Discharge Uncertainty in Frequency Analysis of Han River Discharge

Bachelor Thesis

Timon Klok

University of Twente

8/19/2009

Discharge Uncertainty in Frequency Analysis of Han River Discharge

Bachelor Thesis

Timon Klok

University of Twente

19 August 2009

Foreword

"There can be as much value in the blink of an eye as in months of rational analysis."

Malcolm Gladwell, 2005

This bachelor thesis contains the final product for the completion of my bachelor at the University of Twente. I conducted the research for my thesis in China, at the Hydrology department of Zhejiang University in Hangzhou. The report contains my research into the uncertainty in the frequency analysis of Han River discharges. A frequency analysis gives information on the return period of certain discharges. The application of the research lies in flood safety: the return periods are used for setting a save level for dike height construction around the Han River.

Hereby I'd like to thank my Dutch and Chinese supervisors, dr. Maarten Krol and dr. Yue-ping Xu for their comments. They have helped me to look critically at my own research. I also thank dr. Xu for giving me the opportunity to go to China for this thesis. I have had a great time in Hangzhou, thanks to you and your students.

Enschede, August 18, 2009
Timon Klok

Contents

Summary

Valuating the credibility of research is normally done by reviewing in which journals an article has been published or by how many times an article has been quoted, but still it doesn't tell us exactly how credible the research is. When assessing the research we might look at the confidence intervals, the size of the confidence interval tell us something about the certainty of quantitative information. But not all the (un)certainty can be expressed in confidence intervals, some quality parts in a research can only be valued by a fellow researcher in the same research area. The problem is that results of research are not only red by other scientists, but also by politicians who are looking for grounds for their decisions. For them and other less informed readers Ravetz and Functowicz (1990) proposed the NUSAP method that assesses the uncertainties in a research model. NUSAP is an acronym for Numeral, Unit, Spread, Assessment and Pedigree. The numeral, unit and spread of a model give all quantitative information about the model. The assessment and pedigree part is more an assessment of the quality if the model. In this research the NUSAP method, proposed by, is used to assess the quality of input information for a frequency analysis of the Han River in China. The main question in this research was: *What is the uncertainty of the propagated discharge with a given return period using a frequency analysis for the Baihe discharge station at the Han River?*
The identification of the different uncertainty sources in the frequency analysis is split up into three stages: Measurement (chapter 4), Time series (chapter 5) and Statistics (chapter 6). In each stage the uncertainty sources have been identified.

In the measurement section different methods for the measurement of water level, river profile, velocity and discharge are assessed: for each the spread and Pedigree score have been estimated. The discharges at Baige station are measured according to the two depth velocity area method (ISO, 2007). The measurement error is computed by calculating the uncertainty in the velocity area method and the uncertainty was 3% (95% confidence) in the computed discharges. The NUSAP Pedigree scores are average to high, which means not so much uncertainty.

The time series handles the assessment of the compilation of the peak discharge series. The selection of peak discharges from the time series is done by selecting the annual maximum discharges, but the Exponential distribution needed a threshold of 12.000m^3/s, therefore the 'peaks over threshold' method is used to select peak discharges. The result was one series of AM discharges and one series of POT discharges. The discharge data has not been reviewed for stationary, because there was no information available. As a result the Pedigree score for the time series is low.

The statics of the frequency analysis are assessed by fitting the Normal, Pearson type III and Gumbel distributions to the Annual Maximum and the Exponential distribution is fitted to the POT. The parameter estimation is done with the Method of Moments (MOM) and Maximum Likelihood Estimation (MLE). The goodness-of-fit is tested with the Chi-square test and the Kolmogorov-Smirnov test. A comparison of the distributions with plot-positions of discharges (visual), confidence intervals and the GOF-tests show that the normal distribution has a good fit for the discharges with a return period < 100 years. The Pearson III with MLE parameter estimation has a good fit for return period >100 years. Q_{100} Normal is 23089±1309 m^3/s and PIII MLE is 25019±2258 m^3/s. This fit is explained by the slight S-curve of the measured flows. The Pedigree scores for the different distributions are average to low. This is because the uncertainty of the fit. The different equations give different distributions with a wide range of possible discharges at a give return period.

The main conclusion is that the uncertainty in the flood frequency analysis for the Han River is too large at this moment so that the frequency analysis in this research is not to be of any practical usage at this moment. In this research all the conclusions are drawn upon the differences in discharges. A significant difference in discharge could have relative small impact on the gauge height. Therefore more research on the effects of discharge changes is recommended.

1 Introduction

Information is present all around us. But what about the reliability of all the information, do we assume all the information we get presented to us is correct? Of course not, information about a new product from the manufacturer is valued less trustworthy. Scientific research also has to be valued for its credibility. To value the credibility of research is no easy task, one may look at the journals in which the research has been published or by how many other researchers the article has been quoted, but still it doesn't tell us exactly how credible the research is. When assessing the research one might look at the confidence intervals of the research outcome, which tells something about the certainty of quantitative information. The smaller a confidence level is, the more certainty. Other statistic methods are also possible for the analysing of quantitative uncertainty, such as a sensitivity analysis. A sensitivity analysis investigates the consequences of changes in input data and changes in size of data series.

But not all the (un)certainty can be expressed in confidence intervals, some quality parts in a research can only be assessed by a fellow researcher in the same research area. The results of research are not only red by other scientists, but also by politicians who are looking for grounds for their decisions. Policymakers also have to value the information they read on its credibility. For them and other less informed readers Ravetz and Functowicz (1990) proposed the NUSAP method that assesses the uncertainties in a research model. NUSAP is an acronym for Numeral, Unit, Spread, Assessment and Pedigree. The numeral, unit and spread of a model are all quantitative information about the model used in the research that is reviewed. The assessment and pedigree part is more an assessment of the quality of the model. In this research the NUSAP method, proposed by, is used to assess the quality of input information for a frequency analysis of the Han River in China.

Flood frequency analysis is used to compute the return period of certain discharges. In order to get to the frequency analysis other steps are needed. First input data has to be gathered by measuring the depth, water level and velocity of the current according to the velocity-area method (chapter 4). This information is brought together in the discharge of the Han River. In this chapter it is important to know the Unit and Spread of the instruments used for measurements. The uncertainty in the discharge is computed with the uncertainty calculation of the velocity area method. A time series of the measured discharges is created and assessed in chapter 5, in this chapter questions like the stationarity and independency of the data arise. Further, the selection of discharge peaks, annual maximum or peaks over threshold, in the time series is discussed. The next step is the fitting of the Normal, Exponential, Gumbel and Pearson type III distributions to the time series. The reason for using these distributions is explained in chapter 6. The fitting is done using two different methods: the Method of Moments and Maximum Likelihood Estimation. The performance of distributions is evaluated by using different statistical tests. The results of the fitting and testing of the distributions can be found in chapter 6. Each step is assessed by the NUSAP method. Chapter 7 discusses the propagation of the uncertainties in the frequency analysis. A methodical reflection is presented in chapter 8. The conclusions and recommendations can be found in chapter 9.

1.1 Research objective and questions

The objective of this research will be the assessment of uncertainty in the design return period for the Han River, where the discharge is calculated using a frequency analysis. The assessment will be done by using the NUSAP method.

The central research question is:
What is the propagated uncertainty in discharges with a given return period using a frequency analysis for the Baihe discharge station at the Han River?

Answering this question will be a three step process: 1) find uncertainty sources in discharge frequency analysis calculation, 2) then analyse the uncertainty sources, 3) propagate uncertainty. The propagated uncertainty gives a measure for the uncertainty in the discharge for the Han River. This results in the following questions:

- What are the uncertainty sources in the total process toward the frequency analysis?
    - Which kinds of instruments are used in the measurement of the water depth, width and velocity?
    - What method is used to make discharge data more stationary and homogeneous?
    - What are the different functions, distributions and parameter estimations used in the frequency analysis?
- How to quantify the uncertainty in these sources?
- What is propagated effect of these quantified uncertainty sources on the discharge calculated from the frequency analysis?

1.2 Research approach

Flood risk is calculated with the use of statistics. The statistical calculations are based on a time series with the annual maximum discharges for the discharge station in Baihe. With these time series an extrapolation is needed to estimate a discharge that will occur once every x years. The uncertainty analysis in this research paper starts with the assessment of uncertainty in the measured runoffs and water levels in the Han River. The next step is the assessment of the uncertainty in the peak discharges in the time series. After that the same can be done in the frequency analysis method. When the uncertainty in each separate step is known the propagated uncertainty can be calculated.

The method that will be used for the assessment of the uncertainty in the frequency analysis is presented in figure 1-1.

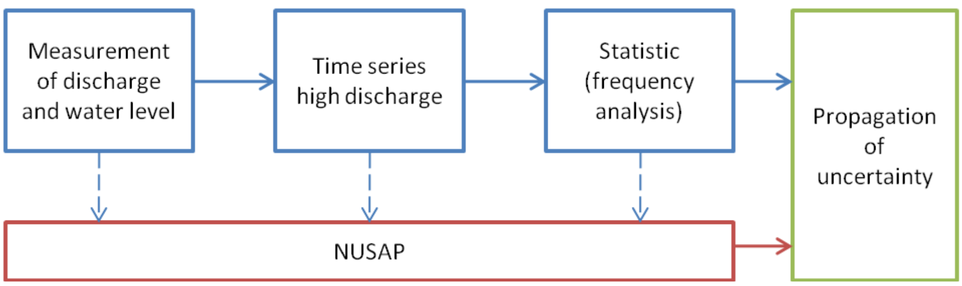


Figure 1.1 Method for assessment of uncertainty

Measurement is the first step. Measurement concerns the discharges and water levels, also a rate of flow and surface measurement of the cross-section profile of the river. Rate of flow is expressed in m/s and is multiplied by the river's cross-section (m^2); the result is the discharge in m^3/s. The most important uncertainties of the measurement phase are:

- Uncertainty in measurement data
- Uncertainties with regard to the execution of these measurements
- Uncertainties with regard to the functioning of the measuring instrument.

The data contain mostly daily runoff and some water levels. But it is important to have knowledge of the uncertainty sources in the discharges.

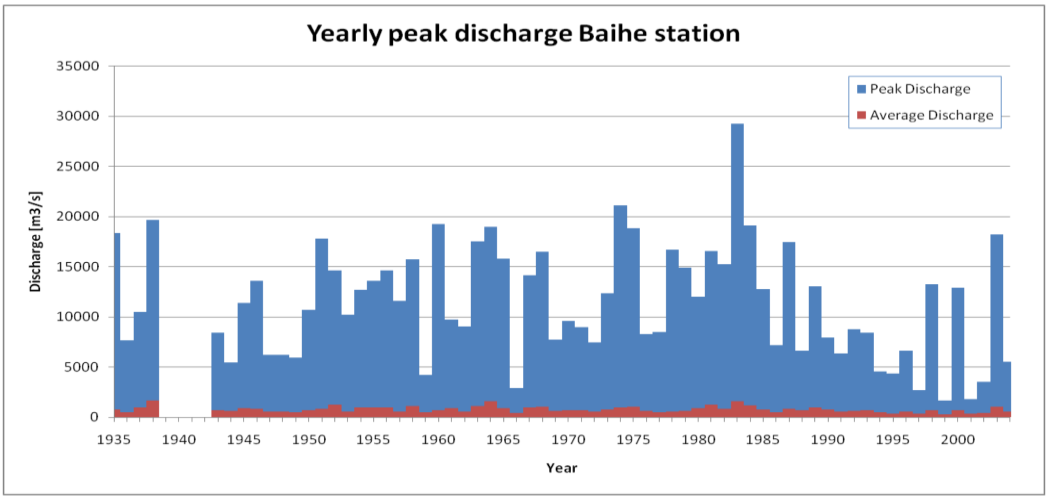


Figure 1.2 Time-series with yearly peak discharges in period 1935-2004

The second step consists of combining of the derived peak discharges in a time series. This time series is used as an input for the statistical calculations. The time series have to be adapted for statistic usage, the purpose of these corrections is a more homogeneous discharge series. The discharges appeared in various conditions, non homogeneity of data can be caused by (Booij & Otter, 2007a):

- Difference in the measurement methods
- Changes in the geometry of the main river
- Changes in the geometry of the tributaries
- Changes in human activities like urbanisation and dams.
- Changes in precipitation because of climate change

The data must be stationary and independent for the statistic calculations. The discharge data from the 'Baihe' discharge station are not stationary yet. The uncertainty of the time series is dependent on the derived Q. The existing knowledge of how to correct the data is also an uncertainty.

After the correction of the time series a frequency analysis can be done, this is the third step. The frequency analysis makes it possible to calculate the recurrence time of certain peak discharges. When the recurrence times for certain peak discharges are known, the flood safety of the present dikes can be assessed. But this will not be done in this research because there is not enough time.

Jansen (2007) analyzed every step in the previous explained process according to the NUSAP method (Sluis et al., 2003). With the use of the NUSAP method the qualified and quantified uncertainties of every step could be assessed. The NUSAP method is explained in the theoretical framework. The first step in the NUSAP method is a traditional standard in the uncertainty analysis. Every input parameter is a possible uncertainty source. This step can be done rather quickly, without spending too much time on it. Much of this step is already known, because of the use of proven models.

The next step is extra in the NUSAP method. That second step starts with the classification of uncertainties, using the NUSAP matrix, much like Walker's (2003) classification, and that is input the identification, rather than the other way around. Every input in the model left out in the first step can now also be identified. The identification is a process that has to be done carefully. The exact way of using the NUSAP method is described in the theoretical background. With the uncertainty in the five steps known, the propagation of these uncertainties can be calculated. The result of the process is the knowledge of the uncertainty in frequency analysis.

For this research the same method as Paul Jansen (2007) will be used. In each step the uncertainty sources will be identified and, if possible, quantified. In 2007 Paul Jansen (2007) used this method to make an assessment of the flood risk uncertainty in the Meuse River in the Netherlands, see also figure 1-2. During this project there was not enough information about the QH-relation. The data already gives the discharge needed for the third step in the process. The last step (flood safety) also won't be done, because of lack of time.

2 Problem analysis

The territory of the People's Republic of China accommodates one of the longest rivers in the world, the Yangtze River (Chinese name: Chanhjiang river), with a total length of 6.380 kilometre and a basin area of 1.9 million km^2. The river basin extends over a vast area. The Yangtze River receives water from many tributaries and thus the average discharge gradually increases, discharge at Wuhan (about 1200 km from mouth of Yangtze) is roughly 24.000 m^3/s. At the mouth of the Yangtze the average discharge has increased to an astonishing 311.000 cubic metres per second (Yangtze River, 2009).

One of the greatest and most important tributaries of the Yangtze River is the Hanjiang River (Han Shui). The Han River has a total length of 1.532 km and a basin area of 170.400 km^2. The basin has a sub-tropical monsoon climate and has, as a result, dramatic diversity in its water resources (Chen et al., 2007). The river changes names a few times from its source, Yudai, the Yang, below Mianxian the name changes to the Mian, at Hanzhong it becomes the Han River (Han River, 2009). The lower course of the Han River flows through lowland, the area is so flat that a small change in the level of the river may inundate a considerable area, and extensive dikes are required. Above Xiangfan at Jun Xian, where the Han receives the Dan River, a dam completed in 1970 stabilizes the water flow, prevents flooding, extends the range of navigation, and permits irrigation. Further downstream at Xiangfan the river receives its largest tributary, the Baishui River. In the 1950s, in order to prevent flooding, a large retention basin was built at the confluence with the Baishui to accumulate floodwaters and to regulate the flow of the Han itself; four extensive irrigation projects were also built in the area. Toward the junction of the Han with the Yangtze, the river narrows sharply. That area, too, has been known to frequent and disastrous flooding, and, to prevent this, in 1954 a second retention basin was built south of the junction with the Yangtze (Han River, 2009). The location of the various dams and weirs can be found on the more detailed map in Appendix A.

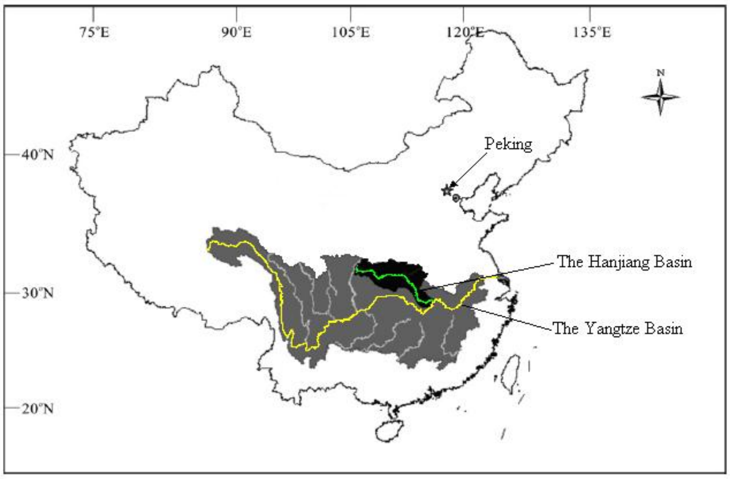


Figure 2.1 Location of Han River in China

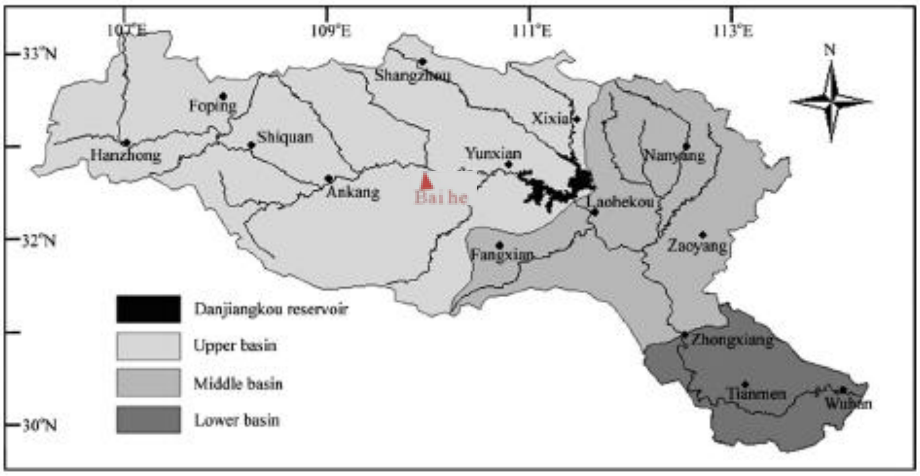


Figure 2.1 Map of Hanjiang basin (Chen, Guo, & Xu, 2007)

2.1 Case study area

The Danjiangkou reservoir is the largest water reservoir in the Han River. The reservoir is used for the 'South to North Water Diversion Project' in China. This means that water from the Yangzte is transported to the dry North region of China, on the height of Beijing. The reservoir and the extraction of large quantities of water have a great influence on the discharge data. For this project a discharge station above the Danjiangkou reservoir has been chosen. The time schedule of the project prevents any deep insight analysis on the stationary of the discharge data, before and after the completion of the reservoir. Other important criterion for the selection of a discharge station is the availability of uncertainty data. Only four stations in the Hanjijang basin give information about the uncertainty in the data they provide. Baihe discharge station is the only station that satisfies both criteria. The station is a relative old one in China, since 1935.

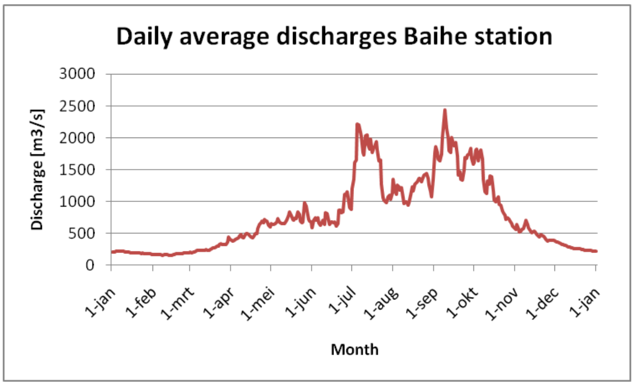


Figure 2.2 Daily average discharge Baihe station

The river upstream of Baihe station is mostly feed by precipitation. The river has yearly two distinctive high precipitation seasons, one from mid June to the end of July and one from late August to early October. The high discharges and flood treats occur mostly in July and September, although this is not a guarantee. Because the chance of an overlap in peak discharges is small, a Gregorian calendar year is used, instead of a hydrological year. Discharges at Baihe station are monitored daily. The normal path of the precipitation is according with the course of the river, because of that combined high discharge waves can occur, which pose a greater flood treat. The frequency of precipitation with an intensity of about 100mm is highest in July, second September and third August. The last decades show a light shift of this peak towards October, but it is not certain if this is a permanent shift.

The basin upstream of Baihe station is a mountainous; this means the ground is rocky, which means low permeability. Combined with the characteristics of the precipitation as well as the small capacity to store water in the rivers, a peak discharge wave resulting from the precipitation may last for 5 to 7 days, with a sharp peak shape.

Baihe station has had different locations in the past. The station was built in 1935. In 1943-1947 the station was sometimes closed for several weeks. In 1950 the station moved 300 meters downstream. In 1957 the station moved 1000 meters upstream from its last position. The station is still on that same position today. It is not known if there are significant water inflows in the sections over which the Baihe station moved. If there are significant tributaries the data series have to be corrected for theses flows.

3 Theoretical framework

3.1 Frequency Analysis

Frequency analysis is the estimation of how often a specified event will occur. Estimation of the frequency of extreme events is often of particular importance. Because there are numerous sources of uncertainty about the physical processes that give raise to observed events (a flooding), a statistical approach to the analysis of data is often desirable (Hosking & Wallis, 1997, p. 1). Let's say we want to construct a dike that may only fail once every 10.000 years, then we need to know the river's discharge with return period of 10.000 years, but we only have measured the rivers runoff for about 50 years. Than the 50 years discharge data are analysed in frequency analysis, to compute the return period of certain discharges. Generally speaking the computation of the return period is done with the frequency analysis. So the frequency analysis gives an idea of the return period of certain peak discharges. These high discharges are derived from measured discharges from data collected since 1950, so the peak discharges are not actually measured; only the lower discharges are measured. The lower discharges are then extrapolated to find discharges that will occur once every ten thousand years or so. Distribution and extrapolation of measured discharges may cause large (more than 5%) and unwanted uncertainty (Morgan & Henrion, 1990).

Between the measurement of actual discharges and the determination of the peak discharges is a model, the measured data point can be used as input for the model. The model is also called a distribution function. .There are different kinds of distribution types, the most common distribution families used for return period calculations of discharges are: Normal distributions, the Gamma family and Extreme value (Gumbel) distributions. Other distributions are Wakeby and Logistic distributions (Rao & Hamed, 2000). The distribution functions have multiple, mostly two or three, parameters so that the distribution functions can be fitted to the measured discharges.

The estimation of the parameters can be done by using different parameter estimation methods. A small list by Rao & Hamed (2000) of different methods: Method of moments (MOM), maximum likelihood estimation (MLE), probability weighted moments method (PWM), least squares method (LS), maximum entropy (ENT), mixed moments (MIX), generalized method of moment (GMM) and incomplete means method (ICM). The method of moments is a relative easy parameter estimation method. Because of its simplicity, the estimates are of inferior quality. Distributions with three or more parameters that have to be estimated are more likely to have biases, especially in combination with smaller data series. The maximum likelihood estimation method is considered the most efficient method compared to other methods (Rao & Hamed, 2000).

The performance of distributions is evaluated by using different statistical tests. The goodness of fit of the distribution is assessed by using goodness-of-fit tests. The most common tests used for the selection of probability distribution functions are the Chi square test and the Kolmogorov-Smirnov test (Rao & Hamed, 2000). Based on the result of the X^2 test and the KS-test and a visual comparison between the distribution and the plot positions of the discharge data a distribution can be selected to have a good fit, which means the propagated discharges with a larger return period than measured ones will be estimated correctly. And correct an estimation of discharges with large return periods is, after all, the goal of a frequency analysis.

3.2 Uncertainty

There are two groups who both use uncertainty, but look at it in a different way; scientists and decision makers. Scientists work often with uncertainties in knowledge, for instance uncertainty in model outcome. Decision makers have to deal with uncertainty in decision variables and priorities, but decision is also based on scientific research; politicians therefore need to keep an eye on those uncertainties too. Wind, De Blois, Kok, Peerbolte, & Green (1997) divided uncertainty in decision-making process into two types of uncertainty, namely outcome uncertainty and decision uncertainty. Outcome uncertainty is the earlier described uncertainty originating from model selection, data availability and scenario development. Decision uncertainty is always present. It is the uncertainty of not knowing everything, of conflicting interests. In multi-criteria analyses measures are commonly prioritised, this can be done in different ways, with different outcomes. Methods to do this have an uncertainty too (Xu, 2005, p. 10). This research focuses on outcome uncertainty, and does not focus on the decision-making process.

In case of an uncertainty analysis a systematic identification and classification of the most important uncertainties has to be made. Walker et al. (2003) classifies uncertainty in three different dimensions. The three dimensions of uncertainty distinguished by Walker et al. (2003) are Location, Level and Nature. The *location* of uncertainty is an identification of where uncertainty manifests itself within the whole model complex. The *level* of uncertainty is a particular determinant for an uncertainty source if it is quantifiable. The *nature* of uncertainty is uncertainty due to the imperfection of knowledge or due to the inherent variability of the phenomena being described.

The identification of the most important sources of uncertainties is based on a sensitivity analysis. After the completion of the sensitivity analysis the uncertainty analysis can start. First step is to quantify the most important uncertainties. Walker et al. (2003) tells us that whether or not a variable or parameter can be quantified depends on the nature of this variable or parameter and the nature of the uncertainty. If literature doesn't provide suitable information about the quantifiability of uncertainties, then expert opinions can be used. The method used in this research for the quantification and assessment is the NUSAP- method, this will be discussed later.

Next step is to determine the propagation of the uncertainties. The aim in propagating uncertainty is to be able to quantify the uncertainty in model outputs. Methods that describe propagation techniques are mentioned by Morgan and Henrion (1990) and include: response surface and Monte Carlo simulation.

3.3 NUSAP method

Issues of uncertainty, and closely related, those of quality of information are involved whenever research related to policy is utilized in the policy process. Up to now, tests for the quality of quantitative information have been much undeveloped. There are standard statistical tests on sets if numbers in relation to a hypothesis; and there are highly elaborated formal theories of decision-making in which "uncertainty" is manipulated as one of the variables. But none of these approaches help with an important question: is this reliable, can I use this information safely? (Ravetz & Funtowicz, 2009)

"Science is based on numbers, therefore numbers are necessary for the effective study of the world; and we assume that numbers, any numbers, are sufficient as well. We still use statistics, usually quite uncritically, because there is nothing better to hand." (Ravetz & Funtowicz, 2009)

The NUSAP method is proposed by Ravetz & Funtowicz (1990) and can be classified as a notational system for quantitative information, by which these difficulties can, to some extent at least, be

overcome. It is based in large part on the experience of research work in the matured natural sciences.
When using models, of all sorts in various sciences, scientists should be aware of the uncertainties and their propagation in the model. Uncertainties in the input should be suppressed if possible, else the outputs become indeterminate.
The NUSAP method allows both quantitative and qualitative aspects to be analyzed in the uncertainty analysis. The method has been used before by the Dutch National Institute for Public Health and the Environment (RIVM) and by the Netherlands Environmental Assessment Agency (PBL). The following description about the NUSAP method is partially copied from Van der Sluijs (2005a).

The NUSAP method is based on five categories, which generally reflect the standard practice of the matured experimental science. By providing a separate box for each aspect of the information, it enables a great flexibility in their expression. The name "NUSAP" stands for Numeral, Unit, Spread, Assessment and Pedigree. The first three are the normal quantitative aspects of the analysis; the last two boxes are the more qualitative part of the method.

3.3.1 Quantitative

Numeral: When analyzing a data string the dimensions of these numbers are relevant. It shows the importance of large numbers. 1E6 + 5E0 = 1E6. The 5E0 doesn't matter because of the much larger number 1.000.000.
Unit: The conventional sort. In this research it will be the water level (meter), velocity (m/s) and the discharge (m^3/s). These data has one important extra piece of information attached; the date they where produced. The date can tell us something about the circumstances in which the data where obtained. The Unit is inherent to the analysis of the data and therefore will be analyzed once.
Spread: generalalizes from the "random error" of experiments or the "variance" of statistics. Although Spread is usually conveyed by a number (either ±, % or "factor of") it is not an ordinary quantity, for its own inexactness is not the same sort as that of measurements.

3.3.2 Qualitative

Assessment: The qualitative assessment is correlated with the Pedigree table, which is discussed next. The Pedigree table makes a distinction between empirical, methodological and statistical assessment criteria. Before using the table, these aspects have to be analyzed first.
Pedigree: The pedigree is an evaluative description of the mode of production of the information. Each sort of information has its own pedigree. The pedigree is expressed by means of a matrix. The columns represent the empirical, methodological and statistical assessment criteria, and within each column there are modes, normatively ranked descriptions. These are numerically graded, so that with a coarse arithmetic, a "quality index" can be calculated for use in Assessment if desired. The grades start with 4 in the top row (ranked high) to zero in bottom row (poor). The assessment is done by finding similarities between the qualities described by NUSAP and the qualities observed in the Assessment analysis.

For each part of the total process the way a method is used has to be identified. Then the Pedigree matrix can be used.

Score	Statistical quality	Empirical quality	Methodological quality
4	Excellent fit to well-known statistical model (Normal, Lognormal, Binomial)	Controlled experiments and large sample direct measurements (n≥50)	Approved standard in well-established discipline
3	Good fit to a reliable statistical model by most fitting test, but not all	Historical/field data, uncontrolled experiment, small sample direct measurements (n≤50)	Reliable method, common within discipline
2	Fitting test not significant, model not clearly related to data, or model inferred from similar data	Modeled data, indirect measurements, handbook estimates	Acceptable method, but limited consensus on reliability
1	No statistical tests or fitting, subjective model	Educated guesses, very indirect approximations, “rule of thumbs” estimates	Unproven methods, questionable reliability
0	Ignorance model (uniform)	Pure guesses	Purely subjective model

Table 3.1 Pedigree matrix (Ellis, Li, Yang, & Cheng, 2000)

The individual scores in the matrix are good indications of the gaps in the total process of flood risk calculation.

4 Measurement of discharge

4.1 Introduction

This chapter will evaluate the methods used for the measurement of the water-level, discharges and cross-section of the Han River at Baihe station. The river hydrometric work at Han River is carried out by the bureau of hydrology, Changjiang Water Resources Commission. Discharges in the years 1939-1942 are not measured, also the years 1948-1949 have gaps in discharge data. During this research no exact data was available about the measurement methods and also no data was available during which years specific instruments were used for measurements at Baihe station.

The methods for data gathering have changed since the first measurement at Baihe station. Until 1950 data about the depth and velocity of the Han River were gathered using a wooden boat. These boats didn't have a motor. The measurements were therefore very labor intensive. Also the accuracy of the data was less because of long duration of measurement sessions and the inexactness in positioning of the boat on the river. In the 1960s and 1970s, motor boats were used. As the river channel is wide and shallow in some places, especially the lower part of the river, it was difficult for small motor boats to orient into the main current for measurement. The deep keel of the ship prevented them from reaching shallow river regions for measurement. At the end of 1970s, the use of motor boats, anchored by a large-span cableway, was introduced. This method has been used for at least 12 years. This method is also used during flood periods. During a flood the cable also spanned across the flood plain, so that boats can measure the flood plains too. In the flood plain the cable is every 150 meters anchored to the riverbed.

	Wet river profile	Dry river profile	Sediment	Velocity	Water level	Positioning
Cross-section						
Theodolites		x				x
GPS		x				x
Water level						
Staff gauge					x	
Stilling well					x	
Velocity						
Currents meters				x		
Bridges	x		x	x		
Boats	x		x	x		
Cableway	x		x	x		
Ultra sonic debt sounders	x		x	x		
ADCP	x			x		

Table 4.1 Measurement Instruments used in China today (Cui et al., 2008)

The measurements of the velocity are preferably done at stationary circumstances, but because of rapid changes in river discharge during summer season this becomes difficult. The fluctuations in discharge have an influence on the accuracy of the measurements. Normal duration of one discharge measurement session was about 5 hours in 1983. Today it takes about 3 hours. Shorter session time means fewer changes in the river discharge during session, so the uncertainty becomes less. Still the accuracy during peak discharges can fluctuate with hundreds of cubic meter within a few hours. Rainfall in summer month is the main perpetrator of peak discharges. That is also why the peak discharges have high yearly fluctuation.

The uncertainty in the measurement of an independent variable is normally estimated by taking N observations and calculating the standard deviation. Using this procedure to calculate the uncertainty in measurement of discharge would require N consecutive measurements of discharge with different current meters at constant water level which is clearly impractical. An estimate of the true value of uncertainty has therefore to be made by examining all various sources of error in the measurement. The different measurement methods used in the period of 1935-2006 each have their own uncertainty during (peak) discharges. The uncertainties of each known used instrument will be assessed as good as possible. The specific aspects that will be assessed with the NUSAP method will be explained in the next section.

4.1.1 Assessment of uncertainties

The quantitative elements in the NUSAP method: numeral (N), unit (U) and spread (S) will be assessed together. The qualified uncertainties will be assessed (A) with the Pedigree matrix (P). The uncertainty sources of the measurement uncertainty will be evaluated with two criteria: methodical and empirical quality. Uncertainty sources which have influence on the methodical and empirical quality are:

- Uncertainty in measurement data (empirical quality)
- Uncertainty in the execution of the measurement (methodical quality)
- Uncertainty in the performance and functioning of the measurement equipment (methodical quality)

Differences between measured data and real data are caused by systematic errors and variance. In streamflow it is sometimes difficult to distinguish between random and systematic errors as some errors may be a combination of the two. For instance, where a calibration group rating is used for current meters, each of the meters forming the group may have a plus or minus systematic error which is randomized to obtain the uncertainty in the group rating. A method to assess the systematic error is the calibration of measurement instruments in a controlled environment with possibilities to set an exact discharge, like a laboratory. The systematic error will not be assessed in this research. The variance is the spread and depends on the margin allowed in duplicate measurements. The performance and functioning of the measurement equipment during an experiment cannot be assessed. Logs about measurement sessions should be examined for this purpose, but these logs are not available for this research.

The uncertainty caused by the execution of the measurement is especially relevant during peak discharge measurement sessions. During these sessions regulatory requirements cannot always be followed, because of extraordinarily circumstances. Regulations are important because they standardize the measurements. If not followed the result could be that discharges can be less compared to each other, which results in greater uncertainty.

4.2 Measurement instruments

4.2.1 Determination of cross-section

The cross-section of the river changes constantly, therefore it is important to measure the cross-section frequently, so information keeps up-to-date. The riverbed at Baihe station has a natural course; this means that the river bed (and thus the cross-section) can change because of riverbed erosion / sedimentation and vegetation changes. During a flooding the river profile can change. Unnatural reasons why the river's cross-section changes are: construction of new wharfs or dredging. During peak discharges the riverbed won't change because of sand waves. Sand waves would have an influence on the gauge height of the water level. A temporarily rise in riverbed would cause a lowering of the measured water level.

Information about the cross-section of Han River at Baihe station is available since 1982 to 1985 mostly during high discharges in August and September. Also data about the situation in September 2006 is available. Thus the riverbed change between 1982 and 2006 can be compared, because the location of Baihe station has not changed.

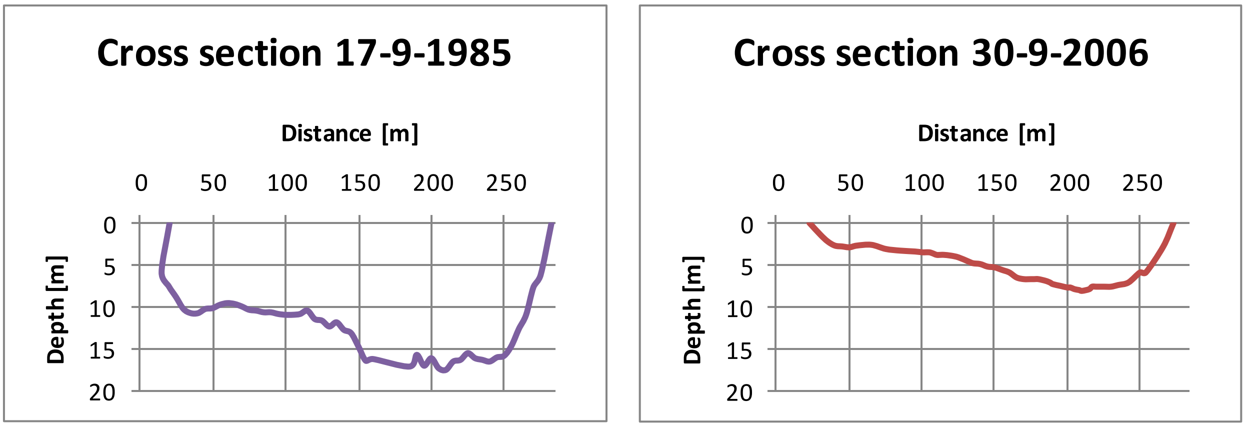


Figure 4.1 Cross section of Han River at Baihe station. Depth is measured from water surface

The cross-sections on 17-9-1985 are measured during peak discharges. Depth is measured from water surface. Data about the depth and width of the Han River are only available between the month May and October. This means that river profile at low discharge in winter month is unknown. The rivers cross-section is measured monthly, sometimes multiple times per month. In this way the cross-section is up to date, and thus the cross-section has less uncertainty. Cross-sections are also measured between and after the years 1985-2006, but this information was not available during research. It is also unknown which exact measurement method and instruments are used for cross-section determination. The cross-sections in figure 4.1 show the difference between the smooth profile in 2006 and the more rough profile in 1985. The main current has stayed on the right side of the river between 1985 and 2006.

One advantage of peak discharges is relative unimportance of faults in the measurement of the cross-section. The absolute measurement errors are expected to grow, but the relative error percentage will become less, because the cross-section grows with larger discharges. So uncertainty may become less with larger discharges. This can be illustrated with an example. Say that depth is measured with error of 30cm. so 40% to low, 50% to high and 10% the right depth. Width is 200 meter and actual depth is 5 meter. The absolute measurement error would be 0,030 m, thus 6m^2 difference in cross-section. The relative error is only 0,6%. If depth would increase to 10m and thus doubling the discharge, the absolute error would still be 6m^2, but the relative error would be 0,3%, which is a 50% decrease of measurement error in cross-section measurement.
The uncertainty in the measurement of the cross-section is part of the uncertainty in the velocity-area method. Therefore the cross-section uncertainty will not be calculated separately; otherwise this uncertainty will be twice accounted for.

4.2.1.1 Gauging-rods and Theodolite

A theodolite is an instrument for measuring both horizontal and vertical angles and can be used to measure surface level. This instrument developed somewhere in the 16th century. The accuracy of a theodolite is high if it is properly used; therefore field procedures have been issued. Horizontal axis error, collimation error, and index error are regularly determined by calibration and are removed by mechanical adjustment at the factory in case they grow overly large. Their existence is taken into account in the choice of measurement procedure in order to eliminate their effect on the measurement results. A few other possible sources for errors are:

Figure 4.1 Theodolite

- A clear line of sight between the instrument and the measured points.
- The precision of the instrument is dependent on the raw repeatability of the angle measurement.
- A well defined measurement point or target/prism is required to obtain the maximum accuracy. This is mostly obtained by a brightly colored gauging-rod.

Assessment

Spread: The systematic error depends on the theodolite model, but according to different manufacturers the error in the measure angles is between +/-0.8" and +/-10" (Qualitest International Inc., 2009). There is also an error when people use the theodolite, and that is about +/-1".

Empirical quality: Even tough measured outside a laboratory, the field experiments are controlled have enough direct measurements. According to NUSAP the Pedigree score would be between 3 "Historical/field data, uncontrolled experiment, small sample direct measurements" "and 4 "Controlled experiments and large sample direct measurements ", so final Pedigree score is 3.5.

Methodical quality: The theodolite can be a very accurate method, if it is used according to field procedures. With assumption that people are dealing with the theodolite in a professional way the following NUSAP assessment is made: 3 "Reliable method, common within discipline".

4.2.1.2 Total Digital Station

A total station is an electronic/optical instrument used in modern surveying. The total station is an electronic theodolite (transit) integrated with an electronic distance meter (EDM) to read distances from the instrument to a particular spatial entity. Some models include internal electronic data storage to record distance, horizontal angle, and vertical angle measured, while other models are equipped to write these measurements to an external data collector, which is a hand-held computer. Most modern total station instruments measure angles by means of electro-optical scanning of extremely precise digital bar-codes etched on rotating glass cylinders or discs within the instrument. The best quality total stations are capable of measuring angles to 0.5 arc-second. Inexpensive "construction grade" total stations can generally measure angles to 5 or 10 arc-seconds. Measurement of distance is accomplished with a modulated microwave or infrared carrier signal. The typical total station can measure distances to about 3 millimeters. Because the Total Digital station is much similar to the theodolite, the same errors apply.

Figure 4.2 Total Digital Station in use

Assessment

Spread: The same error in the measure angles as the theodolite applies: between +/-0.8 arc-seconds and +/-10 arc-seconds, the error in the measured distances is about 3 millimeters, in a range of a few hundred meters (Qualitest International Inc., 2009).

Empirical quality: same as theodolite, so Pedigree score is 3.5

Methodical quality: also the same as theodolite, so Pedigree score is 3

4.2.1.3 Measurementship and GPS

For measurements of the river's wet-profile a boat is common used. The boat uses a depth sounder system in combination with Global Positioning System (GPS).The accuracy of the exact location depends on the accuracy of the GPS. According to a manufacturer of GPS systems, early GPS systems can have an error of several meters, but new GPS systems have errors of a few centimeters to 1 meter[1]. The density of the measurements with the depth sounder system is high, therefore accuracy becomes better.

Assessment

Spread: Assumed the GPS is of a newer model, the accuracy of the GPS will be between a few centimeters - 10 meter. The spread of the depth sounder will be very small.
Empirical quality: Because the accuracy and the number of points where depth is measured is both high, the Pedigree score will be "Controlled experiments and large sample direct measurements ", 4
Methodical quality: The method used for depth sounding is commonly used all over the world, but there is room for error. Therefore Pedigree score is between 3 and 4, so 3.5.

4.2.1.4 Ultrasonic depth sounder

Normal Doppler systems are not always usable in China due to high concentrations of sediments in the rivers; therefore an ultrasonic time-difference flow-meter has been developed. This method is evaluated in section 6.2.44

4.2.2 Water level measurement

Water level measurement is most commonly done by measuring the water surface elevation. The water surface elevation, referred to some arbitrary or predetermined gauge datum, is known as the gauge height. Gauge height is also used interchangeably with the more general term 'stage'. The gauge height is usually expressed in meters and hundredths of thousandths of meter if a more accuracy is required. The water-level is used for the determination of the stage-discharge relation. "The uncertainty in the stage-discharge relation depends largely on the uncertainty in the water-level measurement. It can be stated that, in methods of streamflow measurement where a correlation is established between stage, fall or slope and discharge, the uncertainty in the measurement of stage has a significant effect on the overall uncertainty in the record of discharge." (Herschy, 2008, p. 20).
The water level can be recorded by observation from staff gauges or continual and automated with water level recorders.

[1] Garmin Ltd. (2009). *What is GPS?*. Retrieved June 30, 2009, from Garmin: http://www8.garmin.com/aboutGPS/

4.2.2.1 *Staff gauge*

The non-recording reference gauge is the basic instrument for the measurement of water-level. The staff gauge is used for flow measurement site where only incidental observations are made or sometimes for regular used sites, where other water-level gauges are not available or usable. The gauge can be used as a control instrument for the normal water-level recorder. The disadvantages of a staff gauge is the need for an observer and because of that the loss of accuracy. The accuracy is also less than continuous recording gauges because fewer observations are made during the day. The change that the exact peak of a discharge wave is measured is very small, therefore corrections should be made, which give more uncertainty. Most staff gauges have standard designs, like the one present in figure 4.4. A staff gauge is not a stable construction. The gauge is often exposed to movement or damage, especially during floods. The gauge has to be verified and corrected regularly.

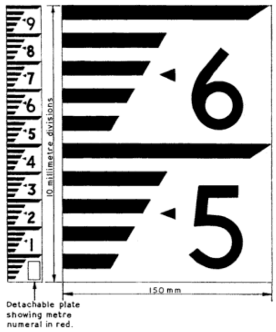


Figure 4.3 Design of staff gauge

Figure 4.4 inclined staff gauge, Yangtze River

A special gauge is the inclined gauge. As the name suggests, multiple staff gauges are placed on a riverbank. The multiple gauges provide more accurate readings if the bank has variations in its slope (figure 4.5). Assumed is that the staff gauge is properly installed, so that height is according to Chinese standards, otherwise a systematic error occurs. Another systematic error occurs when the staff gauge is installed in a curve in the river. The energy level of the water will be higher in the outside of the curve. When installing the staff gauge this has to be taken into account.

Assessment

The judgment of Jansen (2007) will be used for the assessment of the staff gauge.

Spread: The reader may be mistaken when reading the staff gauge in bad weather conditions, therefore a maximum error of 3cm = about 0.005% if average depth is 600 cm is assumed.

Empirical quality: The readout is direct, there are no further calculations needed for the readout of the staff gauge. When reading the staff gauge multiple gauge heights are recorded. The quality of the gauge heights is discussable. But a trained eye will be able to make accurate estimates. The Pedigree score will be between “controlled experiment and large sample direct measurements”, 4 and “historical/field data, uncontrolled experiments, small sample direct measurements”, 3. Pedigree score will be 3.5.

Methodical quality: because of the simplicity of the readout there is no real methodology. The staff gauge is common in the hydrology. Therefore the following NUSAP Pedigree scores are given: The staff gauge is 3, ‘reliable method, common within discipline’.

4.2.2.2 *Water level recorders*

The principle of the stilling well with a water level recorder (float-type recorder) was developed in the first half of the nineteenth century. But the water level recorders were installed around 1980 for the first time. The purpose of the stilling well is to dampen water level fluctuation and protect the float sensor components. The water level is registered with the use of an automated recorder actuated by a float within a stilling well. The floater is attached to a recording mechanism (such as a pen) which can produce either analogue or digital output. There are two types of analogue recorders: strip chart recorders and drum recorders. A clock movement controls the rate at which a strip chart advances. Most strip chart recorders will operate for several month without servicing, drum recorders weekly or monthly checking. Digital water level recorders have the advantage that they

have the ability to record and store information in database ready digital format. Digital data can be transmitted directly to a monitoring centre, so the data is directly available.
Sometimes the floater is not directly attached to a pen, but reduction gearing is used, common scales are 1:2.5 1:5 1:10 1:20. During peak discharge the water level changes rapidly multiple times, still the water level is registered continuously. The advantage of the water level recorder is the constant recording of the river's stage.

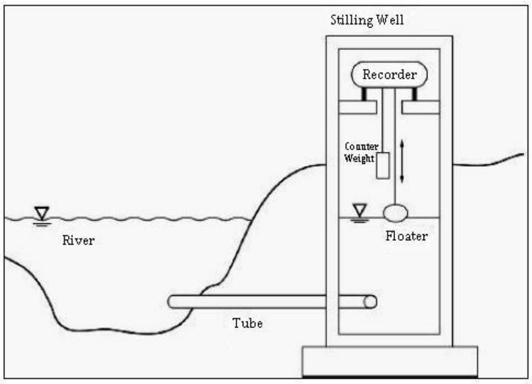


Figure 4.5 Schematic of stilling well

In any recording system using a float device to sense water level is a possible source for errors, resulting in uncertainty between the measured water level and the actual water level. A study done by Herschy (2008) showed that the measurement uncertainty of the water level recorder is about 1 cm. The high accuracy of the recorder can be explained by the decreasing effect of tube at the bottom of the stilling well on waves in the river. The water level is measured without disturbance from outside elements. Errors in the results of the water level recorder can be caused by various combinations of faults, such as: a change in the initial setting of the floater, friction in the mechanism, build-up of silt in the well on the float pulley. The result of these errors is a decrease in the response of the float to water level changes. The errors can be systematic and they are in particular important during low water, because small changes in water level occur more in this period. Small errors in water level measurement during low water have relative large errors in discharge. Other systematic errors in the water level recorder can occur because of a mistake during installation of the recorder, so that the calibration of the recorder is incorrect. The placement of the recorder is assumed to be without mistakes, so that the systematic error of the installation will be insignificant.

As stated before sediment can have influence on the water level measurement with a stilling well. The well can get silted up which can even results in uselessness of the stilling well because the floater cannot move. The status of the silting up of sediment in the stilling well has to be checked on a regular base. An other option for this problem is the use of a slope lifting water level measurement well, as used in the Yangtze River.

Assessment
The assessment of the analogue and digital recorders will be done separately.
Spread: Both the strip chart recorders and the digital recorders have a spread of 1 cm, according to research from Herschy (2008).
Empirical quality: The analogue water level recorder with strip chart is unmanned, but the water level is measured constantly. So the Pedigree score for the analogue recorder is between 3 and 4, so 3.5. The digital recorder can send its recorder water levels directly to a control center. Therefore the measurements are directly controlled, so for the digital recorder with data transmission a Pedigree score of 4 is given.
Methodical quality: Water level recorders are common in the hydrology. If the stilling well is maintained in a good condition (so no silt) the pedigree score for both the analogue and the digital water level recorder is the highest score 4, "Approved standard in well-established discipline"

4.2.2.3 Slope shifting water level measurement well

The slope lifting stilling well is used in case of high sediment transport by a river. The stilling well is built on a sloping track and the water-level recorder can be drawn by a winch (Keijang, 1993). So, the gauge can move with the changing water level. The flow into the gauge well is from the bottom of the river and, because of the small inlet and hydrostatic pressure, sediment deposition in the well is avoided. The system is suitable for recording changes in water-level and eliminates the error of water-level measurement resulting from the sediment transport in the river. Uncertainties in this method are similar to the water level recorder in the previous section. Main difference is reducing of sediment influence on the measurements. But this stilling well is more subjected to errors, because it has to be managed more carefully by people: they have to keep an eye on the water level, and if necessarily raise or lower the stilling well. If this is not done properly, the water in the stilling well may become out of range for the floater to record its water level.

Assessment
Spread: the spread for this stilling well is the same as the water level recorder: 1 cm.
Empirical quality: The recorded gauge height is not directly controlled, but someone has to be near the stilling well in case the water level recorder has to be winched. The Pedigree score will therefore be between 3 and 4, so 3.5.
Methodical quality: The slope shifting gauge is not common in hydrology; there is not much literature available. Therefore a Pedigree score of 3 is given.

4.2.3 Slope-area method

The slope-area method is one of the oldest methods to determine velocity. The method is only used to calculate the velocity after a flood has passed. The slope-area method will mostly be used as back-up for other instruments. The method is not used in Han River. The Chézy formula is used to calculate the velocity or discharge in a river and is as follows: $Q = dwC\sqrt{hi}$. Where:
Q= discharge [m^3/s]
C= Chézy coefficient [$m^{1/2}/s$]
d= water depth of river [m]
w= width of river [m]
i = bottom slope of the river [m/m]
By reading the staff gauges a peak discharge wave can be followed. In this way the distance and speed of the peak discharge wave is computed. An estimate for the roughness of the riverbed has to be made, this is an error source. The method is relative inaccurate, because actual measurements are missing.

Assessment
Spread: Herschy (2008) has computed the percentage of uncertainty in the discharge when using the slope-area method. This will be about 10-20% with 95% confidence with a single measurement of discharge.
Empirical quality: The data used for the Chézy formula are mostly direct (velocity, water level) or indirect measurements (discharge). But there is also a roughness coefficient C in the formula. The roughness coefficient is estimated, so this will be an educated guess. Therefore NUSAP gives a Pedigree score between 1, “Educated guesses, very indirect approximations, “rule of thumbs” estimates” and 2, “Modeled data, indirect measurements, handbook estimates”, so 1.5.
Methodical quality: The slope-area method is well known, but there is also consensus that the method has great errors and should therefore not be used for flood risk calculations. Because of that the Pedigree score will be 1, “Unproven methods, questionable reliability”.

4.2.4 Velocity-area methods

The principle of this method consists of determining velocity and cross-sectional area and is also known as the ‘reduced point method’. On the measurement site the width is measured. The depth is measured at a number of points (known as verticals) across the width. The amount of verticals needed depends on different things, like river width, river profile and estimated discharge. ISO 748:2007 provides standards for the use of the Reduced Point Methods. The ISO 748:2007 poses the following: “for channel widths > 5 m, the number of verticals shall be chosen so that the discharge in each segment is less than 5% of the total, insofar as possible, and that in no case should exceed 10%” (ISO, 2007, p. 6). At Baihe station 50 verticals are measured. More verticals won’t give more accuracy; the accuracy gain is reversed logarithmic with the number of verticals. ISO (2007) gives tables with the uncertainty in the computed mean velocity due to the number of verticals.

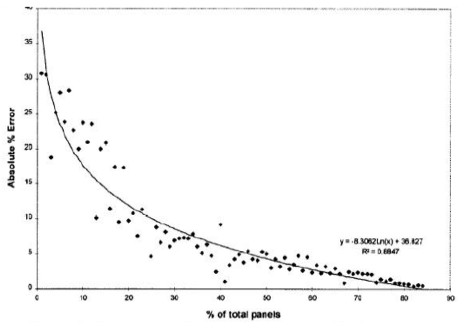


Figure 4.6 Increase in error from verticals reduction

Number of verticals	Uncertainties %
5	7,5
10	4,5
15	3,0
20	2,5
25	2,0
30	1,5
35	1,0
40	1,0
45	1,0

Table 4.2 Percentage uncertainties in the measurement if mean velocity due to the limited number of verticals (68% confidence)

The number of points taken in one vertical has an influence on the velocity uncertainty. If one point is measured the depth of that point will be at 0.6d, where d is measured from the bottom up. The idea is that the mean velocity in a river is at about this height, this is based on general current profile for open rivers as shown in figure 4.8 (ISO, 2007). More common is the use of two point depth method. Here the velocity is measured at two points 0.2d and 0.8 in a vertical. The average of the two values is taken as the mean velocity of the vertical. First the river profile has to be determined; this can be done by the methods described in section 6.2.1. The knowledge of the cross-section at the measurement site is used to determine the 0.2d and 0.8d depth from the riverbed up. More points taken in a vertical translates into better accuracy, but also more possibilities for mistakes. Therefore it is not necessarily to measure the velocity in 5 depth points in a vertical in small rivers. Another option is a continuous lower/raise traverse. This method records the velocity continuously while the current meter is moved bottom-surface or surface-bottom. Which method is used is not of great importance, because for every method has its (dis)advantages. At Baihe station the ‘reduced point method’ is used with 50 verticals and two measurement points at 0,2d en 0,8d of the depth below the surface. The measurements for this method can be made from bridges, cableways and boats. These methods will be discussed further down in the following sections.

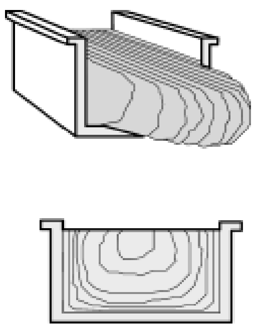


Figure 4.7 Typical current profiles and contours

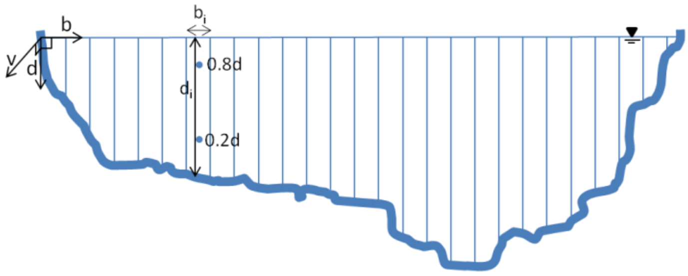


Figure 4.8 Diagram illustrating velocity-area method

When the width, depth and velocity in each vertical are known to total discharge in the cross-section can be computed. The flow is computed as follows: $Q = F \sum b_i d_i v_i$. F is an extra factor, because the measurement is done at a few points, while the ideal situation would be an infinite number of verticals, with infinite points taken at each vertical. Although mathematically this would be ideal, but the measurement time would be to large, resulting in more uncertainty, because velocity can change during session time. Equipment with more calculating capabilities and better sensors may change this in the future.

Information about the reduced point method is available since 1982 to 1985 mostly during high discharges in August and September. Also data about the situation in September 2006 is available. Data about the use of the reduced point method are only available between the month May and October. This means that river profile at low discharge in winter month is unknown. This means uncertainty about discharges in that period. But for frequency analysis only the peak discharges are relevant. A number of possible errors that may arise when using velocity-area methods (ISO, 2007, p. 9):

- If the flow is unsteady
- If material in suspension interferes with the performance of the current meter
- Is skew flow occurs, and the appropriate correction factors are not known accurately
- If the current-meter is used for measurement of velocity outside the range established by the calibration
- If the set-up for measurement (such as rods or cables suspending the current-meter, the boat, etc) is different from that used during the calibration of the current-meter, in which case a systematic error may be introduced.
- If there is significant disturbance of the water surface by wind
- If the current-meter is not held steadily in the correct place during the measurement, which is the case when gauging from a boat which is drifting, or when an oscillating transverse velocities gives rise to serious positive errors.

Some errors may occur due to environmental influences, which cannot be changed. Other errors may occur due to human factors. These last errors can be minimized with proper training of the personnel handling the measurement equipment.

Assessment

Spread: ISO 748:2007 provides a method to calculate the total uncertainty in the discharge when the measurements are taken and calculated using velocity-area methods. The calculation of the uncertainty for Baihe station can be found in Appendix C. Calculation according to the method described in ISO 748:2007 gives an uncertainty of 3% in discharge with 95% certainty at Baihe station.

Empirical quality: During normal discharges the velocity area methods are quite reliable, but during peak discharges and flooding the accuracy becomes less. The measurements are controlled, but are less reliable, because they are measured in field. Therefore the Pedigree score will be between 3 and four, so 3.5.

Methodical quality: The velocity area method is much described in different ISO standards. The 2-point measurement method is common accepted. Although a limited number of verticals is used, the number is greater than the ISO 748 recommends. Therefore the Pedigree score will be 4, “Approved standard in well-established discipline”.

4.2.4.1 Current meter

The current meter is the most universally used instrument for velocity measurements. The principle is based upon the relation between the speed of the water and the resulting velocity of the impellor. During a measurement session the current meter is placed at a point in the stream and the number of revolutions of the rotor is counted for a fixed time. The velocity of the water at that point can now be determined.

Problem at this moment is the inaccuracy of the current meter at low velocities, which gives greater uncertainties. There exist tables to help with dealing with this uncertainty (ISO, 2007). The minimal velocity required for most current meters (accept mini-meters) is normally about 0.03 m/s. The lower the minimum speed of response of current meters the lower the speed of flow which is measurable with confidence, always accepting that the uncertainty at this speed will be of the order of +/- 20%.

Calibration of the meter will be less consistent if the conditions at measuring site will defer from that at the calibration site, like another angle (of the cable) in which the current meter is used. The current meter is used in combination with the velocity-area method.

Assessment
Spread: Herschy (2008) determined an uncertainty of 5% (95% confidence) in current meter measurements. Arnold (2004) also finds an uncertainty of 5%.
Empirical quality: The current meter is used according to the velocity-area method. Therefore the same Pedigree score is given: 3.5
Methodical quality: The current meter is well known with multiple ISO standards, but the method is not infallible. Because of that Pedigree score between 3 "Reliable method, common within discipline "and 4 "Approved standard in well-established discipline"

4.2.4.2 Bridges

Although cableways are generally preferred to bridges for current meter measurements, highway or railway bridges are often used to advantage. Bridges do not so often offer the right conditions for stream gauging but measurements from them may be necessary where suitable sites for wading or for cableway are not present. However, contracted sections, piers and other obstructions have an effect on the current and it is therefore necessary to use a larger number of verticals as well as more velocity observations in each vertical, especially close to bridge piers and banks. Generally there are two types of bridge measurement, namely rod suspension and line suspension. An advantage of measuring from a bridge is that the natural flow of water is not disturbed when measuring as is often the case when measuring from a boat. Small streams are easy to measure from small bridges. Larger rivers have often higher bridges. This gives some extra problems. There are no rules for the selection of the upstream or downstream side of a bridge for discharge measurements. The advantage of upstream side of the bridge is that the hydraulic conditions on the upstream side of the bridge opening are usually more favorable. It is also possible to spot floating materials and avoiding them more easily (Herschy, 2008).

Assessment
Spread: The spread depends on the current meter and the velocity-area method. But in this case the uncertainty of the current meter, 5%, is used.
Empirical quality: The empirical quality depends on the current meter and the velocity-area method. But because bridges have a large influence on the current the data becomes less reliable, but the number of experiments is great, so therefore a Pedigree score between 3.5 is given.
Methodical quality: Bridge measurement has no standards (yet), for example choosing the upstream or downstream side of the bridge for measurement. On the other hand is it a common method and also reliable with extra precautions. So a Pedigree score of 3 is given.

4.2.4.3 Cableway

The cableway is not only used for the anchoring of boats as described in the section about current measurement, but also for lowering a lodging device. In smaller rivers and in the shallow regions of the river current and depth measurement can be conducted using the cableway. In rivers that are too wide to use a tag line a cableway can also be used. There are two basic types of cableway:

1. A cableway with only an instrument carriage controlled from the bank of the river with the use of a winch. The controlling is done either manually or electrically (preferred with larger rivers).
2. Another option is a cableway with a manned personnel carriage, also known as cable-car. The operator in the carriage travels across the stream to make the necessary measurements. This carriage can be operated either manually or electrically.

Figure 4.9 Example of a manned cableway

The general gauging procedure is similar to the measurements done from a boat. The cableway also provides information about both distance from river bank, depth of the river and velocity of current. The main advantage of a cableway compared to a boat is the fixed position. There will be no uncertainty repeatability of the measurement selection site. Of course the fixed position is also a

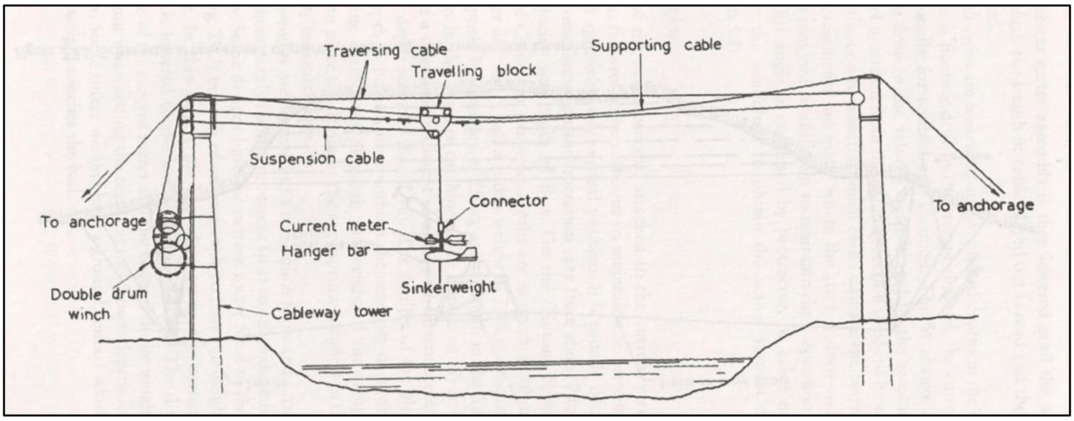


Figure 4.10 Schematic of unmanned cableway

disadvantage. During a flooding the span of the cableway might prove not to be large enough. The cableway has to be used with care. Several institutes have made guidelines for the measurement of velocities from a cableway. The most striking part is the care for floating drift during measurement sessions. This floating drift can have an effect on the measured velocities, because the current meter won't have a fixed position. Especially during peak discharges there is more floating drift in the water, this can cause errors in the measured velocities, without proper care for the floating drift. Under the current meter a sinkerweight is attached. The purpose of the sinkerweight is keeping the current meter in place. The line of which the current meter is attached should be as vertical as possible in the water; otherwise the current meter will not function properly, resulting in faulty velocity measurements. In rapid discharge changes of the river, the sinkerweight might prove not enough. Uncertainties may rise if a fixed position of the current meter cannot be assured.

Assessment

Spread: Just like bridges the uncertainty depends on the current meter attached to the cableway and the velocity-area method. Therefore the uncertainty will be assumed 5%.

Empirical quality: A cableway has much less influence on the current, unlike a bridge. But the sinkerweight needed for the current meter makes it less reliable. There is much written about cableway measurements, and corrections needed for the angle of measurement, wet line/dry line

corrections. Al those corrections are possible error sources. The number of experiments is great, so therefore a Pedigree score between 3.5 is given.
Methodical quality: There is no international standard, but a cableway is very common in China and within the hydrology. Therefore a Pedigree score of 3.5 is given.

4.2.4.4 Ultrasonic depth sounder

The Bureau of Hydrology in China has introduced a combined sediment sampler and current meter equipped with an ultrasonic depth sounder (figure 4.12). The facility if the ultrasonic transducer enables depth to be measured without necessity if having to make air-line and wet-line corrections (Cui, et al., 2008). The depth of the current meter is not calculated from the length of the line, which is the case with normal cableway measurements. Although the ultrasonic depth sounder has to be in the water at a constant depth and angle, so the weight of the sampler requires being sufficient to place the device at the appropriate location in the vertical to make the velocity measurements. In the Yangtze River an ultrasonic depth sounder is used with a sinkerweight of more than 200kg. The sinkerweight is placed under the depth sounder, so that the sinkerweight protects the sounder in case it would touch the bottom of the river. The depth sounder part fits in the cone of the current meter, so the sinkerweight gives no nuisance when measuring depth.

Figure 4.11 Combined sediment sampler, current meter, sinker weight and ultrasonic depth gauge

Assessment
Spread: There is no usable data about the accuracy of the ultrasonic depth sounder that is used in China; therefore the same uncertainty as a normal current meter, 5%, is assumed.
Empirical quality: The ultrasonic depth sounder is used in field measurements but gives large amounts of data, so the Pedigree score will be 3.5.
Methodical quality: The ultrasonic depth sounder used in China is developed by only for Chinese rivers. International approved standards are not available. Still it is a reliable method, because it combines already existing and proven methods. The ultrasonic depth sounder is common for the Yangtze and Yellow River basin. Therefore a Pedigree score of 3 is given.

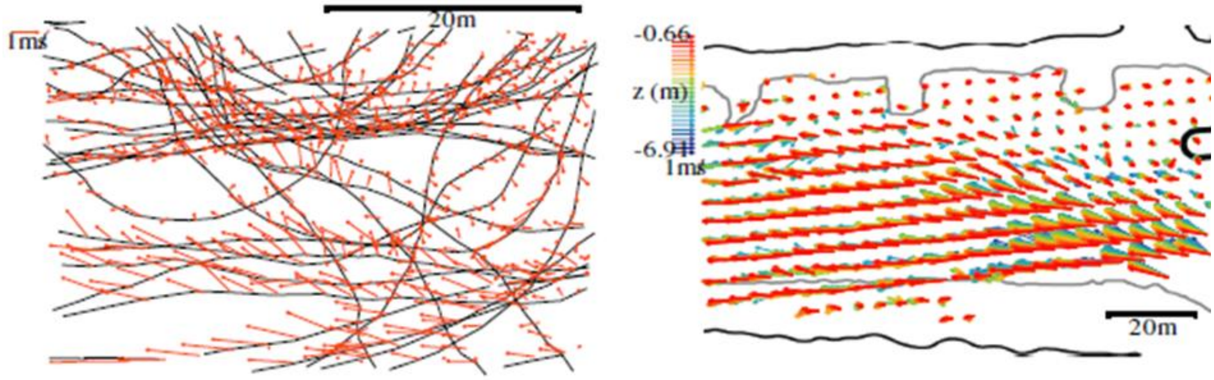
4.2.4.5 Boats

Where the river is too wide for cableway installation, discharge measurement are made from boats. One limiting factor in the use of boats is the high velocity of water, especially during floods, as safety of personnel on the boat has to be considered. Where the river is sufficient narrow to use a tag line, this cable is spanned across the river at the measuring section. The tag line serves two purposes: it anchors the measuring boat in positioning during measuring sessions; the line can also be used for measuring width of river and monitoring the verticals. The tag line is attached at the stern of the boat. During measurement sessions the line can be a problem for river traffic, so the line has to be lowered sometimes during sessions. As a result the tensions on the line during a single session can change; also the duration of the measurement becomes greater. Both consequences have an influence on the uncertainty in the measurement. Sometimes a tag line cannot be used; the boat can be kept in place by anchoring it to fixed point on the river bank further upstream. This positioning is less accurate than when a tagline is used, but the position of the boat is more accurate known than when the boat would be held in place solely by power from the boat self, without any connection to any riverbank. The hull of a boat has influence on the upper part of the current. In shallow water the influence can become relatively large, which influences the uncertainty in the measured velocities and even depth.

Assessment
Spread: When a current meter is used from a boat the uncertainty will be assumed the same as a normal current meter measurement, so 5%.
Empirical quality: The boat obviously has influence on the data, but the amount of data compensates so the Pedigree score will be 3.5.
Methodical quality: There are now international standards. This is also almost impossible, because different rivers need different kinds of measuring boats. Still, current measurement from boats is widely used, with much research and literature about the use, so a Pedigree score of 3.5 is given.

4.2.5 ADCP

The Acoustic Doppler Current Profiler (ADCP) was originally designed and manufactured in the 1970s especially for application in oceanography; research and investigation bringing the system successfully into the present day rivers application has been carried out since the early 1990s. The ADCP is not used at Baihe station in the Han River; however the ADCP is used in the Yangtze River. The ADCP is too important not to be evaluated, because in the future it may be used at Baihe stati on. The ADCP can be used to record current patterns in the river. The ADCP can be attached to a boat or it can be used from the riverbank, bridge or cableway. When the ADCP is used in combination with a boat the result can be plotted like figure 4.13a. After interpolation a velocity profile in the river can found, showed in figure 4.13b. The principle of calculating the discharge with this method is the same as the velocity-area method. The velocity in the river is measured at different depths and verticals, multiplied by the cross-section of each vertical gives the total discharge.



a) Randomly measured data b) Interpolation into regular grid

Figure 4.12 Readout of ADCP (Tsubaki & Fujita, 2007)

Measurement principle of the ADCP is the recording of the Doppler shift. The instrument sends out a pulse with a fixed frequency. The pulse is reflected by particles (sediment) in the water, the frequency of the pulse changes when it is reflected. The ADCP records the frequency change, because the magnitude of the frequency gives the velocity of the particles in the river. The ADCP sends out four beams, which gives the ADCP also the possibility to not only give the size of the velocity but also to give the direction of the velocity. Problem with the ADCP mounted on a boat is the visibility loss. The first meter from the surface down cannot be measured, because that’s the depth of the boat+ ADCP. The same reason prevents the boat from moving too close to the bank. The ADCP is sensitive for moving sediment on the bottom of the river. The lower part of the river can therefore also not be measured. A maximum of 30% of the rivers cross-section is not measured. Interpolation techniques come in hand to compute the immeasurable regions of the river.

There are a few advantages of the ADCP over velocity-area methods:

- An ADCP measurement session is normally quicker than existing current meter methods.
- The ADCP can be used in many sorts of rivers. The manufacturer says: 'The largest of the worlds rivers may be measured as well as the smallest'
- During flooding a river's width will expand, a cableway might not be usable any more. But with a boat with ADCP the river can still be measured quickly, which is important because of rapidly changing velocities in a flooded river.
- The ADCP can be used as a possibility to check existing velocity measurement methods, but this goes also the other way around.
- The ADCP is not only usable for measuring velocity, but it can also be used for sediment discharge measurement (because of the reflecting particles)

Uncertainties

Uncertainty calculation of an ADCP measurement of discharge is more complex than that of the conventional methods of streamflow measurement, but the principle is the same in that the velocity in each of *n* verticals is measured and the discharge estimated from the mid-section method. From experiments so far, it is estimated that the uncertainty of an ADCP measurement under good conditions is similar to the uncertainty of a current meter measurement. For a stationary ADCP measurement the estimation of the uncertainty is similar to that of the uncertainty in a current meter measurement with minor modifications. Generally, the error sources in an ADCP measurement consist essentially of the following (Herschy, 2008):

1. Spatial resolution (velocities estimated by monastatic diverging multibeam geometry)
2. Noise (may be large in low flows with high turbulence)
3. Velocity ambiguity (ADCP measures phase angle difference between pulses)
4. Side lobe interference (estimated by power curve fitting)
5. Temporal resolution (velocity data sampled as time series at equally spaced intervals)
6. Sound speed (ADCP assumes speed of sound and salinity constant)
7. Beam angle (like u1 is due to instrument tolerances
8. Boat speed (high ratios of boat-to-flow velocity may affect this error
9. Sampling time (may not be as critical for discharge measurements as it is for estimating mean velocity)
10. Near transducer(ringing waiting time-blanking period causing errors in velocities in upper bins)
11. Reference boat velocity (boat-mounted ADCPs measure in water column relative to boat movement)
12. Depth (transmit time for bottom tracking profiling and immersion depth of ADCP)
13. Cell positioning (maintaining to of first cell at constant position across section)
14. Rotation (pitch, roll, heading, attitude, and motion related to instrument configuration)
15. Time (needed to establish boat velocity and gating the return signal)
16. Edge (distance of ADCP from bank from assumed velocity distribution and discharge algorithm)
17. Vertical velocity profile model (depends on moving of fixed boat)
18. Discharge model (velocity area methods of estimation may be used in the algorithm)
19. Finite summation (as in velocity area the uncertainty in number of verticals may be taken)
20. Site selection and operation (secondary currents, aspect ratios, bed, turbulence, etc.)

The systematic error thus also depends on the cross-section where the ADCP is used. In this research the systematic error cannot be computed, because the ADCP is not used at Baihe station. Research and experience of providing values for each of the above uncertainty components is at an early stage but components 3,4,7,13 may be considered small or insignificant or may be included with other sources: the total uncertainty may then be estimated by the root-sum square method as follows

$u(Q) = \sqrt{u_1{}^2 + .. + u_{20}{}^2}$

The above component uncertainties are at the 68% confidence level and u(Q) is multiplied by two to present the results at the 95% confidence level. Values for the component uncertainties in the equation are estimated by both the user and manufacturer of the equipment.

Assessment

Spread: the spread depends on multiple error sources, as described above. Some of them are similar to normal current meter measurement with the velocity-area method. Research between measured discharges from a current meter and ADCP shows that ADCP gives consistently larger discharges than a current meter. The uncertainty of the ADCP will be about 5%.

Empirical quality: The density and controllability of the ADCP is higher than a standard current meter. But like current meter measurements assumptions have to be made about the velocities near the riverbanks. Therefore the Pedigree score is also 3.5.

Methodical quality: The APCP becomes more widely used ever since its introduction. In China the ADCP is used in larger rivers like the Yangtze River. But there are no standards for using the ADCP yet. Therefore the Pedigree score for methodical quality is 3.5.

4.3 Summary of NUSAP assessment

Different measurement methods and instruments have been used since the beginning of recording of water level changes at Baihe station. The first water level recordings were carried out by small wooden boats. A theodolite was used to map the rivers banks during low tide. The staff gauge was also introduced in the early periods. The slope-area method was used for calculating the discharge. Later the velocity-area method was introduced for more accuracy. The measurement instruments also gradually improved. Wooden boats became steel motorboats around 1960. The staff gauge was replaced by a water level recorder for continuous discharge measurements around 1980, the theodolite by a total digital system around 1990. Most of these changes lead to better accuracy, not only because of better instruments, but also by shortening the measurement time. The latest improvement is the use of ADCP. This system reduces the measurement time and difficulty, by eliminating human influence in navigation and discharge calculations.

Assessment of uncertainties in determination of cross-section

Uncertainties	Gauging-rods and Theodolite	Total digital station	Measurement ship and GPS
Spread	0.8”-10”	0.8”-10”	3-60 cm
Empirical quality	3,5	3,5	4
Methodical quality	3	3	3,5

Assessment of uncertainties in water level measurement

Uncertainties	Staff gauge	Water level recorder (analogue)	Water level recorder (digital)	Slope shifting water level measurement well
Spread	3 cm	1 cm	1 cm	1 cm
Empirical quality	3,5	3,5	4	3,5
Methodical quality	3	4	4	3 (miss. 4)

Assessment of uncertainties in velocity-area methods

Uncertainties	Velocity-Area method	Current meter	Bridges	Cableway	Ultra sonic depth sounder	Boats
Spread	3%	5%	5%	5%	5%	5%
Empirical quality	3,5	3,5	3,5	3,5	3,5	3,5
Methodical quality	4	3,5	3	3,5	3	3,5

Assessment of uncertainties in other methods

Uncertainties	Slope-area methods	ADCP
Spread	10-20%	5%
Empirical quality	1,5	3,5
Methodical quality	1	3,5

5 Time series peak discharges

5.1 Introduction

This chapter will give information about different techniques to evaluate the time series. This chapter will also evaluate the time series of the Baihe discharge station, China. The assessment is focused on the different methods that are available to create a series of peak discharges that will be used as an input for the statistical analysis.

5.2 Annual Maximum versus Peaks over Threshold

In a frequency analysis the data that is used as an input for the analysis has to be selected carefully, because the data has a great influence on the distribution of the discharges (see also chapter 6). In general there are two ways of selecting the discharges that are used for the analysis: Annual maxima model (AM) and the peaks over threshold model (POT), also called partial duration series approach (PDS). An AM data series is constructed by extracting from a series of flows the maximum value of each year (annual flood), so only one event per year is used. The POT is based on the selection of retaining all peak values that are above a certain “base level” or “threshold”. Hence, the POT approach is not limited to only one event per year. The main advantage of POT modeling is that it allows for a more rational selection of events to be considered as peak discharges. Unlike the AM modeling, which includes only one event per year, the POT approach provides the possibility to control the number of flood occurrences to be included in the analysis by an appropriate selection of the threshold. In fact, some annual floods may not even be selected as flood events in the POT approach (Lang, Ouarda, & Bobée, 1999).

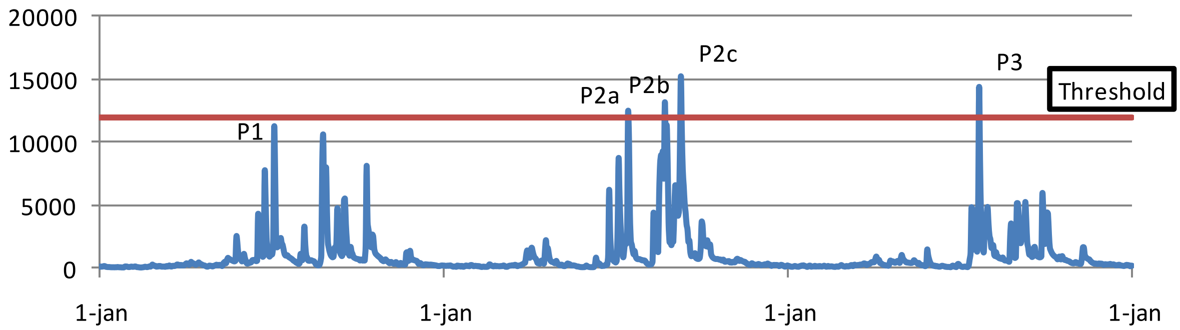


Figure 5.1 Annual maximum and peaks over threshold

The selection of discharge peaks is visualized with figure 5.1. With annual maximum selection the peaks P1, P2c and P3 will be selected, with POT the peaks P2a, P2b, P2c and P3 will be selected. By considering peak events instead of yearly maxima, the number of datapoints for statistical processing may be increased considerably. However, excessive lowering of the threshold to obtain more data may lead to substantial bias (Lang, Ouarda, & Bobée, 1999). However, the additional flexibility of the POT approach is often associated with an additional statistical complexity. Furthermore, there are no guidelines for the application of the POT, there are also some unsolved questions concerning the various details of the approach (Lang, Ouarda, & Bobée, 1999).

The choice of threshold and selection of criteria for retaining flood peaks are two elements that are of great importance since they are crucial for the independence and distribution of discharge peaks. In contrast, the AM approach is based on the selection of the largest discharge for each year of the record, which naturally leads to discharge peaks that are generally identically distributed.

Two different approaches can be used for threshold selection: the first one is based on physical criteria such as the identification of the flood level (overflowing) for a specific river, and the second one is based on purely mathematical and statistical considerations. Intuitively, and based on mathematical considerations, threshold values should be selected high enough to meet the basic model hypothesis, namely peak exceedances should be independent, and the occurrence process should be described by a Poisson process (Lang, Ouarda, & Bobée, 1999). A variation on the first approach for POT selection is a threshold set on the discharge on which the dams and weirs in the river open their overflow, which normally means the discharge in the river is larger than usual.

Give below are the two different time series. One is the annual peak discharges (figure 5.2) and the other is the time series which is the result of using the peaks over threshold method (figure 5.3) with a threshold of 12.000 m^3/s. This threshold is chosen because the number of peaks will also be 66, just like with the annual maximum method. Because there is no information available about critical thresholds for overflowing in surrounding dams and weirs in the Han river, this criteria is not used in selecting a level for the threshold.

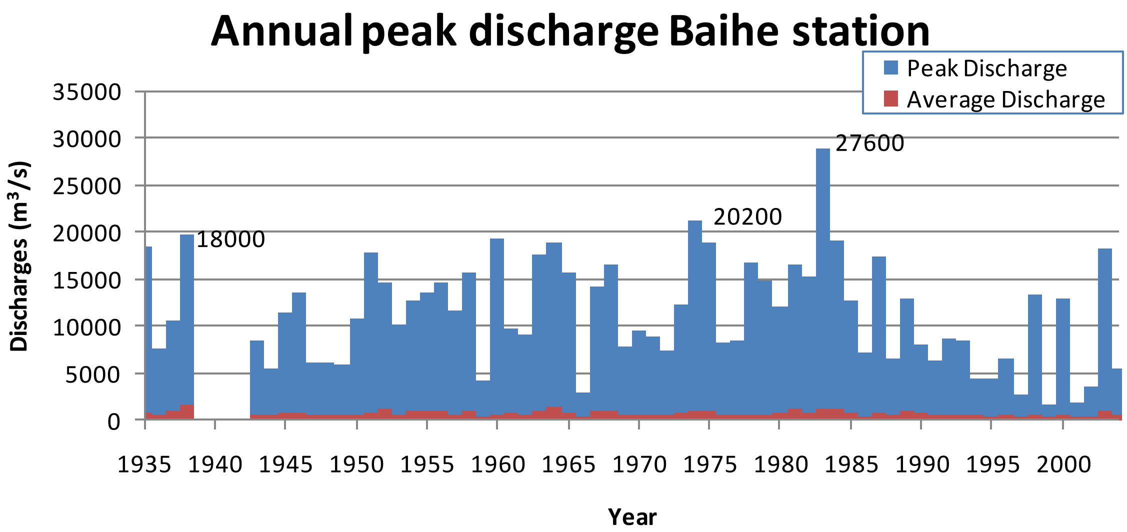


Figure 5.2 Annual peak discharge Baihe station

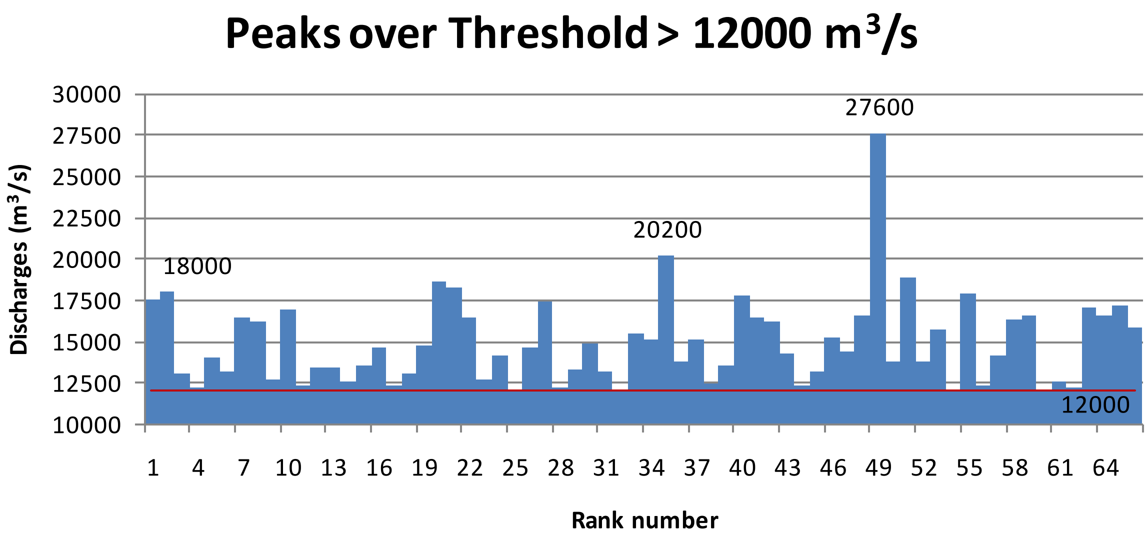


Figure 5.3 Peaks over threshold

Discharges in the years 1939-1942 are not measured, also the years 1948-1948 have gaps in discharge data. In the rest of this research the AM method will be used, because of its simplicity and proven usability. The annual maximum discharges can be found in appendix B.

The location of Baihe station has moved in 1950 and 1957. One might argue that, when looking at figure 5.2, a significant change in measured flows can be seen between 1950 and 1957. This research does not further inquiry. If the change is found significant, than the data for these years might have to be adjusted. This falls under the subject of stationarity. The discharges in this research are not analyzed for stationarity and independency, because there is no information about the different factors (see chapter 1.2) that have an influence on the stationarity at Baihe station. More information about this can be found in chapter 6.

5.3 Trends in time series

In order to make trends visible in the time series the measured discharges and the predicted discharges (with regression analysis) are plotted. When using annual maximum peak discharges the discharges are decreasing over the years. When using POT12000 the trendline is slowly increasing. The R square in both trend plots is very small, 0,03 for AM and 0,02 for POT. Therefore the regression is not significant, but the best fit suggests that a conclusion that the overall discharges are decreasing, but that when a peak discharge occurs, the discharge is larger than in the past. It is also arguable that because the data has no significant trend, the data can be treated as stationary (Rao & Hamed, 2000).

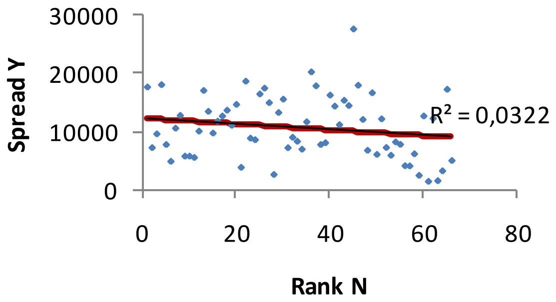


Figure 5.4 Trend in peak discharges (AM)

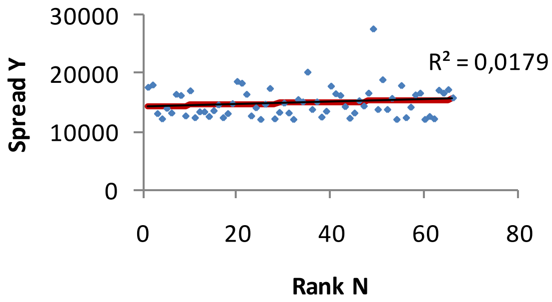


Figure 5.5 Trend in peak discharges (POT)

5.4 NUSAP assessment

Empirical: The data that are used for the time series are measured discharges, so Pedigree category is historical / field data. But also indirect measurements and sometimes with aid of handbook estimates because of the derivation of discharges from velocity/area method, thus Pedigree score between 2 and 3, so 2,5.

Methodical: Because non-stationary data is used, the reliability is questionable. The usage of data without screening them is not accepted in the hydrology discipline. Therefore the Pedigree score is 1.

6 Statistics

6.1 Introduction

The statistics is what mostly is referred to as the frequency analysis. The measured discharges from chapter 4 have reviewed in chapter 5 so that there is now one data series, the Annual Maximum, which will be used as an input for the frequency analysis. But before that the discharge data will be reviewed for any outliers, because outliers might have an influence on the distribution of the discharges. After the time series is tested it is possible to fit different distributions to the discharge data. The fitting is done by estimating the different parameters in the distribution functions. This chapter will assess the fitting of the different distribution models.

According to Rao & Hamed (2000) there are two types of error associated with return period estimations. The first error type is involved in the assumption that the observed discharge data follow a particular distribution. This error can be checked with a goodness-of-fit test, this will be discussed in section 8.4. The second error source is inherent in parameters estimated from small samples, in other words, there is not enough discharge data (yet) measured to estimate a valid distribution function. This error can be reduced by using a method which gives minimum variance parameter estimates, which in turn would result in the smallest variance in the expected return periods for discharge. It is then possible to construct confidence intervals for the estimated discharges by using information about sampling variance of the parameter estimates. The confidence intervals are expressed as the standard error. For every distribution function with different parameter estimates it is needed to determine the standard error. This is because the standard error does not take the goodness-of-fit of a distribution into account. The most efficient method is that which gives the smallest standard error of estimate. In this research the standard error is determined for each distribution function with different parameter estimation methods. The calculation methods and results can be found in appendix E.

6.2 Hydrologic input data

Two basic assumptions in statistical flood frequency analysis are the independence and stationarity if the data series (Rao & Hamed, 2000). In addition, the assumption that the data come from the same distribution (homogeneity) is made. In this research the stationarity of the data cannot be guaranteed, because there is no information about chances of the Baihe basin that have an influence on the stationarity. Possible influences are given in chapter 1.2. Therefore all computations are carried out with the measured data.

6.2.1 Tests for outliers

An outlier is an observation that stands out significantly from the rest of the data series. The outlier may be a result of errors in the data collection or may be caused by natural causes. The presence of outliers in the data causes difficulties when fitting a distribution to the data. Low and high outliers are both likely to occur, and they have both different effects on the analysis.
Outliers may be detected from a plot of e_i versus $\widehat{Y}_i$, also known as a residual plot (Rao & Hamed, 2000). The criterion to test whether an observation is an outlier is whether it has a residual value greater than multiple times the standard deviation. The residual value is the difference between predicted and observed values. Variance in discharge data is likely to be high, the whole 95% range is used. Sometimes outliers of more than 7,5 times the standard deviation can be found (McCormick and Rao, 1995). Other tests are also possible to detect outliers, like the Grubbs and Beck (1972) (G-B) test or the least median square (LMS) method.

For this test the peak discharge data (appendix B) are used. The residual plot is shown below. The black lines represent the 2σ bench mark.

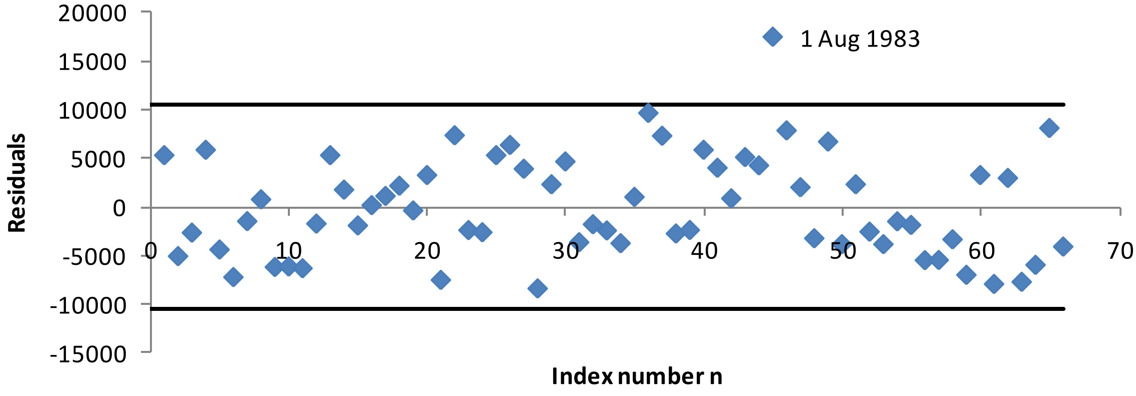


Figure 6.1 Peak discharge Residual Plot with outlier

One peak stands out. This peak is more than 3 times the standard deviation above the mean. The peak corresponds to the discharge of 27600 m^3/s on august 1st 1983. At first glace this peak should be removed in order for a better fit for the distribution functions. But research on discharge stations nearby Baihe station tells us that on that day there was an extremely high tide in the river, the discharge doubled within 24 hours. Other discharge stations recorded similar proportionally discharges, so proportions between the different discharge stations remained constant. The reason this flood occurred was because of heavy rainfall: 600mm within 10 hours. The water level rose with multiple meters above normal level. The duration of the peak was about five days, as can be seen in figure 6.2. There is no data available on hour basis, but an analysis of this discharge peak is not relevant for this research. The important thing to know is that the outlier is valid and thus will be used in the return period analysis and for the fitting of the distributions.

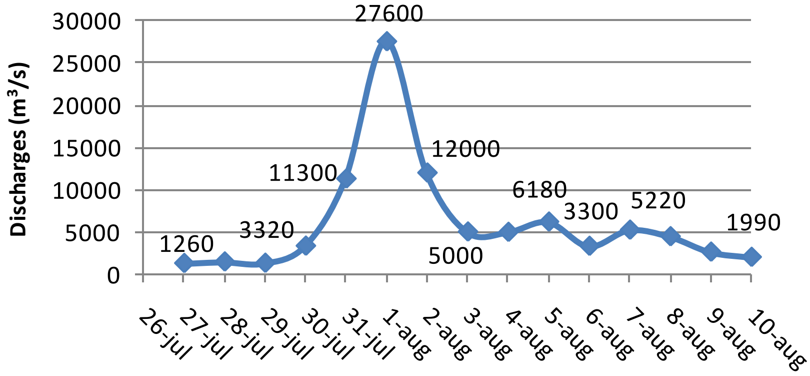


Figure 6.2 Discharge peak at August 1st 1983

6.3 Distribution functions

Flood peaks do not occur with any fixed pattern in time or magnitude. Time intervals between floods vary. The definition of return period is "*the average of these inter-event times between flood events*" (Rao & Hamed, 2000, p. 6). Large floods naturally have large return periods and vice versa. The definition of the return period has especially no remark about probability of the occurrence of peak discharge or floods. But probability is widely used in hydrology for the computing of the return period, this can be explained. A given discharge Q with return period T may be exceeded once in T years. Thus the probability of exceedence is $P(Q_T>Q)= 1/T$. Normally a cumulative probability of non-exceedence, $F(Q_T)$, is used. The relationship with T is: $F=1-1/T$. For example if T = 100 years then, Probability of non-exceedence $F=1-1/T = 1-0{,}01=0{,}99$.

There are different kinds of distribution types, the most common distribution families used for return period calculations of discharges are: Normal distributions, the Gamma family and Extreme value distributions. Other distributions are Wakeby and Logistic distributions. In this research the distribution functions for assessment are selected so that each group is assessed. The distribution functions that will be assessed are:

- Normal (Normal distributions)
- Exponential (Gamma family), with peaks over threshold
- Pearson III (Gamma family)
- Gumbel or Extreme Value Type I EV(1) (Extreme value distributions)

The distribution functions have to be fitted to the discharge data. This fitting is done by estimation different parameters in the distribution functions; this is explained in the next section. The exponential distribution is not usable with annual maximum time series (Jansen, 2007). The fit of the distribution is largely depended on the x_{min}, the lowest discharge in the discharge series. With the AM method the x_{min} is 1330 m^3/s and the mean is 10646 m^3/s. The difference between the x_{min} and the mean is too large. The POT would give a better fit, because the x_{min} would be above the threshold, and thus closer to the mean of the POT, this method is commonly accepted (Rao & Hamed, 2000). Therefore the exponential distribution is used with the POT series; the threshold is 12.000 m^3/s, this threshold was determined in chapter 5.2.

After the parameters are estimated, the fit of the distribution function in the discharge data is evaluated. This evaluation can be done in two main ways. The first is purely mathematical, the second a combinations of mathematical and visual estimations. The assessment of the goodness-of-fit of probability distributions for flood frequency analysis is done by using chi-square, Kolmogorov-Smirnoff tests (KS-test) (Kite, 1977). Studies done by Moon et al. (1993), Turkman (1985) and others show no real best probability distribution for all data series. In the chi-square test, data are first divided into class intervals. In each class interval the number of events and the expected number of events that exceed a probability are computed. The KS-test is based on the deviations of the sample distribution function F(x) from the completely specified continuous hypothetical distribution $F_0(x)$ (Rao & Hamed, 2000, p. 43). The KS-test also uses class intervals, like the chi-square test.
Other goodness-of-fit measures are also possible, such as the least-square test (Kite, 1977) and the probability plot correlation coefficient test (Filliben, 1989). Rao & Hamed (2000) note that goodness-of-fit tests have very low statistical power. Since the parameters of the tested distributions are estimated from samples, it follows that several candidate distributions may be considered to be similair. Consequently there is a very high probability that real differences will not be detected by these tests. The results of the different goodness-of-fit tests are likely to have high variations. Arora and Rao (1985) did multiple goodness-of-fit tests to see how well different distributions would fit a data series, the conclude that even in the light of the weak tests, distributions are not acceptable in many cases. Consequently a single distribution is not acceptable for all the data.

The goodness-of-fit can also be estimated visually. In this case the distributions are plotted against discharge plotted with plotting positions. This plotting is done with using yet another formula. Commonly used plotting formulas are Weibull, Gringorton, Hazen, Blom and others. Research on plotting positions has had a long history and the work is still continuing (Rao & Hamed, 2000). In this research only the Weibull and Gringorton plotting-postions are used, because they are the most common types. In the past the visual estimation was done by drawing the different plot positions on logarithmic plotpaper. A line was drawn trought these plotting positions. As earlier shown, plotting poitions vary with different types of plotting positions funtions. So these plotting positions had an influence on the extrapolation of the peak discharges. Today computerprograms, like MS Excel, are available to draw the

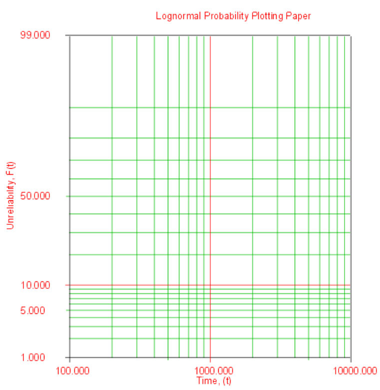


Figure 6.3 Example of lognormal probability plotting paper

distributions. Additionaly the use of computerprograms gives more accurate visual fittings, because they do no longer have to depend on the plot positions. The plot positions are only used for a visual confirmation for the selection and fitting of the right distribution function. The distribution is fitted with the parameter estimation.

For this research different distribution functions, as mentioned earlier, have been fitted to the annual maximum peak discharges. The exponential distribution is used with 'peaks over threshold', with a threshold level of 12.000 m^3/s. The choice for this threshold is explained in chapter 5.2. The exact method and calculations can be found in appendix E The resulting distributions are plotted with two types of plot positioning methods, Gringorten and Weibull. The result is shown in figure 6.4. A larger plot can be found in appendix E. In the next section the different parameter estimations will be discussed in more detail.

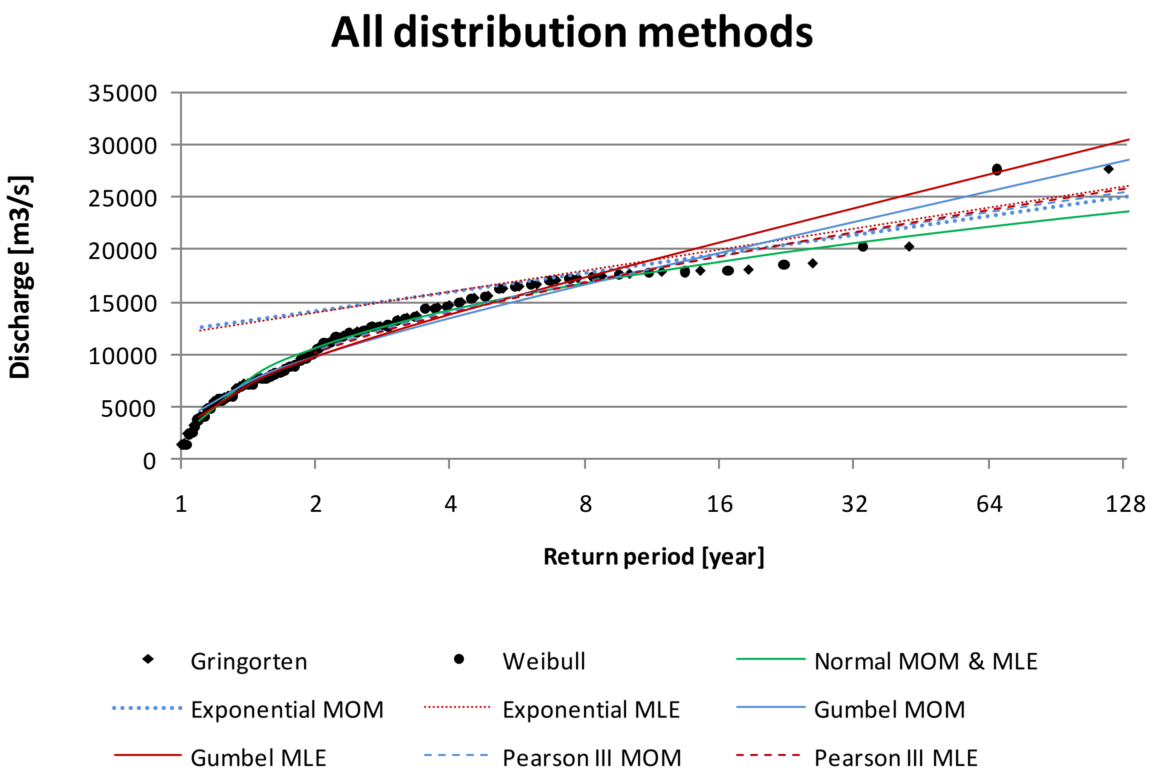


Figure 6.4 Results of different distribution methods and discharges plotted with plot positioning functions

6.3.1 Parameter estimation

After several distributions are selected to fit the data, their parameters must be estimated. Some of the common methods will be discussed in this chapter. The chosen distributions for this research are: Normal, Exponential, Gumbel and Pearson III. In this chapter the estimation of the parameters in the different distribution functions will be presented. The calculations can be found in Appendix E. There are different methods that can be used for parameter estimation. A small list: Method of moments (MOM), maximum likelihood estimation (MLE), probability weighted moments method (PWM), least squares method (LS), maximum entropy (ENT), mixed moments (MIX), generalized method of moment (GMM) and incomplete means method (ICM).

The method of moments is a relative easy parameter estimation method. Because of its simplicity, the estimates are of inferior quality. Distributions with three or more parameters that have to be estimated are more likely to have biases, especially in combination with smaller data series.

The maximum likelihood estimation method is considered the most efficient method compared to other methods (Rao & Hamed, 2000). Since it provides the smallest sampling variance of the estimated parameters, and thus also the smallest variance in the estimated return period for discharges. However, in some particular cases, such as the Pearson type III distribution the optimality of the MLE is only asymptotic and small sample estimates may lead to estimates of inferior quality (Rao & Hamed, 2000). Also the MLE is known to give often biased estimates, although these can be corrected, but it gives larger uncertainty. Furthermore, it may not be possible to get MLE with small data series, especially if there are a lot of parameters (like Pearson III).

In China the MLE method is commonly used. In this research two methods, the Method of Moments (MOM) and Maximum Likelihood Estimation (MLE), will be evaluated. The combination of MOM and MLE will be used, because the MLE is used in China and it can be compared with parameter estimates from the MOM to see which one would perform better, so with the least standard error. Also a Pedigree score for the usability and accuracy will be given to both methods. Special attention goes to the Pearson type III distribution, because this distribution is standard in Chinese distribution calculations. As stated before both parameter estimations will possibly have difficulties with the parameter estimates in the Pearson III distribution.

The following information about the discharge series is obtained, before the parameter estimation can start. Input was the normal – non-stationary discharge data – as given in Appendix B. This information is obtained with function 'Descriptive Statistics' in MS Excel. To the right a summary of the results is shown. The symbols are added for better understanding and reference for the parameter estimation calculations. The calculations and results can be found in Appendix E.

Discharges		*Symbols*
Mean	10646,52	$\hat{\alpha}_1 = m'_1 = \mu'_1$
Standard Deviation	5347,31	$m_2^{1/2}$
Sample Variance	28593715,36	μ_2
Kurtosis	0,0885	C_k
Skewness	0,4362	$C_s = \gamma_1$
Minimum	1330	X_{min}
Maximum	27600	X_{max}
Count	66	N
Confidence Level(95,0%)	1314,53	

Table 6.1 Results of Descriptive Statistics

6.3.2 Results and Conclusions

A summary of the discharges at T=100 and T=500 for different distribution methods can be found in table 6.2. The standard error in the determination of runoff with a specified recurrence time for different distributions and parameter estimation methods is given below in table 6.3, more detailed tables can be found in Appendix E.

	T_{100} (Q in m^3/s)	Relative Error	T_{500} (Q in m^3/s)	Relative error
Normal	23089	5,67	26039	5,92
Exponential MOM	24378	7,73	28593	9,15
Exponential MLE	25326	6,48	29964	7,40
Pearson III MOM	24783	9,46	28919	11,86
Pearson III MLE	25019	9,03	29400	10,87
Gumbel MOM	27419	10,51	34146	11,23
Gumbel MLE	29208	8,72	36613	9,07

Table 6.2 Discharges at T=100 for different distribution methods

Based on the results in the tables 6.2 and 6.3 can be concluded that the exponential distribution has the best fit, but this is not entirely right. The exponential distribution is left out the fitting of the lower discharges, because a threshold of 12.000 m^3/s is used, the discharges of a specific recurrence interval are based on the POT. Figure 6.4 shows that the normal distribution is best fitted to the lower discharges, but not the peak discharges with a return period of about 100 years. Therefore this distribution is not the best to use in determination of discharges with a return period >100 years. This fit can be explained by the slight S-curve of the measured flows (fig 6.4). Different distribution functions give different importance to lower or upper part of the S-curve. The Exponential distribution with MLE parameter estimation has a good fit for larger discharges with return period larger than 100 years although, Gumbel MLE distribution and Pearson III MLE distribution is also close. The main conclusion is that a best distribution method cannot be appointed without doubt.

Relative errors in Q_t (%)		Normal distribution		Exponential distribution		Pearson type III distribution		Gumbel distribution	
Recurrence interval T (Years)	Probability	MOM	MLE	MOM	MLE	MOM	MLE	MOM	MLE
10	0,10	5,24	5,24	4,65	4,38	6,12	6,28	8,70	7,89
20	0,05	5,36	5,36	5,75	5,15	6,99	7,05	9,39	8,19
50	0,02	5,54	5,54	6,97	5,97	8,38	8,18	10,09	8,52
100	0,01	5,67	5,67	7,73	6,48	9,46	9,03	10,51	8,72
200	0,005	5,78	5,78	8,40	6,91	10,52	9,85	10,85	8,89
500	0,002	5,92	5,92	9,15	7,40	11,86	10,87	11,23	9,07

Table 6.3 Standard errors in the determination of runoff with a specified recurrence time

6.4 Goodness-of-Fit Tests

The choice of distributions which can be used in flood frequency analysis can be a tricky one. A literature study from Rao & Hamed (2000, p. 41) concludes that "most of the methods available for selection of distributions from small samples are not sensitive enough to discriminate among ditributions." The selection of distributions is based on the goodness-of-fit of different distributions. The most common tests used for the selection of probability distribution functions are the Chi square test and the Kolmogorov-Smirnov test. The KS-test is used to test whether or not a given distribution is significantly different from one hypothesized. The test is a more powerful alternative to chi-square goodness-of-fit tests when its assumptions are met. The Chi-square test is used to test if the observed distribution is not significantly different from the hypothesized one. On the other hand, KS-test tests takes also the most deviant values of the criterion variable into account, because these values have a larger influence on some distributions, also it may be important for the propagation (Rao & Hamed, 2000). Thus, One-Sample Kolmogorov-Smirnov Goodness-of-Fit Test is a slightly better test. The D value in the KS-test is the largest absolute difference between the cumulative observed proportion and the cumulative proportion expected on the basis of the hypothesized distribution. The computed D_n is compared to a table of critical values of D of KS, for the given sample size. In this research both the X^2-test and the KS-test were used to evaluate the goodness-of-fit for the different distributions that are computed in section 8.3.

6.4.1 Results and conclusions

The results of both the X^2-test and the KS-test are shown below. The rank gives the order for best fit according to the specific test, with 1 the best fit and 'R' means the distribution is rejected at the given certainty range (95%).

	X^2 test			Kolmogorov-Smirnov			Total score (sum rank)
Distribution	Rank	X^2	$X_{0,95}{}^2$ significance level	Rank	D_n	$D_{critical,\,0,95}$	
Normal	1	4,30	8,34	2	0,0636	0,1674	3
Exponential MOM ($POT_{12.000}$)	R	10,62	8,34	6	0, 1310	0,1786	6R
Exponential MLE ($POT_{12.000}$)	R	11,88	8,34	4	0,0687	0,1674	4R
Pearson III MOM	3	5,82	8,34	1	0,0545	0,1674	4
Pearson III MLE	2	5,52	8,34	1	0,0545	0,1674	3
Gumbel MOM	4	8,32	8,34	5	0,1032	0,1727	9
Gumbel MLE	R	8,50	8,34	3	0,0688	0,1700	3R

Table 6.2 Results of goodness-of-fit tests

Based on the goodness-of-fit results the following conclusions can be drawn.

- The Normal distribution gives best fit according to X^2 test
- The Pearson III gives the best fit according to Kolmogorov-Smirnov test
- Both Normal and Pearson III MLE have the best total score according to both X^2 test and Kolmogorov-Smirnov test
- The fitting of the Exponential distribution tells us something about the fitting of the exponential distribution to the POT, not to the AM.

6.5 NUSAP assessment

The assessment of the distribution functions and their parameter estimations is done according to the NUSAP method. The Pedigree scores are given based on table 3.1. The criteria for the assessment of the uncertainties are based on studies of Rijkswaterstaat (2002) and have been used by Jansen (2007).

Statistical quality

Distribution type uncertainty

The distribution type uncertainty is used to deal with the possibility that two or more distribution functions have the same goodness-of-fit, but give different distribution results. The other way around is also possible, small differences in the distribution results, but large differences in the goodness-of-fit results. The different distribution functions give a reliable fit on the measured discharges, except the exponential distribution, because this distribution uses the peaks over threshold method, instead of the annual maximum. The conclusion from the computation of the distributions was that there it is not possible to point out one best distribution function. For the statistical quality the NUSAP Pedigree score will be 2.0, because the fitting tests do not give a definite answer to the best distribution fit.

Goodness-of-fit of distribution

This criterion tells something about the quality of the distribution function in describing the discharge data. In this research the goodness-of-fit is assessed by reviewing the results of the goodness-of-fit tests and the standard error in the discharges of the distribution function at a specified recurrence time. When the standard error is relatively large in comparison with the other standard errors, then the Pedigree score will be less, according to the rank in table 6.3.

Normal distribution

The standard error is relative small compared to the other distribution functions, therefore a good fit is assumed. The normal distribution has, together with the Pearson III MLE, the best ranking in the goodness-of-fit tests. The following Pedigree score is given: 2,5, this score is given because the normal distribution is not fitted for the larger discharges (figure 6.4), resulting in Pedigree score 2. But the fit to the lower discharges is good. The rankings in both the relative standard error and GOF-tests show the fit is the best, which results is Pedigree score 3 for a good fit to a reliable statistical model. Resulting in a total score of 2,5.

Exponential MOM

The standard error has a third rank, which is good. The standard error is small. The GOF test shows that the fit is not really well. The exponential distribution did not pass the Chi-square test with a 95% certainty level and the KS-test is just past with almost the lowest ranking. The resulting Pedigree score is between 1 and 2, so 1.5, because fitting is not really great, not all tests are past.

Exponential MLE

The standard error has a second rank, which is better than the exponential distribution with the Method of Moments parameter estimation. The standard error is small. The GOF tests give different results. The Chi-square test is not satisfied ($\alpha = 0{,}05$), the KS-test gives an average fit, with ranking 4 which is the same as the Pearson III MOM. The resulting Pedigree score is 1.8, because fit is a bit better than that of the exponential MOM, but the distribution is rejected by the Chi-square test and the size of the standard error is small.

Pearson III MOM

The standard error in the estimated runoff is larger than other distribution functions. The GOF tests give acceptable results. Both the X^2-test and the KS-test are satisfied, with the best fit of all distributions according to the KS-test. The Pedigree score that is gives is 2.0. This is because the GOF-tests are passed, but the standard error is relative large.

Pearson III MLE

The standard error is smaller than the Pearson III MOM, which gives higher certainty. The GOF tests are both passed fine. The fit according to the Chi-square test is second best and according to the KS-test the fit is the best of all distributions. The resulting Pedigree score is 2.5. This is the same as the normal distribution, because the Pearson III better fits for discharges with larger recurrence period (>100 years). Also both GOF-tests are passed well with good fits.

Gumbel MOM

The standard error in the Gumbel discharges at with a given recurrence period is the largest of all the distributions, resulting in the largest uncertainty of this part. The GOF test shows the same result as the uncertainty in the standard error. Positive is that Gumbel MOM is accepted by both goodness-of-fit tests, but the ranking in both test is the lowest. The Pedigree score is the lowest of all the distributions, because the standard error and the GOF tests are both not really great. A Pedigree score of 1 is too low, because the Gumbel MOM distribution is not totally rejected, but it is also not really accepted, therefore a Pedigree score of 1.1 is given.

Gumbel MLE
The standard error of Gumbel MLE distribution is comparable with the Pearson III distribution, average score. The goodness-of-fit tests give different results. The Chi-square test is not satisfied, while on the other hand the KS-test is satisfied, with third ranking. This distribution gives the largest discharges at a given return period. Because not all the GOF test are satisfied, and an average standard error the Pedigree score is 1.8, which is the same as the exponential MLE.

Empirical quality
The Annual Maximum method for the construction of the data series is an accepted method in frequency analysis research. The margin for error is relative small (Lang, Ouarda, & Bobée, 1999). The Peaks over Threshold method is less common, and there have not been made standard approaches for the POT. Peaks that are close together can have an influence on each other's magnitude, so peak discharges measured within a few days might not be selected for the POT. The approach used depends on the researcher, this gives more room for error, and thus larger uncertainty. The AM method gets a higher Pedigree score, because the data can directly be used, but the data is still subject to possible errors related to field measurements. Therefore a Pedigree score for the AM of 3 is given. The POT method has a Pedigree score of 2, because that are selected largely depends on the researcher. It is possible that two researchers get different data series with the POT method from the same measured discharge data series.

Methodological quality
The methodological quality is mostly based on the assessment of the parameter estimations. The different distributions have been fitted with two different parameter estimation methods, the Method of Moments and the Maximum Likelihood Estimation. Both these estimation methods are common accepted and used in flood frequency analysis. The uncertainty is not so much in the method itself, but in the possible choices. In this research only the MOM and MLE have been investigated, but other parameter estimation methods are also possible (section 8.3.1). There is not one standard estimation method with a single distribution possible, each dataset requires different ones. The investigation and use of different parameter estimation methods is a 'reliable method, and common within discipline', so Pedigree score for methodological quality is 3.

6.5.1 Summary of NUSAP assessment
The results of the NUSAP assessment show that the quality of the parameter estimation methods for the distribution functions have differences. Based on only the quality aspect, which is a part of the certainty of the propagations, the Normal and Pearson III MLE show similar scores. These two distribution functions could be preferred above the other functions.

	Normal	Exponential MOM	Exponential MLE	Pearson III MOM	Pearson III MLE	Gumbel MOM	Gumbel MLE
Statistical quality							
• Distribution type uncertainty	2	2	2	2	2	2	2
• Goodness-of-fit of distribution	2,5	1,5	1,8	2	2,5	1,1	1,8
Empirical quality	3	2	2	3	3	3	3
Methodological quality	3	3	3	3	3	3	3

Table 6.3 Results of NUSAP assessment

7 Propagation of uncertainties

In chapter 4, 5 and 6 the uncertainty in different sources has been identified: not only the quantitative but also the qualitative uncertainties. All these sources have an influence on the predicted distributions. The fitted distribution functions are propagated in order to give an idea about the discharges with a specific return period larger than the measurement period of 66 years. But with the propagation of the distributions, the uncertainty is also propagated. Some sources will have an influence on other uncertainties, they may increase or decrease. For uncertainty purposes it is important to know exactly how the uncertainties behave when propagated (Morgan & Henrion, 1990). If the uncertainty in a distribution is proving to be too large, it may be useful to gather more information. The NUSAP method can give an insight in where the focus on the gathering of more/better information should be, because each step in the frequency analysis is assessed with NUSAP (Ravetz & Funtowicz, 2009).
Chapter 6 'Statistic' gives a better insight into the stochastic error in the selected distributions. This section will deal with the propagation of the measurement error.

7.1 Quantitative uncertainties

7.1.1 Propagation of measurement error

A considerable variety of methods for propagation have been developed. But research done by Morgan & Henrion (1990) shows that it is not possible to point out one propagation simulation method. "The choice of a propagation method should depend on both the nature of the problem and the resources available to the analyst." (Morgan & Henrion, 1990, p. 172).
There are roughly two ways to propagate the measurement uncertainty in a flood frequency analysis. The first is Monte Carlo simulation (MC); the second is Latin Hypercube Simulation (LHS).In this research the LHS is used. For the LHS the input is 66 years of annual maximum discharges. The measurement error is computed in chapter 4 according to the velocity area method and is 1,56%. Each measured discharge in the data series is assumed to have a Normal spread. For each data point a sample is made, with the data point as mean and the uncertainty in measurement the sigma. When repeating this for 10.000 times, a normal spread can be seen. This is done for each data point, thus creating a matrix for the 10.000 random samples in each of 66 annual maximums. For each row in the matrix a distribution function is fitted with parameter estimation. In this research the Pearson type III distribution with parameter estimation according to the Method of Moments. The Pearson III distribution is chosen, because it is commonly used in China. A visual comparison between the Pearson III MOM and MLE show not a large difference in the distribution of discharges at a given return period: only 481 m^3/s (1,6%) at a return period of 500 years (Appendix E). The loss of accuracy is found acceptable. The parameters α, β and γ in the Pearson III function have a Normal probability density, but the β and γ have some skew. The result of the parameter estimation is shown below.

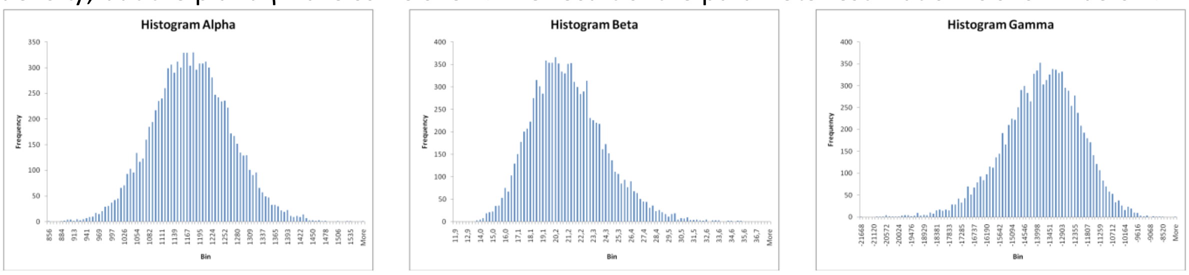
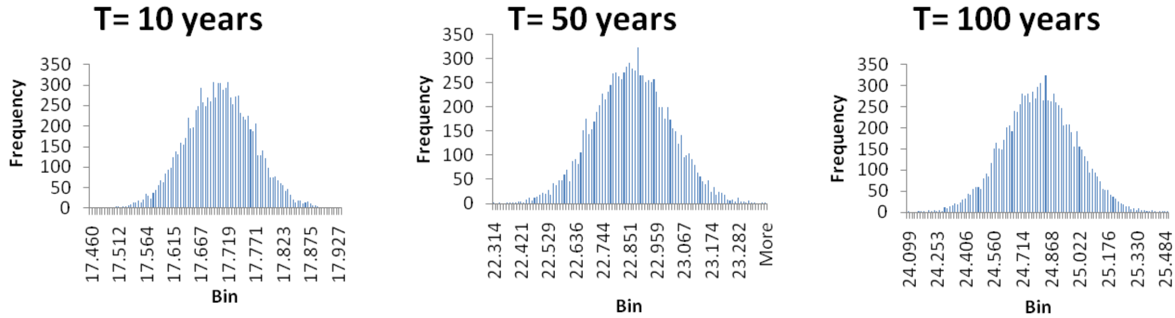
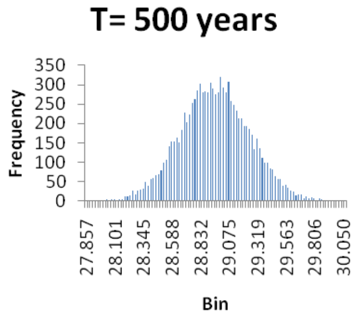


Figure 7.1 Probability densities of Pearson III parameters

The parameters have the following mean: α=1179 ±88 β= 21±3 and γ=-13856±1666

Continuing with the method the quantile estimations and standard errors are computed. The results are shown below for the 10, 50, 100, 500 and 10.000 years return period.





T	Mean	Standard Error	-95%	95%
10	17701	1085	15532	19870
50	22856	1920	19016	26695
100	24801	2351	20098	29504
500	28948	3440	22067	35828
10.000	35821	5658	24504	47138

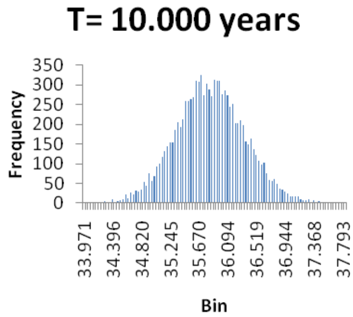


Figure 7.2 Probability densities of different return periods T

The distributed discharges at a given return period are almost identical to the discharges distributed in chapter 6. This may be the result of the small measurement error as computed with the velocity area method. The small measurement error is most likely because it is a relative error. Say the absolute measurement error in depth measurement is 30cm. In a river of 3 meters deep, this would cause a relative error of 0,3/3=10%. In a river of 10 meters deep the relative error would be 0,3/10=3%, so a relevant difference. It is worth doing a more in-depth analysis of the measurement error and its propagation. The result of the computation of the fitting of the Pearson III with the Method of Moments to the Annual Maximum is presented below.

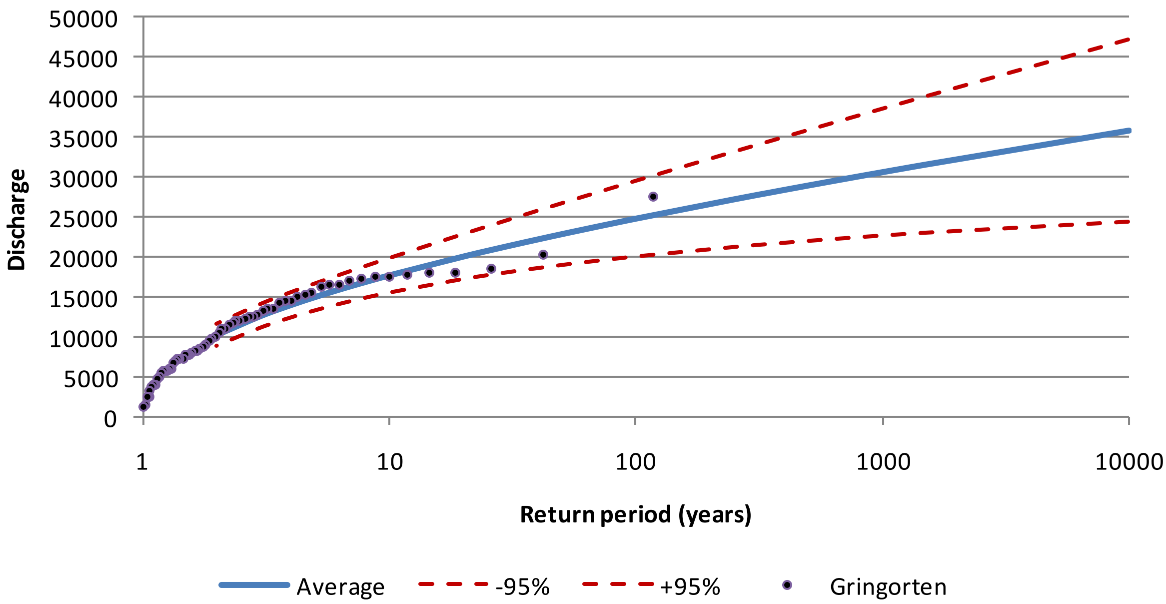


Figure 7.3 Pearson III distribution

7.2 Qualitative uncertainties

The qualified uncertainties are computed according to the NUSAP method and with the Pedigree table (table 3.1). The different uncertainty sources have been divided into three main quality groups: statistical, empirical and methodological. Each of these groups has been given a Pedigree score. For policy decisions it is important to know what the total quality of a frequency analysis is, in other words how trustworthy the analysis is. The NUSAP Pedigree scores can help with answering this problem. With a technique introduced by Van der Sluis et al. (2005b) and used by Ellis (2000) and Jansen (2007) the propagation of the individual Pedigree scores can be calculated. In order to do this, it is important to know exactly which methods are used to calculate the return period of discharges. So the methods in each step, as proposed in this research, will have to be identified. This research doesn't involve the propagation of the NUSAP scores, because the exact methods are unknown. The individual Pedigree scores have been presented in the different chapters 4-6. The Flowchart in figures 7.4 and 7.5 gives an insight in the logical order of the computation of a frequency analysis. With this chart it is possible to identify the parts in the frequency analysis with the least quality, and thus the largest uncertainty. In order to reduce the total uncertainty, the uncertainties in each section have to be minimized.

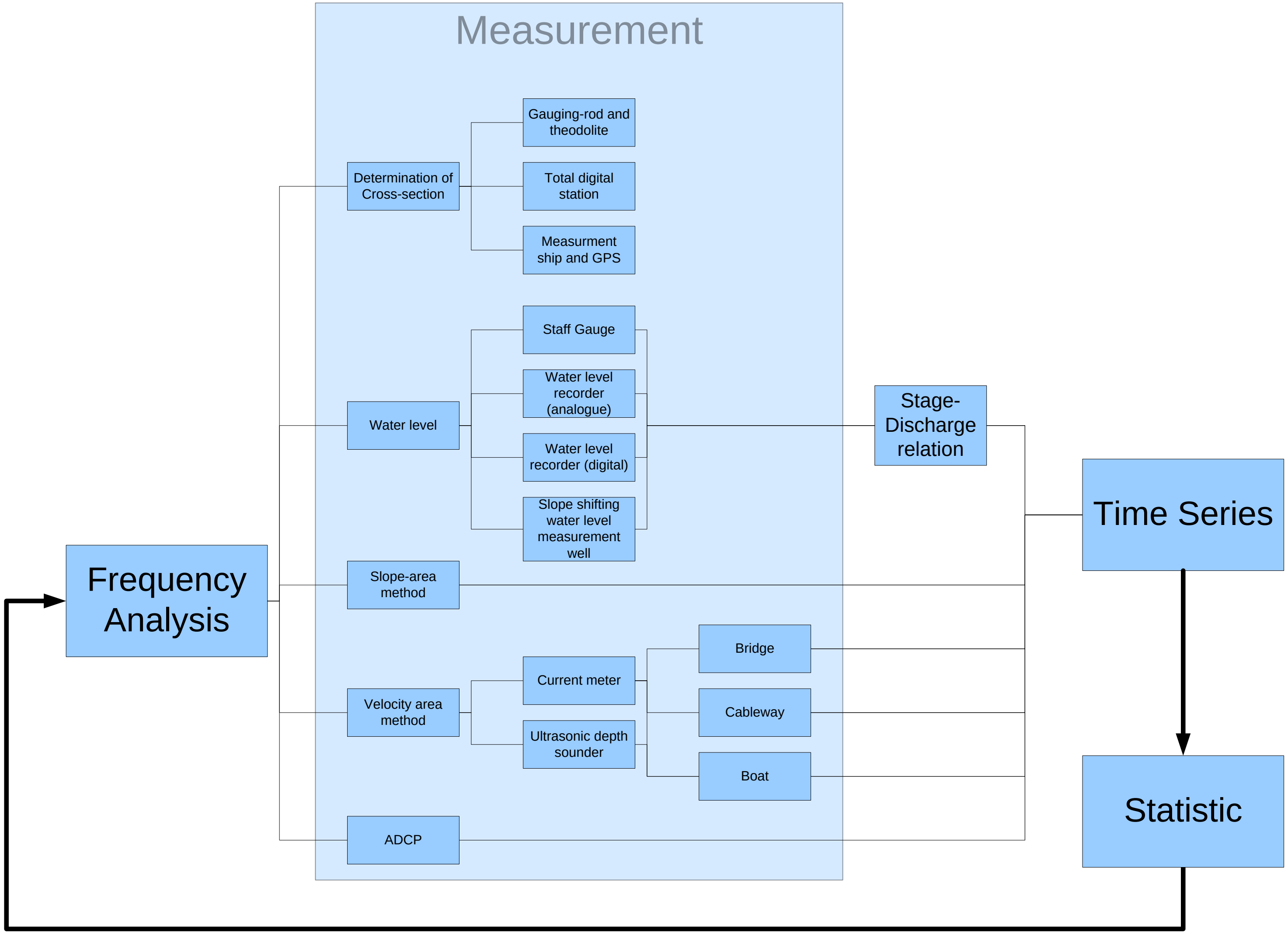
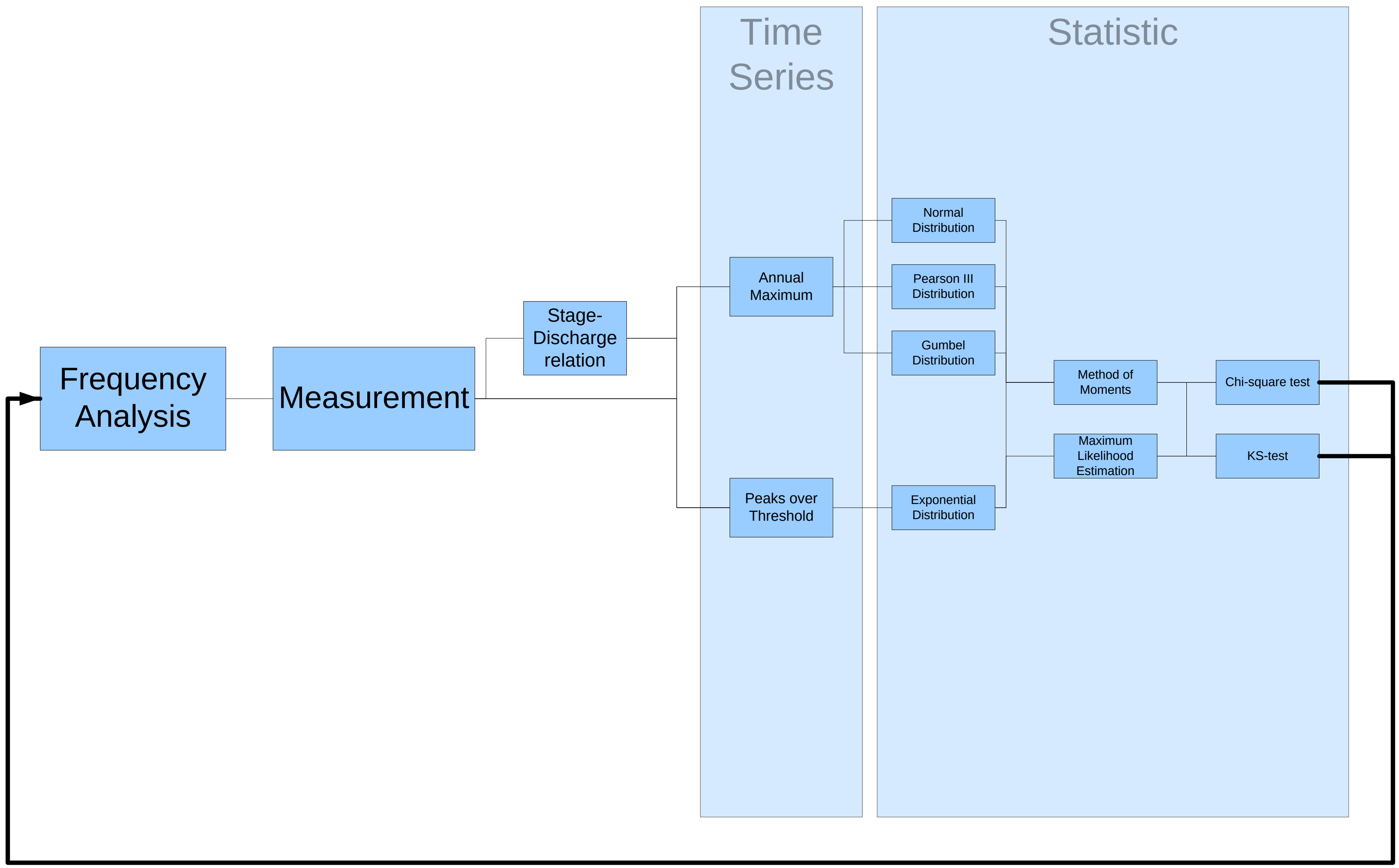


Figure 7.4 Detailed flowchart of Frequency Analysis



Figuur 7.5 Detailed flowchart of Frequency Analysis

8 Methodical reflection

The assessment of the uncertainties in this research is based on the NUSAP method. This method has been used before by Jansen (2007) to assess the design flood of the Meuse River in the Netherlands. The assessment of the uncertainty is based on two separate assessments: a quantitative and a qualitative one. The assessment is not specifically assessed by the NUSAP method, this uncertainty assessment is more based on classical statistic analysis. A good thing is the use of ISO standard calculation methods for the computation of uncertainty in the measurements of discharges by the velocity area method. This calculation is explained by an ISO standard, which gives more confidence to the outcome of the calculations, because the calculation method is assumed to contain little errors. The ISO explains the different factors and their uncertainties involved in the velocity area method. The use of ISO standards for the velocity area method has, as far known, not been used much in similar researches. This would mean less importance for the ISO standards. Still for this research the ISO has proved to be a good source for insight into the different parameters involved, and so aided in the assessment of the qualitative uncertainties in the measurement section.

The qualitative assessment in NUSAP is mostly bases on the Pedigree table. Van der Sluis et al. (2005b), Ellis et al. (2000) and Jansen (2007) don't go into details about the Pedigree score; only the resulting score is given. In their articles the recommendation is to give the Pedigree score according to the assessment of that particular model section by an expert on that section. In this research the establishment of the Pedigree scores is not based on the opinion of different experts, which makes the given Pedigree scores subjected to a larger error. The quality assessment is not completely objective, because the scores are based on the Pedigree table. The real situation is not always the same as the theoretical one in the Pedigree table. Human involvement in the establishment of the Pedigree gives more subjective results, especially when the researcher also is its own reviewer (Ravetz & Funtowicz, 2009). The quantitative uncertainty is more objective than the qualitative uncertainty, because this uncertainty is mostly based on the spread of numbers and can be calculated objectively by proven statistics.

The NUSAP method can be used as a starting point for the assessment of qualitative uncertainties, and to help with prioritizing the uncertainty reduction process and to improve the quality of the frequency analysis.

9 Conclusions & Recommendations

9.1 Conclusions

The main question in this research was: What is the uncertainty of the propagated discharge with a given return period using a frequency analysis for the Baihe discharge station at the Han River? This chapter will try to answer that question by answering the other research questions first.

The identification of the different uncertainty sources is split up into three stages: Measurement, Time series and Statistics. In each stage the uncertainty sources have been identified. The question which instrument has been used for measuring water depth, width and velocity of the Han River has been partially answered. The exact instruments used at Baihe station at different times is not known, but an investigation of possible used instruments give a good overview of the uncertainty, both quantitative and qualitative uncertainties. The measurement error is calculated using the velocity area method and proves to be small, only 1.5%. The use of the velocity area method in combination with the quantitative uncertainty computation of the ISO (2007) standard gives more confidence in the outcome. Because the ISO standards are internationally recognized, the confidence in the method is large, this results in larger Pedigree scores in the quality assessment. If the computations are done correctly, the statistic outcome is found trustworthy.

The small error of 3% in the velocity area method is explained by the size of the Han River, because absolute uncertainties are almost the same as in smaller rivers, the relative error is less. The quality of the measurement instruments and methods is good, according to the NUSAP method. Almost all the instruments that are likely to be used are widely used in hydrology fieldwork. A part that is missing in this research is the assessment of the stage-discharge relationship. But the discharge is mostly derived from that relationship. This is another source of quality uncertainty.

The discharge data time series is not made stationary or homogeneous in this research. The reason is that during this research information about changes in the river basin came not available. The uncertainty about whether or not the data is completely usable gives a low NUSAP Pedigree score for the Time series. The selection of peak discharges from the time series is done by selecting the annual maximum discharges; also the 'peaks over threshold' method is used to select peak discharges. The result was one series of AM discharges and one series of POT discharges.

In China mainly the Pearson type III distribution function is used to distribute the discharges for larger return periods. The parameter estimation method that is most common in China is the Maximum Likelihood Estimation (MLE). For this research also the Normal, Exponential and Gumbel distributions are used to compare with the Pearson III. All these distributions are fitted to the AM, with exception for the Exponential distribution, which is used with a POT threshold of 12.000m^3/s. The parameter estimation is done with both the Method of Moments (MOM) and MLE. The standard error for each fitted distribution is calculated. The goal was selection of one best distribution function, but it proves not to be possible. The uncertainty is the reason why the frequency analysis has no definite distribution. The margin for errors are too large and are overlapping each other in the selection of a distribution type in chapter 6. The goodness-of-fit test shows that the Normal distribution and the Pearson type III distribution both have an acceptable fit, but visual comparison of the Normal distribution with plot positions shows that the distribution doesn't fit for discharges with return period larger than 100 years. What can be concluded with more certainty is that the Gumbel distribution with MLE parameter estimation is fitted worst compared to the other investigated distributions. It may be needed to set a threshold for the Gumbel distribution, but the threshold should be lower than that of the Exponential distribution.

The propagated effect of all the uncertainty sources is not computed for all the distributions, only for the Pearson III. The measurement error is propagated by a Latin Hypercube Simulation (LHS). The result is a propagation of the return periods of certain discharges. In the LHS method the measurement uncertainty (of 1.5%) is used in the propagation of the measurement error in the rest of the frequency analysis. This 1.5% measurement error proves to be too small to have a real effect on the distribution function. The fitted distribution proves to be almost exactly the same as the fitted distribution with only one annual maximum series. This leads to the conclusion that the measurement error is not much significant for the frequency analysis.

The range in which the propagated discharges lie is between the lowest accepted distribution function and the highest. This would be the Normal resp. the Gumbel MOM distribution.
The Normal distribution has an average Q_{100} of 23089±1309 m^3/s. The Gumbel MOM distribution has Q_{100} of 27419±2882 m^3/s. The difference in the average distributed discharge with 100 years return period = 27419-23089= 4330 m^3/s.

When using the 95% confidence intervals the maximum difference between the upper 95% limit of the Gumbel MOM distribution and the lower 95% limit of the Normal distribution at T=100 is 12712 m^3/s, which is more than 50% of the average discharge of the Normal distribution. The main conclusion is that the uncertainty in the flood frequency analysis for the Han River is too large at this moment so that the frequency analysis is not to be of any practical usage. In this research all the conclusions are drawn upon the differences in discharges. The result of different discharges on water level changes has not been part of the research. A significant difference in discharge could have relative small impact on the gauge height. Therefore more research on the effects of discharge changes is recommended.

9.2 Recommendations

In order to minimize the uncertainty, the following recommendations are made for further research.
A more in-depth research into the velocity area method can be done. Research into the factors that influence the outcome of the VA calculations, so a sensitivity analysis, for a better understanding of the relative small measurement error.

A step that is missing in this research is the computation of a stage-discharge relationship. During this research there was no information about measured water levels, so this part was left out. The stage-discharge also has an influence on the uncertainty error.

The reviewing of the discharge data over time was not thorough. There was no information about the factors that had an influence on the discharge measurements in the past. It is worth doing this research with stationary data, so that the certainty in the outcome will be higher.

A sensitivity analysis can be carried out; so that the response of the different distributions for changes in input data are computed. The sensitivity analysis will provide more information about the validity and credibility of the computed distributions in this research. This sensitivity analysis should also look at the effects of changes to the water level. Significant differences in discharge could have relative small impact on the gauge height.

In this research the peaks over threshold method is used for selecting peak discharges from the time series. The POT is used as an input for the exponential distribution. At this moment the influence of different peaks occurring shortly after each other has not been taken into account. The exponential distribution could become more usable if the POT is reviewed and maybe the threshold is changed. The fit of the exponential distribution might be increased in this way.

The Pedigree scores in this paper are not propagated. For the qualitative uncertainty the benefit of a propagated Pedigree score would give more insight in the total quality of the frequency analysis. If the propagation is carried out for different periods in time, changes in the quality of the frequency analysis are known.

Altogether, this research gives multiple directions in which this research can be continued.

References

Abramowitz, M., & Stegun, I. (1965). *Handbook for Mathematical Functions.* New York: Dover Publications.

Booij, M., & Otter, H. (2007a). *Inleiding waterbeheer (Introduction watermanagement).* Enschede: University of Twente. [In Dutch]

Chen, H., Guo, S., & Xu, C.-y. (2007). Historical temporal trends of hydro-dimatic variables and runoff response to climate variability and their relevance in water resource management in the Hanjiang basin. *Journal of Hydrology , 344*, 171– 184.

Cui, Y., Chen, S., Zhang, L., & Huang, Y. (2008). Hydrometric technology and development in China. *Proceedings of Symposium SK, Hydrometry China.* Bureau of Hydrology.

Ellis, E., Li, R., Yang, L., & Cheng, X. (2000). Long-term change in village-scale ecosystems in China using landscape and statistical methods. *Ecological Applications , 10* (4), 1057-1073.

Han River. (2009). *Encyclopædia Britannica*. Retrieved March 29, 2009, from Encyclopædia Britannica Online: http://www.britannica.com/EBchecked/topic/253960/Han-River

Herschy, R. W. (2008). *Streamflow Measurement.* London: Taylor & Francis.

Hosking, J. R., & Wallis, J. R. (1997). *Regiona Frequency Analysis, an Approach on L-moments.* Cambridge: Cambridge University Press.

ISO. (2007). *ISO 748 Hydrometry - Measurement of liquid flow in open channels using current-meters of floats.* Geneva: International Organisation for Standardization.

ISO. (2007). *ISO/TS 25377 Hydrometric uncertainty guidance.* Geneva: International Organistation for Standardization.

Jansen, P. C. (2007). *Beoordeling van de kwaliteit van hoogwaterparameters berekend voor de Nederlandse Maas [MSc Thesis].* Enschede. [In Dutch]

Keijang, C. (1993). *Flow measurment in large rivers in China.* Butterworth-Heinemann Ltd.

Lang, M., Ouarda, T. B., & Bobée, B. (1999). Towards Operational guidelines for over-threshold modeling. *Journal of Hydrology* , 103–117.

Morgan, M., & Henrion, M. (1990). *Uncertainty. A guide to dealing with uncertainty in quantitative risk and policy analysis.* Cambridge: Cambridge University Press.

Qualitest International Inc. (2009). *Optical Theodolites*. Retrieved June 30, 2009, from Qualitest: http://worldoftest.com/opticaltheodolite.htm

Rao, R.A., & Hamed, K. H. (2000). *Flood Frequency Analysis.* Boca Raton, Florida, United States of America: CRC Press.

Ravetz, D. J., & Funtowicz, S. O. (2009). *NUSAP - The managent of uncertainty and quality in quentitative information*. Retrieved april 9, 2009, from NUSAP.net: http://www.nusap.net/sections.php?op=viewarticle&artid=14

Ravetz, J., & Functowitcz, S. (1990). *Uncertainty and Quality in Science for Policy.* Dordrecht: Kluwer.

Sluijs, J. v., Craye, M., Funtowitcz, S., Kloprogge, P., Ravetz, J., & Risbey, J. (2005a). Combining Quantitative and Qualitative Measures of Uncertainty in Model-Based Enviromental Assessment: The NUSAP system. *Risk Analysis , Vol. 25* (No. 2).

Sluijs, J. v., Risbey, J., & Ravetz, J. (2005b). Uncertainy Assessment of VOC emissions from paint in the Netherlands. *Environmental Monitoring and Assessment* , 229-259.

Sluis, J. V., Kloprogge, P., Risbey, J., & Ravetz, J. (2003). Towards a synthesis of qualitative and quantitative uncertainty assessment: apllications of the Numeral, Unit,Spread, Assessment, Pedigree (NUSAP) system. *International Workshop on Uncertainty, Sensitivity, and Parameter Estimation for Multimedia Environmental Modeling.* Rockville.

Tsubaki, R., & Fujita, I. (2007). Interpolation and correction of 3-D velocity distribution from randomly measured ADCP data.

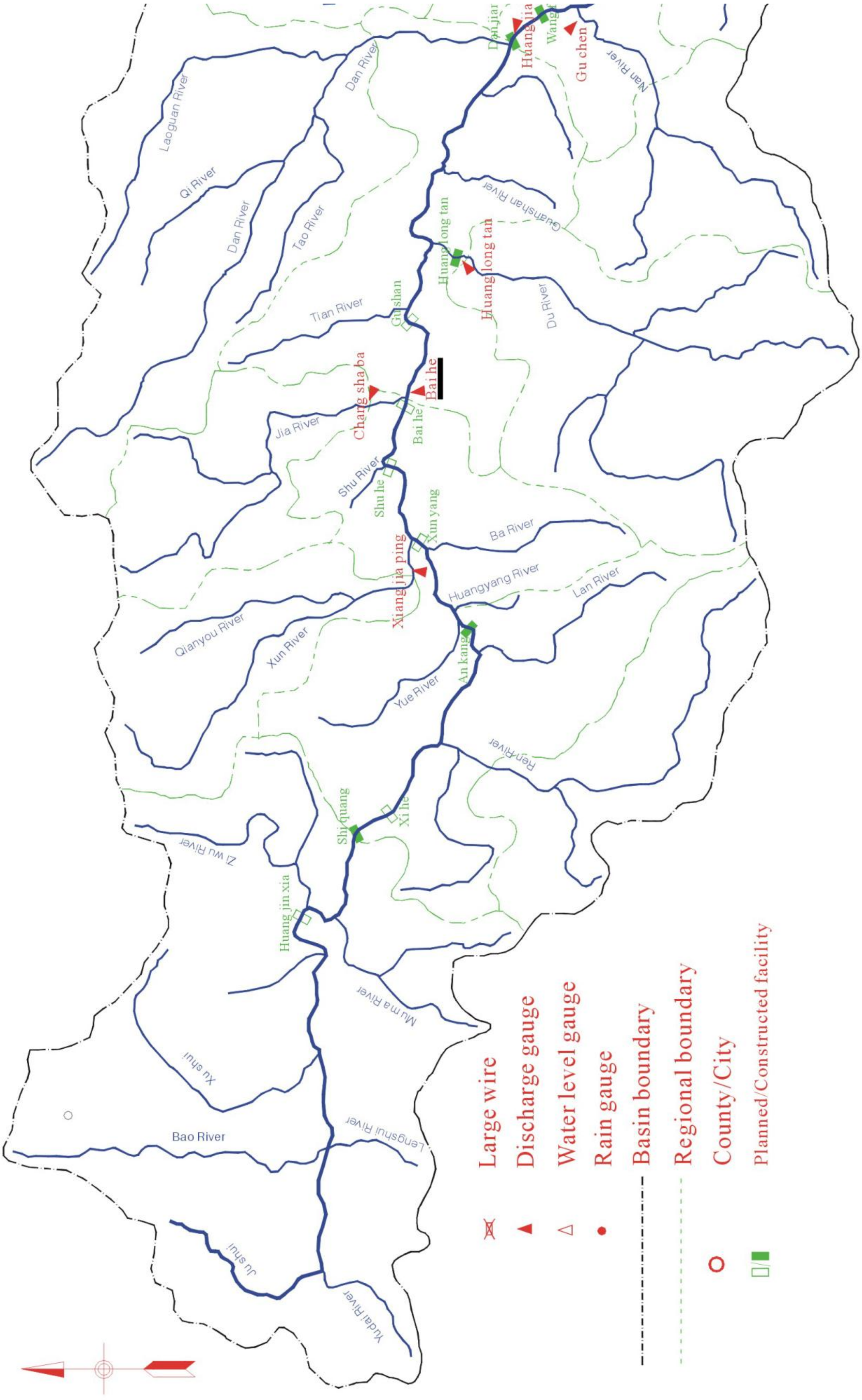
Walker, W., Harremoes, P., Rotmans, J., Van der Sluijs, J., Van Asselt, M., Janssen, P., et al. (2003). Defining Uncertainty. A conceptual basis for uncertainty management in model-based decision support. *Integrated Assessment , 4*, 5-17.

Wind, H. G., De Blois, C. J., Kok, M., Peerbolte, E. B., & Green, C. H. (1997). Uncertainty in flood damage assessment: When does it matter? An European perspective. *Proceedings of the First International RIBAMOD Workshop.* Delft, the Netherlands: Delft Hydraulics.

Xu, Y. (2005). *Appropriate modelling in decision suppot systems for river basin management.* Zutphen: Wohrmann Print Service.

Yangtze River. (2009). *Encyclopædia Britannica.* Retrieved March 21, 2009, from Encyclopædia Britannica Online: http://www.britannica.com/EBchecked/topic/651857/Yangtze-River

Appendix A – Detailed map of Hanjiang Basin



Appendix C – Calculation of uncertainties in the velocity-area measurement

The following calculation is presented in the international standard ISO 748:2007

The measurement method consists of dividing the channel cross-section under consideration into segments by *m* verticals and measuring width, depth and mean velocity associated with each vertical *i*. The mean velocity at each vertical is computed from point velocity measurements made at each of two depths (0.2 and 0.8) on the vertical.

Discharge measurement equation

$$Q = \sum_{i=1}^{m} b_i d_i v_i$$

Where

Q	is the (in cubic meters per second)
b_i	is the width of vertical
d_i	is the depth of vertical
v_i	is the mean velocity in vertical
m	is the number of verticals
i	vertical

Uncertainty in discharge measurement

$$u(Q) = \pm\sqrt{u_m^2 + u_s^2 + \frac{\sum_{i=1}^{m}\left[(b_i d_i v_i)^2 \left(u_{b_i}^2 u_{d_i}^2 u_{v_i}^2\right)\right]}{\left(\sum_{i=1}^{m} b_i d_i v_i\right)^2}}$$

Where

$u(Q)$ is the relative (percentage) combined standard uncertainty in discharge
u_{b_i} is the uncertainty in measurement of width
u_{d_i} is the uncertainty in measurement of depth
u_{v_i} is the uncertainty in measurement of velocity
u_s is the systematic uncertainty. An estimated practical value of 1% may be taken for this expression
u_s is the uncertainty due to the limited number of verticals

u_{v_i} is an estimate of the average of point measurements of velocity. The uncertainty is computed as follows:

$$u_{v_i} = \sqrt{u_{p_i}^2 + \frac{1}{n}\left(u_{c_i}^2 + u_{e_i}^2\right)}$$

N is the number of depths at which velocity measurements were made
u_{p_i} is the uncertainty in velocity in the vertical *i* due to limited measurements at each vertical
u_{c_i} is the uncertainty in velocity in the vertical *i* due to repeatability of the current meter
u_{e_i} is the uncertainty in velocity in the vertical *i* due to velocity fluctuations (pulsations)

For the cross section as measured on 30 September 2006 the following VA calculation is made :

Number of verticals	50
Number of pionts taken	2
Average velocity in measured section	1,61 m/s
Exposure time of current meter	61,88 sec -> 1 min

Above gives the following input:

Um	1
Us	1
Ub	0,5
Ud	0,5
Up	3,5
Uc	1
Ue (0,2)	3
Ue (0,8)	3
Ue	4,24
m	50
n	2

Using MS Excel the calculations for each vertical are easily made. The result is the following measurement uncertainty:

u(Q)	1,56%
coverage factor k	2
U95(Q)	**3,13%** at 95% confidence level

Appendix E –Distribution function parameter estimation

The following information about the discharge series is obtained, before the parameter estimation can start. Input was the normal – unstationary discharge data – as given in Appendix B. This information is obtained with function 'Descriptive Statistics' in MS Excel. Below a summary of the results is shown. They symbols are added for better understanding and reference for the parameter estimation calculations.

Discharges		*Symbols*
Mean	10646,51515	$\hat{\alpha}_1 = m_1' = \mu_1'$
Standard Error	658,2082585	
Median	10250	
Mode	7680	
Standard Deviation	5347,30917	$m_2^{1/2}$
Sample Variance	28593715,36	μ_2
Kurtosis	0,088492524	C_k
Skewness	0,436210721	$C_s = \gamma_1$
Range	26270	
Minimum	1330	
Maximum	27600	
Sum	702670	
Count	66	n
Confidence Level(95,0%)	1314,53265	

Tabel E.1 Results of Descriptive Statistics

Normal distribution

The normal distribution follows the following (Rao & Hamed, 2000, p. 90):
$\hat{x}_T = \hat{\alpha}_1 + u\hat{\alpha}_2$. $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are parameters that have to be estimated. This is done by two methods: Method of Moments (MOM) and Maximum Likelihood Estimation (MLE). For a given return period T, the corresponding probability of non-exceedence is F=1-1/T. It is easy to calculate the standard normal variate u corresponding to a probability F of non-exceedence. Abramowitz and Stegun (1965) give this value of u as,

$$u = W - \frac{C_0 + C_1 W + C_2 W^2}{1 + d_1 W + d_2 W^2 + d_3 W^3} + \varepsilon(P)$$

Where

$C_0 = 2{,}515517$ $d_1 = 1{,}432788$
$C_1 = 0{,}802853$ $d_2 = 0{,}189269$
$C_2 = 0{,}010328$ $d_3 = 0{,}001308$

And $W = \sqrt{-2\ln(P)}$ for P<0,5 and $\varepsilon(P)$ is less than 4,5 x 10^{-4}

Parameter estimation for normal distribution

Method of Moments (MOM)

$\hat{\alpha}_1 = m_1' = \mu_1'$= 10646,52
$\hat{\alpha}_2 = m_2^{1/2} = C_v * \mu_1'$ with $C_v = z = \mu_2^{1/2} / \mu_1'$ =28593715,36$^{1/2}$/10646,52=0,5023
Thus $\hat{\alpha}_2 = C_v * \mu_1'$ = 0,5023*10646,52=5347,31

Maximum Likelihood Estimation (MLE)
This method is also known as Maximum Likelihood Method (MLM)
Rao & Hamed (2000, pp. 86-87) shows that the parameter estimation for the normal distribution with both the MOM and MLE will give the same result, so the parameter estimates, quantile estimates and standard errors are the same in both methods.

Standard error

The standard error is given by the following equation (Rao & Hamed, 2000, p. 95), $S_T = [1 + u^2/2]^{1/2} \frac{4710}{\sqrt{48}}$, so with u as input, than the results are given in table E.2

Exponential distribution

The exponential distribution is part of a larger group of distributions: the Gamma distribution which also includes the Pearson type III (Rao & Hamed, 2000, p. 127). The exponential distribution can be expressed as follows, $\hat{x}_T = \hat{\varepsilon} + \hat{\alpha}\log(T)$. The $\hat{\varepsilon}$ and $\hat{\alpha}$ are parameters that have to be estimated by the MOM and MLE (Rao & Hamed, 2000, p. 134). T is again the return period of a given discharge.

Parameter estimation for exponential distribution

Method of Moments (MOM)

$\hat{\alpha}^2 = m_2$ and $\hat{\varepsilon} = m_1' - \hat{\alpha}$

$\hat{\alpha} = m_2^{1/2} = m_1' * C_v$ = 14937,88 * 0,1753 = 2618,56

$\hat{\varepsilon}$= 14937,88 - 2618,56= 12319,32

Maximum Likelihood Estimation (MLE)

N=66 m_1'=10646,52 and $x_1 = X_{min}$= 1330

$\hat{\alpha} = \frac{N(m_1' - x_1)}{(N-1)} = \frac{66(14937{,}88 - 12100)}{66-1}$=2881,54

and $\hat{\varepsilon} = \frac{Nx_1 - m_1'}{(N-1)} = \frac{66*12100 - 14937{,}88}{66-1}$=12056,34

Standard error

MOM

The standard error in the parameter estimation with MOM is given by Rao & Hamed (2000, p. 135),

$S_T^2 = \frac{\alpha^2}{N}[1 + 2K_T + 2K_T^2]$

K_T is given as $K_T = \log(T) - 1$

MLE

The standard error is given by, $S_T^2 = \frac{\alpha^2}{N}\left[1 + 2K_T + \frac{NK_T^2}{N-1}\right]$: so the standard error in the MLE is similar to the MOM.

Pearson III distribution

The Pearson Type III, or Gamma, distribution describes the probability of occurrence of a given event in a Poisson process. The Pearson Type III Distribution was first applied in hydrology to describe the distribution of annual maximum discharges (Foster, 1924). The Pearson Type III probability density function is

$$p(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)}(x-\gamma)^{\alpha-1}e^{-(x-\gamma)/\beta}$$

where Y is the lower bound of the distribution, β is a scale parameter, α a shape parameter, and Γ () the gamma function. Theses three parameters are related to the mean μ, variance σ^2, and skewness g_x of the distribution as

$\mu = \gamma + \alpha\beta$, $\sigma^2 = \alpha$, $g_x = 2/\sqrt{\alpha}$

Quantile estimation is carried out using the frequency factor K_T, and the quantile x_T is given by $x_T = \hat{\alpha}\hat{\beta} + \hat{\gamma} + K_T\sqrt{\hat{\alpha}^2\hat{\beta}}$. The $\hat{\alpha}$, $\hat{\beta}$and $\hat{\gamma}$ are the parameters that are estimated with the MOM or MLE. K_T is given by the Wilson-Hilferty Transformation (Rao & Hamed, 2000, p. 147).

$$K_T = \frac{2}{C_s}\left[\left\{\frac{C_s}{6}\left(u - \frac{C_s}{6}\right) + 1\right\}^3 - 1\right], C_s > 0$$

This approximation is quite accurate for $C_s \leq 1{,}0$, but can also be used for higher values. The C_s = 0,4362. u has already been computed with the normal distribution.

Parameter estimation for Pearson III distribution

MOM

N=66, m_1'= 10646,52, C_v=0,5023, C_s = 0,4362, Page 157

$\hat{\beta} = (2/C_s)^2$ = $(2/0{,}4362)^2$ = 21,0217

$\hat{\alpha} = \sqrt{(m_2/\hat{\beta})}$ with $m_2 = (m_1' * C_v)^2$

m_2 = $(10646{,}52*0{,}5023)^2$ = 28593715,36

$\hat{\alpha} = \sqrt{28593715{,}36/21{,}0217}$ = 1166,2768

$\hat{\gamma} = m_1' - \sqrt{m_2\hat{\beta}}$ = $10646{,}52 - \sqrt{28593715{,}36 * 21{,}0217}$ = -13870,5768

MLE

For this method the same method as described in Rao & Hamed (2000, pp. 157-162) is used.

$$\hat{\alpha} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \gamma) - \frac{N}{\sum_{i=1}^{N}\frac{1}{(x_i - \gamma)}}$$

$$\hat{\beta} = \frac{1}{N\alpha}\sum_{i=1}^{N}(x_i - \gamma)$$

For a given initial value of γ we can evaluate $\hat{\alpha}$ and $\hat{\beta}$ from the equations above and the objective is to satisfy the following equation:

$$F = \sum_{i=1}^{N}\log(x_i - \gamma) - N\log\alpha - N\psi(\beta) = 0$$

With

$$\psi(\beta) = \log(\beta) - \frac{1}{2\beta} - \frac{1}{12\beta^2} + \frac{1}{120\beta^4} - \frac{1}{252\beta^6} + \frac{1}{240\beta^8} - \frac{1}{132\beta^{10}}$$

This problem is solved using a Newton iteration method to improve the initial value of γ as,

$$\gamma_{n+1} = \gamma_n - \frac{F}{F'}$$

where

$$F' = \sum_{i=1}^{N} \frac{1}{(x_i - \gamma)} - \frac{N}{\alpha}\frac{\partial \alpha}{\partial \gamma} - N\psi'(\beta)\frac{\partial \beta}{\partial \gamma}$$

Where

$$\frac{\partial \alpha}{\partial \gamma} = -1 + \frac{N \sum_{i=1}^{N} \frac{1}{(x_i - \gamma)^2}}{\left[\sum_{i=1}^{N} \frac{1}{(x_i - \gamma)}\right]^2}$$

And

$$\frac{\partial \beta}{\partial \gamma} = \frac{-1}{\alpha} - \frac{1}{N\alpha^2} \sum_{i=1}^{N} (x_i - \gamma) \frac{\partial \alpha}{\partial \gamma}$$

And

$$\psi'(\beta) = \frac{1}{\beta} + \frac{1}{2\beta^2} + \frac{1}{6\beta^3} - \frac{1}{30\beta^5} + \frac{1}{42\beta^7} - \frac{1}{30\beta^9} + \frac{10}{132\beta^{11}}$$

This results in the following parameters:

$\hat{\alpha}$=1417,9932

$\hat{\beta}$=13,9470

γ=-9130,182

F=-2,09 *10^{-2}

$\psi'(\beta)$=0,0743

Standard error

MOM

Standard error is given by:

$$S_T^2 = \frac{\mu_2}{N}\left[1 + K_T\gamma_1 + \frac{K_T^2}{2}\left(\frac{3\gamma_1^2}{4} + 1\right) + 3K_T\frac{\partial K_T}{\partial \gamma_1}\left(\gamma_1 + \frac{\gamma_1^3}{4}\right) + \left(\frac{\partial K_T}{\partial \gamma_1}\right)^2\left(2 + 3\gamma_1^2 + \frac{5\gamma_1^4}{4}\right)\right]$$

With

$$\frac{\partial K_T}{\partial \gamma_1} = \frac{-2}{C_s^2}\left[\left\{\frac{C_s}{6}\left(u - \frac{C_s}{6}\right) + 1\right\}^3 - 1\right] + \frac{2}{C_s}\left[3\left\{\frac{C_s}{6}\left(u - \frac{C_s}{6}\right) + 1\right\}^2 \left\{\frac{u}{6} - \frac{2C_s}{36}\right\}\right]$$

$$\gamma_1 = C_s$$

MLE

Standard error is given by:

$$S_T^2 = \left(\frac{\partial x}{\partial \alpha}\right)^2 var\,\alpha + \left(\frac{\partial x}{\partial \beta}\right)^2 var\,\beta + \left(\frac{\partial x}{\partial \gamma}\right)^2 var\,\gamma + 2\left(\frac{\partial x}{\partial \alpha}\right)\left(\frac{\partial x}{\partial \beta}\right)cov(\alpha,\beta) + 2\left(\frac{\partial x}{\partial \alpha}\right)\left(\frac{\partial x}{\partial \gamma}\right)cov(\alpha,\gamma) + 2\left(\frac{\partial x}{\partial \beta}\right)\left(\frac{\partial x}{\partial \gamma}\right)cov(\beta,\gamma)$$

The different equations for the var and cov calculations can be found in Rao & Hamed (2000). The results are as follows:

$\left(\frac{\partial x}{\partial \gamma}\right)$=1

$var\ \alpha$= 607584,5645

$var\ \beta$= 197,5278811

$var\ \gamma$= 8,71E+07

$cov(\alpha,\beta)$= -1,08E+04

$cov(\alpha,\gamma)$= 6894829,88

$cov(\beta,\gamma)$= -129229,6079

Gumbel or Extreme Value Type I EV(1) distribution

Gumbel

$x_T = \hat{\beta} + \hat{\alpha}\ln\left[-\ln\left(1-\frac{1}{T}\right)\right]$ and $K_T = \frac{\sqrt{6}}{\pi}\left[-0{,}5772157 - \ln\left\{-\ln\left(1-\frac{1}{T}\right)\right\}\right]$.

Parameter estimation for Gumbel or Extreme Value Type I EV(1) distribution

MOM

N=66, m_1'= 10646,52, C_v=0,5023, $m_2^{1/2}$= 5347,31

$\hat{\alpha} = \frac{\sqrt{6}}{\pi} * \sqrt{m_2} = \frac{\sqrt{6}}{\pi}$ * 5347,31 = 4169,28

$\hat{\beta} = m_1' - 0{,}45005\sqrt{m_2}$ = 10646,52 – 0,45005* 5347,31= 8239,96

MLE

The $\hat{\alpha}$ in the MLE method is obtained by solving an equation, this is done with Newton's method.

N=66

$$F(\alpha) = \sum_{i=1}^{N} x_i * e^{\frac{-x_i}{\alpha}} - \left(\frac{1}{N} * \sum_{i=1}^{N} x_i - \alpha\right) \sum_{i=1}^{N} e^{\frac{-x_i}{\alpha}} = 0$$

Solving this equation with solver in MS Excel gives $\hat{\alpha}$ = 4589,55.

To verify: 68515,26 – (10646,52-4589,55) * 11,31 ≈ 0

$\hat{\beta} = \hat{\alpha}\ln\left[\frac{N}{\sum_{i=1}^{N} e^{\frac{-x_i}{\alpha}}}\right]$ = 4589,55*(66/11,31)=8095,08

Standard error

MOM

The standard error in Gumbel MOM calculation is given by the following equation:

$S_T^2 = \frac{\alpha^2}{N}\left[1{,}15894 + 0{,}19187Y + 1{,}1Y^2\right]$, with $Y = -\ln\left[-\ln\left(1-\frac{1}{T}\right)\right]$

MLE

The standard error in Gumbel MLE calculation is given by the following equation:

$S_T^2 = \frac{\alpha^2}{N}\left[1{,}1087 + 0{,}5140Y + 0{,}6079Y^2\right]$, with $Y = -\ln\left[-\ln\left(1-\frac{1}{T}\right)\right]$

Summary of results

Normal distribution				
		MOM / MLE		
Parameters		$\hat{\alpha}_1 = m'_1 = \mu'_1$ = 10646,52 $\hat{\alpha}_2 = \sigma$ = 5347,31		
Recurrence interval T (Years)	Probability	Q (m^3/s)	Absolute error (m^3/s)	Relative error (%)
10	0,10	17500	917	5,24
20	0,05	19444	1043	5,36
50	0,02	21631	1199	5,54
100	0,01	23089	1309	5,67
200	0,005	24422	1413	5,78
500	0,002	26039	1542	5,92

Table E.2 Results parameter estimation Normal distribution

Exponential distribution							
		MOM			MLE		
Parameters		$\hat{\alpha}$ = 2618,56 $\hat{\varepsilon}$ = 12319,32			$\hat{\alpha}$ = 2881,54 $\hat{\varepsilon}$ = 12056,34		
Recurrence interval T (Years)	Probability	Q (m^3/s)	Absolute error (m^3/s)	Relative error (%)	Q (m^3/s)	Absolute error (m^3/s)	Relative error (%)
10	0,10	18349	853	4,65	18691	819	4,38
20	0,05	20164	1160	5,75	20689	1066	5,15
50	0,02	22563	1572	6,97	23329	1393	5,97
100	0,01	24378	1885	7,73	25326	1641	6,48
200	0,005	26193	2199	8,40	27324	1889	6,91
500	0,002	28593	2615	9,15	29964	2216	7,40

Table E.3 Results parameter estimation Exponential distribution

Pearson type III distribution							
		MOM			MLE		
Parameters		$\hat{\alpha}$ = 1166,2768 $\hat{\beta}$ = 21,0217 $\hat{\gamma}$ = -13870,5768			$\hat{\alpha}$ = 1417,9926 $\hat{\beta}$ = 13,9470 $\hat{\gamma}$ = -9130,1908		
Recurrence interval T (Years)	Probability	Q (m^3/s)	Absolute error (m^3/s)	Relative error (%)	Q (m^3/s)	Absolute error (m^3/s)	Relative error (%)
10	0,10	17696	1083	6,12	17657	1108	6,28
20	0,05	20053	1403	6,99	20084	1415	7,05
50	0,02	22842	1915	8,38	22984	1879	8,18
100	0,01	24783	2345	9,46	25019	2258	9,03
200	0,005	26618	2800	10,52	26956	2654	9,85
500	0,002	28919	3430	11,86	29400	3196	10,87

Tabel E.4 Results parameter estimation Pearson type III distribution

Gumbel or Extreme Value Type I EV(1) distribution							
		MOM			MLE		
Parameters		$\hat{\alpha}$ = 4169,28 $\hat{\beta}$ = 8239,96			$\hat{\alpha}$ = 4589,55 $\hat{\beta}$= 8095,08		
Recurrence interval T (Years)	Probability	Q (m^3/s)	Absolute error (m^3/s)	Relative error (%)	Q (m^3/s)	Absolute error (m^3/s)	Relative error (%)
10	0,10	17622	1533	8,70	18423	1454	7,89
20	0,05	20624	1936	9,39	21727	1780	8,19
50	0,02	24508	2474	10,09	26003	2215	8,52
100	0,01	27419	2882	10,51	29208	2546	8,72
200	0,005	30320	3291	10,85	32400	2879	8,89
500	0,002	34146	3834	11,23	36613	3321	9,07

Table E.5 Results parameter estimation Gumbel or Extreme Value Type I EV(1) distribution

Plot of distributions