.routine$estimate[3],
                             est.mu.se=SE[1], est.sigma.se=SE[2], est.beta.se=SE[3],
                             mle.val = -ret.routine$minimum, OptimRoutine=indOptimRoutine)

        params <- unique(params)
        return(params)
```

```
      }

    if(mmin == -1){mmin = floor(min(obsEQ))}
    if(mmax == -1){mmax = ceiling(max(obsEQ))}
    mc <- seq(mmin, mmax, binWidth)

    ret <- mle.est(mc[1], obsEQ=obsEQ, mmin=mmin, mmax=mmax)
    for(i in 2:NROW(mc)){ret <- rbind(ret, mle.est(mc[i], obsEQ=dEQ, mmin=mmin, mmax=mmax))}

    retSampleStats  <- ret[complete.cases(ret),]
    retRecord       <- subset(retSampleStats, mle.val==max(retSampleStats$mle.val))
    retMc           <- retRecord$mc

    ret <- list(mc=retMc, SampleStats=retSampleStats,
                muHat = retRecord$est.mu, sigmaHat=retRecord$est.sigma, betaHat=retRecord$est.beta,
                muHat.SE = retRecord$est.mu.se, sigmaHat.SE = retRecord$est.sigma.se, betaHat.SE =
                    retRecord$est.beta.se)
    return(ret)
  }

  Ret <- ffDetectFunc(dEQ, mmin, mmax)

  if(lngBootStrapResamples != -1){
    stats <- vector("numeric",lngBootStrapResamples)
    for(i in 1:lngBootStrapResamples){
      rData <- sample(dEQ,size=NROW(dEQ),replace=T)
      stats[i] <- ffDetectFunc(rData, mmin, mmax)$mc
      if(blnPrintSteps==T){print(paste("Done : ",i, " of ", lngBootStrapResamples, sep=""))}
    }
    Ret <- list(SampleStats=Ret$SampleStats, mc = Ret$mc, SamplingDist=stats,
                muHat=Ret$muHat, sigmaHat=Ret$sigmaHat, betaHat=Ret$betaHat,
                muHat.SE=Ret$muHat.SE, sigmaHat.SE=Ret$sigmaHat.SE, betaHat.SE=Ret$betaHat.SE)
  }

   return(Ret)
 }
```

./RCode/ThresholdEstimation_EMR_Exp.r

I.1.8 Implementation of MITC threshold estimation algorithm

```
#MITC

#Auxiliry Functions
   {
    f_Mg <- function(m, mmin, mmax, beta){
      f <- function(m, mmin, mmax, beta){
        if(m < mmin || m > mmax){
          ret <- 0
        }else{
          ret <- beta * exp(-beta*m) / (exp(-beta*mmin) - exp(-beta*mmax))
        }
        return(ret)
      }
      return(sapply(m, f, mmin=mmin, mmax=mmax, beta=beta))
    }

    F_Mg <- function(m, mmin, mmax, beta){
      f <- function(m, mmin, mmax, beta){
        if(m < mmin){m = mmin}else if(m > mmax){m = mmax}
        return( (1 - exp(-beta*(m - mmin))) / (1 - exp(-beta*(mmax - mmin))) )
      }
      return(sapply(m, f, mmin=mmin, mmax=mmax, beta=beta))
    }

    ffCNorm <- function(mmin, mc, mmax, beta, mu, sigma){
        C1 <- (1 - exp(-beta*(mmax - mmin)))^(-1)
        mu2 <- mu - beta * sigma^2

        1/(1 - (C1 / pnorm(mc, mu, sigma)) * (pnorm(mc, mu, sigma) - pnorm(mmin, mu, sigma) - exp(-
            beta*(mu - mmin - 0.5*beta*sigma^2)) * (pnorm(mc, mu2, sigma) - pnorm(mmin, mu2, sigma)))
            )
    }

    f_Ma <- function(m, mmin, mc ,mmax, beta, mu, sigma, C_Norm = -1){
      if(C_Norm == -1){C_Norm  <- ffCNorm(mmin, mc, mmax, beta, mu, sigma)}

      f <- function(m, mmin, mc ,mmax, beta, mu, sigma, C_Norm) {
        if(m < mmin || m > mmax){
```

```
          ret <- 0
        }else{
          ret <- C_Norm * f_Mg(m, mmin, mmax, beta) * ifelse(m < mc, pnorm(m, mu, sigma) / pnorm(mc,
              mu, sigma), 1)
        }
        return(ret)
      }

      return(sapply(m, f, mmin=mmin, mc=mc, mmax=mmax, beta=beta, mu=mu, sigma=sigma, C_Norm=C_Norm))
    }

    F_Ma <- function(m, mmin, mc, mmax, beta ,mu, sigma, C_Norm = -1){
      mu2     <- mu - beta*sigma^2
      C1      <- (1 - exp(-beta*(mmax - mmin)))^(-1)
      C2      <- exp(-beta*(mu - mmin - 0.5*beta*sigma^2))

      if(C_Norm == -1){C_Norm  <- ffCNorm(mmin, mc, mmax, beta, mu, sigma)}

      f <- function(m, mmin, mc, mmax, beta ,mu, sigma, C_Norm, mu2, C1, C2) {
        if(m < mmin){m <- mmin} else if(m > mmax){m <- mmax}

        if(m < mc){
          ret <- (C_Norm / pnorm(mc, mu, sigma)) *
            (
            F_Mg(m, mmin, mmax, beta) * pnorm(m, mu, sigma) - C1*(pnorm(m, mu, sigma) - pnorm(mmin,
                mu, sigma) - C2*(pnorm(m, mu2, sigma) - pnorm(mmin, mu2, sigma)))
            )
        }else{
          ret <- C_Norm *
            (
            F_Mg(m, mmin, mmax, beta) - (C1 / pnorm(mc, mu, sigma)) *
              (
              pnorm(mc, mu, sigma) - pnorm(mmin, mu, sigma) - C2 * (pnorm(mc, mu2, sigma) - pnorm(
                  mmin, mu2, sigma))
              )
            )
        }
        return(ret)
      }

      return(sapply(m, f, mmin=mmin, mc=mc, mmax=mmax, beta=beta ,mu=mu, sigma=sigma, C_Norm=C_Norm,
          mu2=mu2, C1=C1, C2=C2))
    }

    loglikl.EMR <- function(param, obs, mmin, mmax, mc) {
          #Can't pass normalising constant to here <- is determined by following params :
          mu    <- param[1]
          sigma <- param[2]
          beta  <- param[3]

          C_Norm <- C_Norm <- ffCNorm(mmin, mc, mmax, beta, mu, sigma)

          n <- NROW(obs)

          q <- rep(1,NROW(obs))
          q[which(obs < mc)] <- pnorm(obs[which(obs < mc)], mean=mu,sd=sigma) / pnorm(mc, mean=mu,sd=
              sigma)

          -(n * log(C_Norm) + n * log(beta) - beta * sum(obs - mmin) - n*log(1 - exp(-beta*(mmax -
              mmin))) + sum(log(q)))
        }
    }

    f <- function(mc, dEQ, mmin, mmax, MaxMethod="O1", optimMethod=""){

      m_belowMC <- dEQ[which(dEQ < mc)]
      m_aboveMC <- dEQ[which(dEQ >= mc)]

      sVal_mu = 0.1 #mean(m_belowMC)
      sVal_sigma = 0.2 #sd(m_belowMC)
      sVal_beta = log(10) #1/(mean(m_aboveMC)-mc)
      sVal_Vec <- c(sVal_mu, sVal_sigma, sVal_beta)

      loglikl.EMR.GlobalParams <- function(param) {loglikl.EMR(param, dEQ, mmin, mmax, mc)}

      #MLE for nuisance parameters - beta, mu, sigma
      {
    if(MaxMethod == "O1" | MaxMethod == "mixed") O1=try(nlm(f = loglikl.EMR, p = sVal_Vec, obs=dEQ,
        mmin=mmin, mmax=mmax, mc=mc, hessian=T), silent=T)

    if(MaxMethod == "O2" | MaxMethod == "mixed"){
      if(optimMethod == ""){oMethod <- "Nelder-Mead"}else{oMethod <- optimMethod}
      #print(optimMethod)
```

```
113       O2 <- try(optim(par=sVal_Vec, fn = loglikl.EMR.GlobalParams, hessian=TRUE, method=oMethod),
              silent=T)
114     }
115
116     if(MaxMethod == "O3" | MaxMethod == "mixed") O3 <- try(nlminb(sVal_Vec, objective=loglikl.EMR,
            obs=dEQ, mmin=mmin, mmax=mmax, mc=mc, hessian=TRUE), silent=T)
117
118     if(MaxMethod == "O4"){
119       O4 <- maxLik(loglikl.EMR.GlobalParams, start=sVal_Vec)
120     }
121
122     if(MaxMethod == "O1"){
123       muHat = O1$estimate[1]
124       sigmaHat = O1$estimate[2]
125       betaHat = O1$estimate[3]
126     }else if(MaxMethod == "O2"){
127       #print(O2)
128       #print(O2$par)
129       muHat = O2$par[1]
130       sigmaHat = O2$par[2]
131       betaHat = O2$par[3]
132     }else if(MaxMethod == "O3"){
133       muHat = O3$par[1]
134       sigmaHat = O3$par[2]
135       betaHat = O3$par[3]
136     }else if(MaxMethod == "O4"){
137       muHat = coef(O4)[1]
138       sigmaHat = coef(O4)[2]
139       betaHat = coef(O4)[3]
140     }else if(MaxMethod == "mixed"){
141       muHat = O3$par[1]
142       sigmaHat = O2$par[2]
143       betaHat = O1$estimate[3]
144     }
145   }
146
147       mbar_A <- mean(dEQ)
148       mbar_A_AboveMC <- mean(m_aboveMC)
149       x <- integrate(F_Ma, lower=mmin, upper=mc, mmin=mmin, mc=mc, mmax=mmax, beta=betaHat, mu=muHat,
              sigma=sigmaHat)
150
151       F_Mg.mc <- F_Mg(mc, mmin, mmax, betaHat)
152       F_Ma.mc <- F_Ma(mc, mmin, mc, mmax, betaHat, muHat, sigmaHat)
153
154       f.mc <- mbar_A - mbar_A_AboveMC * (1 - F_Ma.mc)  + x$value - mc*F_Ma.mc
155
156       ret <- data.frame(func.Val=f.mc, betaHat=betaHat, muHat=muHat, sigmaHat=sigmaHat)
157
158      return(ret)
159     }
160
161     f_Solve <- function(mc, dEQ, mmin, mmax, MaxMethod, optimMethod, blnRet.OnlyFuncVal=T){
162       x <- f(mc=mc, dEQ=dEQ, mmin=mmin, mmax=mmax, MaxMethod=MaxMethod, optimMethod)
163
164       if(blnRet.OnlyFuncVal){
165         return(x$func.Val)
166       }else{
167         return(x)
168       }
169     }
170
171     ffFindFirstSignChange <- function(func, domain.interval, blnRetVal.FunctionValue=TRUE, ...){
172       xCurr <- func(domain.interval[1], ...)
173       xNext <- func(domain.interval[2], ...)
174       i <- 1
175       while((sign(xCurr) == sign(xNext)) & (i + 2 <= NROW(domain.interval))){
176         i <- i + 1
177         xCurr <- xNext
178         xNext <- func(domain.interval[i + 1], ...)
179       }
180
181       if(sign(xCurr) == sign(xNext)){
182         ret <- NA
183       }else{
184         ret <- c(i, i + 1)
185         if(blnRetVal.FunctionValue){ret <- sapply(ret, func, ...)}
186       }
187       return(ret)
188     }
189
190 findMc.MITC <- function(dEQ, mmin, mmax, blnPrintSteps=T, lngBootStrapResamples=-1, MaxMethod="O2",
        optimMethod=""){
191
```

```
  ffDetectFunc <- function(dEQ, mmin, mmax, blnPrintSteps, MaxMethod, optimMethod){
    #Isolate Root
      int.Mag <- seq(mmin+1, mmax-0.1, 0.1)
      intToSearch <- ffFindFirstSignChange(f_Solve, int.Mag, blnRetVal.FunctionValue=FALSE, dEQ=dEQ,
          mmin=mmin, mmax=mmax, MaxMethod=MaxMethod, optimMethod=optimMethod)

    #Continue to find root
      if(NROW(intToSearch) == 2){
      intToSearch <- int.Mag[intToSearch]

      #Estimation
        mcHat <- uniroot(f_Solve, interval=intToSearch, dEQ=dEQ, mmin=mmin, mmax=mmax, MaxMethod=
            MaxMethod, optimMethod=optimMethod)$root

        ret <- f_Solve(mcHat, dEQ=dEQ, mmin=mmin, mmax=mmax, MaxMethod=MaxMethod, optimMethod=
            optimMethod, blnRet.OnlyFuncVal=F)

        ret <- data.frame(mcHat = mcHat, betaHat=ret$betaHat, muHat=ret$muHat, sigmaHat=ret$sigmaHat)

      return(ret)
    }
  }

  Ret <- try(ffDetectFunc(dEQ, mmin, mmax, blnPrintSteps, MaxMethod, optimMethod), silent=T)
  if(inherits(Ret, "try-error")){
    Ret <- data.frame(mcHat = NA, muHat = NA, sigmaHat = NA, betaHat = NA)
  }

   if(lngBootStrapResamples != -1){
     stats <- vector("numeric",lngBootStrapResamples)
     for(i in 1:lngBootStrapResamples){
       rData <- sample(dEQ,size=NROW(dEQ),replace=T)
       x <- try(ffDetectFunc(rData, mmin, mmax, blnPrintSteps=F, MaxMethod=MaxMethod, optimMethod=
           optimMethod), silent=T)

       if(inherits(x, "try-error")){
         stats[i] <- NA
         print("No Return")
       }else{
         if(length(x) != 0){
           stats[i] <- x$mcHat
         }else{
           stats[i] <- NA
         }
       }

       if(blnPrintSteps==T){print(paste("Done : ",i, " of ", lngBootStrapResamples, sep=""))}
     }

     Ret <- list(mc = Ret$mcHat, muHat = Ret$muHat, sigmaHat=Ret$sigmaHat, betaHat=Ret$betaHat,
         SamplingDist=stats)
   }

  return(Ret)
}
```

./RCode/ThresholdEstimation_MITC_Exp.R