

# Trends and Extreme Values of River Discharge Time Series

Malaak Kallache

## Abstract

The central goal of this thesis is the development and application of sound data analysis frameworks to assess trends in mean and extreme river run-off. This includes the consideration of the auto-correlation structure (memory) of the data. Such an approach is very useful to, e.g., assess the anticipated intensification of the hydrological cycle due to anthropogenic climate change. The costs related to more frequent or more severe floods and droughts are enormous. Therefore an adequate estimation of these hazards and the related uncertainties is of major concern. We analyse discharge of basins of the Danube River and Neckar River in Southern Germany. Thereby we compare our frameworks with methods common in hydrology with respect to the calculation of assessment measures used by water management authorities. Furthermore, we look for spatial patterns of trends.

In the first part of this thesis, we evaluate trends of average discharge data. We use a trend test, where the auto-correlation structure of the data is explicitly modelled using stochastic FARIMA processes. This is a crucial task, because auto-correlations are capable of producing spurious trends. We assume that the data consists of an additive combination of natural variability, which is represented by the stochastic process, and potentially a deterministic trend component. The trend is estimated using wavelets and represents the fluctuation of the data on large time scales. In case the trend adds more variability to the data than the natural variability is likely to generate we consider the trend as significant. Hydrological data possesses short- and long-term memory, due to, e.g., weather regimes or consequences of large scale atmospheric patterns. This is confirmed by our results when analysing about 90 discharge records in the Neckar and Danube River basin: We find a spatially heterogeneous auto-correlation structure incorporating short- and long-term correlations. Thus, the complexity of our novel approach is necessary to adequately represent the data. Our results therefore differ from those of the Mann-Kendall trend test which is a common tool in hydrology. This test too often detects significant trends in case memory is present in the data. Trends and memory are interrelated phenomena. Thus, when applying our trend test we find that the chosen stochastic model influences the uncertainty of the trend estimate as well as the trend test result. We detect increasing and decreasing trends in both basins considered and the resulting spatial pattern is not ad hoc interpretable.

In the second part of this thesis, we use a point process to model trends in the extremes of river discharge data. Thereby we use exceedances over a threshold as extremes which we assume to be distributed according to a generalized Pareto distribution. In order to eliminate auto-correlation, the data are thinned out. Contrary to ordinary extreme value statistics, potential non-stationarity is included by allowing the model parameters to vary with time. By this, changes in frequency and magnitude of the extremes can be tracked. The model which best suits the data is selected out of a set of models which comprises the stationary model and models with a variety of polynomial and exponential trend assumptions. Common assessment measures, such as 100-year return levels, can be calculated from this model. Analysing winter discharge data of about 50 gauges within the

Danube River basin, we find trends in the extremes in about one third of the gauges examined. The spatial pattern of the trends in extremes is not immediately interpretable. Importantly, assuming stationary models for non-stationary extremes results in biased assessment measures. This underlines the necessity of the proposed methodology. The magnitude of the bias depends on the trend strength. Furthermore, the proposed approach allows to quantify the uncertainty of assessment measures such as return levels.

Summing up, this thesis provides enhanced trend and extreme value assessment frameworks, which account for the idea of memory in the data and environmental change. Trends in the mean values and extremes are found to be common in discharge when analysing river basins in Southern Germany. However, as yet no ubiquitous signal of increasing or decreasing discharge trends primarily related to climate change can be detected in these river basins. We observe neighbouring gauges often to display distinct behaviour, possibly due to non-climatic factors such as changes in land use or soil conditions. In any case, a sound trend assessment is the necessary basis for any process oriented, physical interpretation. Moreover, common practice of water management authorities can be improved by applying the proposed methods, and costs for flood protection buildings can be calculated with higher accuracy. That is, construction costs are reduced in case return level estimates are evidenced to be too high and possible future damage costs are dampened in case it is shown that those estimates are too low.

Keywords: trend assessment, auto-correlation, extreme value analysis, threshold excesses, declustering, wavelets, stochastic modelling, discharge

## **Zusammenfassung**

In dieser Dissertation wird ein profunder Datenanalyse-Methodenkanon entwickelt und angewendet, um Trends im Mittelwert und den Extremen von Abflusszeitreihen zu untersuchen. Insbesondere findet dabei die Autokorrelationsstruktur der Daten (das Gedächtnis) Berücksichtigung. Ein solcher Ansatz ist zum Beispiel sehr nützlich, um eine mögliche Intensivierung des globalen Wasserkreislaufes aufgrund des anthropogenen Klimawandels zu untersuchen. Die Kosten häufigerer oder schwererer Hochwasser und Dürren sind sehr hoch, deshalb ist eine akkurate Abschätzung dieser Gefahren samt ihrer Unsicherheiten von großer Bedeutung. Wir analysieren Abflüsse von Einzugsgebieten der Donau und des Neckar in Süddeutschland. Hierbei vergleichen wir unseren Methodenkanon mit in der Hydrologie gängigen Methoden, um Unterschiede und Gemeinsamkeiten bezüglich der Berechnung von Bemessungsgrößen, die von Wasserwirtschaftsämtern genutzt werden, feststellen zu können. Außerdem suchen wir nach räumlichen Mustern von Trends.

Im ersten Teil dieser Arbeit untersuchen wir Trends im Mittelwert von Abflussdaten. Beim Trendtest wird hierbei die Autokorrelationsstruktur der Daten mittels stochastischer FARIMA-Prozesse modelliert. Dies ist ein entscheidendes Merkmal des Trendtests, da durch Autokorrelationen unechte Trends vorgetäuscht werden können. Wir nehmen an, dass die Daten additiv zusammengesetzt sind aus natürlicher Variabilität, die durch den stochastischen Prozeß repräsentiert wird, und möglicherweise einer deterministischen Trendkomponente. Dieser Trend wird mit Hilfe von Wavelets bestimmt und repräsentiert die Fluktuationen der Zeitreihe auf großen Skalen. Falls der Trend mehr Variabilität verursacht als die natürliche Variabilität zu generieren vermag, betrachten wir den Trend als

signifikant. Hydrologische Zeitreihen zeichnen sich durch Kurzzeit- und Langzeit-Gedächtnis aus, das zum Beispiel durch Wetter oder Folgen großskaliger atmosphärischer Muster verursacht werden kann. Dies wird durch die Ergebnisse unserer Analyse von ca. 90 Abflusspegeln der Einzugsgebiete von Neckar und Donau bestätigt: Wir finden räumlich sehr heterogene Autokorrelationsstrukturen der Daten, die sowohl Kurzzeit- als auch Langzeit-Korrelationen beinhalten. Dies zeigt, dass die Komplexität des hier präsentierten Ansatzes zur Bewertung von Trends notwendig ist, um die Zeitreihen adäquat zu repräsentieren. Unsere Ergebnisse unterscheiden sich deshalb auch von denen des in der Hydrologie gebräuchlichen Mann-Kendall-Trendtests, der zu oft signifikante Trends diagnostiziert, falls die analysierten Daten Autokorrelationen besitzen. Trends und Gedächtnis sind zusammenhängende Phänomene. Beim Anwenden unseres Tests stellen wir also fest, dass das gewählte stochastische Modell sowohl die Unsicherheit des Trendschätzers beeinflusst als auch das Ergebnis des Trendtests. Wir finden sowohl steigende als auch fallende Trends in beiden Einzugsgebieten und es ist kein ad hoc interpretierbares räumliches Muster auszumachen.

Im zweiten Teil dieser Arbeit modellieren wir Trends in den Extremwerten von Abflussdaten mittels Punktprozessen. Hierbei nehmen wir an, dass Überschreitungen eines Schwellenwertes gemäß der verallgemeinerten Pareto-Verteilung verteilt sind. Autokorrelationen in den Extremwerten werden entfernt, indem man die Daten ausdünt. Im Gegensatz zu konventioneller Extremwertstatistik wird eine mögliche Instationarität über zeitabhängige Modellparameter zugelassen. Auf diese Weise können Änderungen sowohl in der Häufigkeit als auch im Betrag der Extremwerte berücksichtigt werden. Das beste Modell wird dann aus einer Menge gewählt, zu der sowohl das stationäre Modell als auch Modelle mit einer Auswahl an polynomialen und exponentiellen Trendannahmen gehören. Gängige Bemessungsgrößen, wie das Jahrhunderthochwasser, können nun mit Hilfe dieses Modells berechnet werden. Wir analysieren Winter-Abflussdaten von ca. 50 Pegeln innerhalb des Einzugsgebietes der Donau und finden in ungefähr einem Drittel der Daten Trends in den Extremwerten. Das räumliche Muster der Trends in den Extremwerten ist nicht direkt interpretierbar. Wesentlich ist, dass die Anpassung stationärer Modelle an instationäre Extrema zu einer Verzerrung der Bemessungsgrößen führt. Dies unterstreicht die Notwendigkeit der vorgeschlagenen Methodik. Die Größe der Verzerrung hängt von der Stärke des Trends ab. Die hier vorgestellte Methode bietet außerdem die Möglichkeit, die Unsicherheit der ermittelten Bemessungsgrößen, wie zum Beispiel Wiederkehrschwelle, abzuschätzen.

Die vorliegende Dissertation liefert einen Methodenkanon zur verbesserten Trend- und Extremwertanalyse, die die Vorstellung von Gedächtnis in den Daten und einer sich mit der Zeit verändernden Umwelt aufgreift. Bei unseren Untersuchungen von Abflüssen in Flusseinzugsgebieten von Süddeutschland ermitteln wir oft Trends im Mittelwert und den Extremen. Dennoch ist in diesen Flusseinzugsgebieten kein einheitliches, in Beziehung zum Klimawandel stehendes, Signal fallender oder steigender Abfluss-Trends zu erkennen. Oft stellen wir vielmehr fest, dass sich benachbarte Pegel unterschiedlich verhalten. Dies wird möglicherweise durch Faktoren wie Änderung der Landnutzung oder der Bodenbeschaffenheit bestimmt, die primär nichts mit dem Klimawandel zu tun haben. Auf jeden Fall ist eine fundierte Trendbewertung die notwendige Grundlage für jegliche prozeßorientierte, physikalische Interpretation. Weiterhin kann die gängige Praxis von Wasserwirtschaftsämtern verbessert werden indem die vorgestellte Methodik angewandt

wird und Kosten für den Hochwasserschutz können genauer kalkuliert werden. So werden zum Beispiel Baukosten reduziert, wenn die Schätzung zu hoher Wiederkehrschwelen aufgezeigt werden kann und Kosten möglicher zukünftiger Schäden werden vermindert, falls die Unterschätzung solcher Wiederkehrschwelen belegt werden kann.

Stichworte: Trendbewertung, Autokorrelation, Extremwertanalyse, Schwellenwertüberschreitungen, Entclustern, Wavelets, stochastische Modellierung, Abfluss