

**Analysis of discharge series from the Rhine basin
at different levels of aggregation**

P.J.J.F. Torfs en H. Middelkoop

RAPPORT 68

Augustus 1996

Vakgroep Waterhuishouding
Nieuwe Kanaal 11, 6709 PA Wageningen

ISSN 0926-230X

931156

Contents

1	The disaggregation problem for the Rhine	5
1.1	Introduction and problem definition	5
1.2	Objectives	6
1.3	Restrictions	7
1.4	Investigated discharge series	7
2	The techniques	10
2.1	Aggregation and disaggregation	10
2.1.1	Definition of aggregation	10
2.1.2	Definition of disaggregation	10
2.1.3	High and low filtering and computational aspects	12
2.2	The VRF and MVRF function	15
2.2.1	Statistical disaggregation	15
2.2.2	The variance reduction function	15
2.2.3	The marginal reduction function	15
2.2.4	A few characteristic examples	17
2.3	Use of VRF and MVRF in disaggregation	25
2.3.1	Stationarity of VRF and MVRF	25
2.3.2	Stationarity under climate change	25
2.3.3	Extrapolating VRF and MVRF	26
2.3.4	Need for parametric models	27
3	Application for discharges in the Rhine catchment	29
3.1	Difference between summer and winter	29
3.2	Along the Mosel	29
3.3	Along the Rhine	29
3.4	Some tributaries of the Rhine	39
3.5	The Vecht under different climates	39
4	After variance : covariance	44
4.1	The MACF and MCCF functions	44
4.2	The "all independ" choice	44
5	Conclusions and continuation	48
5.1	Conclusions	48
5.2	What to do next	48
5.2.1	Stability under climate change	48
5.2.2	Proposal for parametrisation	49

5.2.3	Simulation with independencies	49
5.2.4	Criteria for evaluation	49
A	Proof of second interpretation of MVRF	50
B	An example of a typical set of computations	51
C	The fractal generation code	53

List of Figures

1.1	Map of locations of discharge measurements	8
1.2	Info-table on the measuring stations	9
2.1	A typical time series : 1024 days of discharges (m^3s^{-1}) of the Rhine at Andernach	11
2.2	An aggregation of the Andernach data with $T = 50$	11
2.3	The process $X_{[50]}$ for the Andernach data, the dotted lines are $X^{[50]}$ and $X^{[25]}$	12
2.4	The $X^{\{50\}}$ -process for the Andernach data	13
2.5	The $X_{\{50\}}$ -process for the Andernach data	13
2.6	The fourier transform of the Andernach data, solid line is the real component, dotted is the imaginary component	14
2.7	The VRF function of the Andernach data	16
2.8	The MVRF function of the Andernach data	16
2.9	The rainfall series	18
2.10	The VRF function of the rainfall series	18
2.11	The MVRF function of the rainfall series	18
2.12	A "sampled" MVRF function	19
2.13	The first part of the rainfall series	19
2.14	The MVRF function of the first part of the rainfall series	20
2.15	The second part of the rainfall series	20
2.16	The MVRF function of the second part of the rainfall series	20
2.17	A white noise process	21
2.18	The MVRF function of the white noise process	21
2.19	A fractal process	22
2.20	The MVRF function of the fractal process	22
2.21	An AR(1) process with $\alpha = 0.98$	23
2.22	The MVRF function of the AR(1) process with $\alpha = 0.98$	23
2.23	An AR(1) process with $\alpha = 0.5$	24
2.24	The MVRF function of the AR(1) process with $\alpha = 0.5$	24
2.25	The sampled MVRF of the Andernach data	26
2.26	A possible extrapolation of the sampled MVRF of the sampled Andernach MVRF, the dashed line is the "true" MVRF	27
3.1	The Rhine at Stilli in the summer	30
3.2	The winter part of the Rhine at Stilli	31
3.3	The Saar in Laneuveville in winter time	32
3.4	The Saar at Fremersdorf in the summer	33

3.5	The Saar at Fremersdorf in the winter	34
3.6	The Mosel at Cochem in the winter	35
3.7	The Rhine at Neuhausen in the winter	36
3.8	The Rhine at Kaub in the winter	37
3.9	The Rhine at Lobith in the winter	38
3.10	The Main at Obernau in the winter	39
3.11	The Neckar at Gundelsheim in the winter	40
3.12	The Lippe at Schermbeck in the winter	41
3.13	The Vecht in the winter	42
3.14	The Vecht in the winter after climate change	43
4.1	The first three stages of the generation process.	46
4.2	The next three stages of the generation process.	47

Chapter 1

The diaggregation problem for the Rhine

1.1 Introduction and problem definition

To assess the impact of climate change on the discharge of the river Rhine, the International Commission for the Hydrology of the Rhine basin (CHR) initiated a research project in 1989. The main purpose of the project is to develop a water management model for the entire Rhine basin. Using the model the effects of changes in climate and land use on both average discharges and extreme discharges should be studied. The model must be valid for such changed conditions, and therefore it should have some physical basis. The spatial resolution of the model should be accordingly fine. To be able to simulate peak discharges accurately, the temporal resolution of the model should be maximum one day.

It was decided to phase the project, and to start with developing a first set of models, following a bottom-up approach as well as following a top-down approach. The Rhine basin can be subdivided into three major hydrological areas: the Alpine area, the middle mountains and the lowland area. Along the bottom-up line, detailed hydrological models with a physical basis are developed for representative sub-basins within each of these three areas. These models use a daily time step, and they are therefore suitable to analyze the effects of changes in climate and land use on both average and extreme discharges. Along the top-down approach a rough water balance model for the entire Rhine basin is developed. This model can be used to determine the effects of climate change on monthly average discharges, but does not provide precise estimations of changes in extreme discharges. In combination with empirical relationships between monthly discharges and peak flows, the Rhineflow model provides a first estimation of the probability of peak flows at the Dutch-German border.

The efforts that were required in the first phase of the project demonstrate that extending the detailed models to the entire Rhine basin will be a very time-consuming task. The data requirement is huge, and not all required data may be available. Also the effort for the modelling itself will be large. Nevertheless, the floods of December 1993 and January 1995 stressed once again that further development of a water management model for the Rhine basin is necessary, particularly because as the result of climate change the probability of peak flows might increase in the forthcoming century. Already during the forthcoming years, a more detailed insight in the effects of changes climate and

in land use on the Rhine discharge and the consequences for the inherent water resources planning processes are needed. Therefore, until the moment that a detailed model for the entire basin becomes available, within phase 2 of the CHR project "Impact of climate change on the discharge of the river Rhine" the gap between coarse and fine hydrological models should be bridged using alternative methods.

In 1994 the Dutch Institute for Inland Water Management and Waste Water Treatment (RIZA) has initiated a preliminary study to investigate the possibilities of downscaling the presently available water balance model for the entire Rhine basin to the required temporal resolution. For modelling peak flows of the Rhine, this is about one day. For this downscaling study, hydrological models with different temporal resolutions are required. In the first place, a water balance model is required to provide discharges at the "coarse" time scale. For this purpose, the RHINEFLOW-2 water balance model with a 10-daily time step is being developed as a second version of the monthly RHINEFLOW-1 model. In addition, detailed models, with a daily time basis are needed to evaluate whether the downscaling methods are valid under changed climate conditions.

The preliminary study indicated two approaches that can be followed to achieve results on a daily basis for the Rhine basin.

1- Using a relatively simple conceptual hydrological model ("event model") to simulate runoff with a time step of 1 day in combination with the water balance model with a resolution of 10 days (RHINEFLOW-2). This method particularly aims at modelling peak discharges: the coarse RHINEFLOW-2 model calculates the water balance on a 10-day basis, and it is used to determine the initial hydrological conditions in the area. On the basis of these initial conditions, the event model calculates daily discharges, using daily weather data.

2- Statistical techniques of downscaling series of 10-day average discharges calculated by the water balance model RHINEFLOW-2 to the required temporal scale of one day. It is impossible to make a perfect, absolute reconstruction at the finer scale. So, techniques should be used to derive the (most) important statistical characteristics of the finer scale. Once a proper algorithm is available to perform this statistical disaggregation, it can be used to simulate series of discharges at the finer scale. Assuming that the statistical characteristics of the disaggregations remain constant, they can also be used to determine the changes in discharge at the finer time scale under a different climate, given the changes in average discharges at the coarse scale. The detailed hydrological models are required to test whether the statistical downscaling provides correct results under changed climate conditions.

1.2 Objectives

The main objective of the project is to investigate whether statistical techniques can be applied for hydrological downscaling in the Rhine basin. For this purpose, the statistical properties of discharge series from various different parts of the

Rhine basin must be analysed first. This is the aim of this study. Of particular importance for disaggregating discharge series is the increase of variation that is obtained when decreasing aggregation level from 10 days to 1 day. From these analyses it is investigated whether robust relationships between variances at different aggregation levels of discharge series can be found, that allow to carry out a statistical downscaling of 10-day average discharges simulated by the RHINEFLOW-2 model to daily discharges. The downscaling should apply to sub-catchments with a size between 5000 and 15000 km².

1.3 Restrictions

This study is a follow-up of a preliminary inventory on candidate techniques for disaggregation of hydrological data that was carried out on request of the RIZA at the Wageningen Agricultural University (Torfs, 1995)¹. The study yielded several methods, both classical and new, that can be applied for the downscaling problem. Although not all of them are readily available, they seemed promising to be tested. The methods identified by Torfs (1995) will be tested for parts of the Rhine basin, with special emphasis on the Saar basin. This sub-basin was chosen for the following reasons: - it is part of the German Middle Mountain area, which comprises an important section of the Rhine basin. - for this area both a coarse model (RHINEFLOW-2) and detailed model (Saar model) will be available. - its size is about 7000 km², which is in accordance with the requirements of the study - an appropriate database is available

In this study only statistical methods of univariate (temporal only) processes are described, since these already comprise several problems that must be solved. Taking into account spatial correlations and covariances, e.g. between discharges from different Rhine tributaries, raises problems that are far from being solved yet. Only when using conditional simulations based on the coarse scale discharges of a water balance model, a minor spatial correlation is achieved as the result of the spatial correlation between discharges in different tributaries at the coarse scale.

1.4 Investigated discharge series

For this study, discharge series from different gauging stations within the Rhine basin were analysed. The stations represent different areas and river stretches: the Alpine area (switzerland), Saar area, stations along the main tributaries of the Rhine in central Germany, a series of stations along the Rhine, and the Overijsselse Vecht in The Netherlands. The location of these stations is given in figure 1.1. The table in figure 1.2 gives more information on these stations.

1. Torfs, P.J.J.F. (1995), Disaggregation techniques for hydrological use. Wageningen, Dept. of Water Resources, Wageningen Agricultural University.

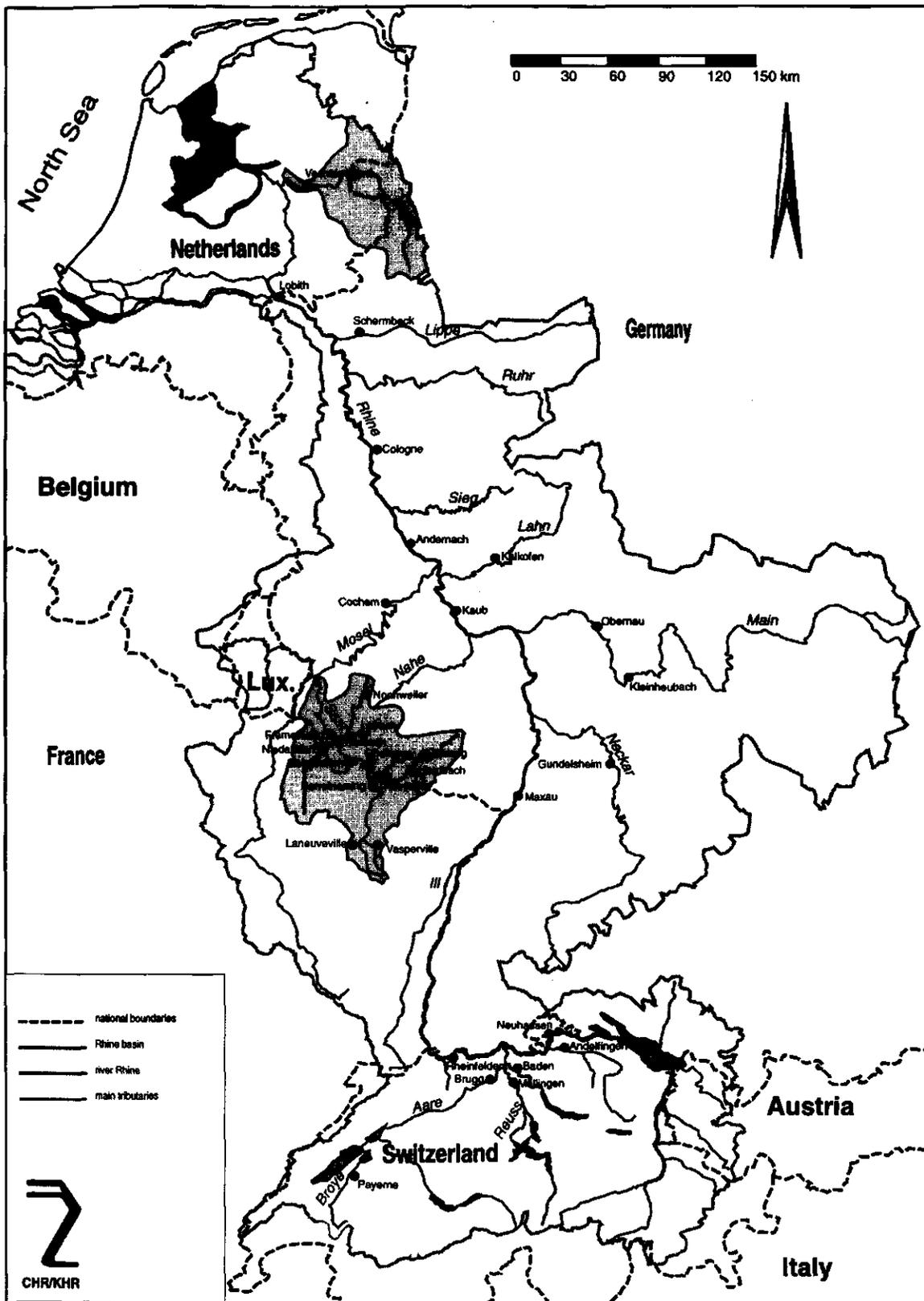


Figure 1.1. Map of locations of discharge measurements.

Figure 1.2. Info table on the measuring stations.

Swiss part of the Rhine basin

nr	station	river	upstream area (km ²)
18	Neuhausen	Rhein	11887
63	Rheinfelden	Rhein	34550
22	Andelfingen	Thur	1696
61	Baden	Limmat	2396
44	Brugg	Aare	11750
62	Stilli	Aare	17625
54	Mellingen	Reuss	3382
37	Payeme	Broye	399

Middle Rhine and main tributaries

nr	station	river	upstream area (km ²)
82	Konstanz	Bodensee	10922
107	Maxau	Rhein	50343
223	Kaub	Rhein	103729
306	Andemach	Rhein	139795
327	Koeln	Rhein	144612
544	Lobith	Rhein	160800
124	Gundelsheim	Neckar	12360
178	Kleinheubach	Main	21505
183	Obermau	Main	22300
298	Cochem	Mosel	27100
250	Kalkofen Up	Lahn	5320
396	Schembeck	Lippe	4762

Saar basin

nr	station	river	upstream area (km ²)
	Laneuveville	Saar	
	Vasperville	Saar	
514	Sarreinsming	Saar	1759
	Guedingen	Saar	
	Fremersdorf	Saar	
269	Hornbach	Schwalb	111
268	Contwig	Schwarz-bach	529
271	Reinheim	Blies	1790
277	Geislautern	Rosel	203
	Nonweiler	Prims	
281	Lebach	Theel	206
282	Nalbach	Prims	713
283	Niedaltdorf	Nied	1332

Overijsselsche Vecht: lowland Rhine basin

nr	station	river	upstream area (km ²)
615	Vechterweerd	Vecht	3779

Chapter 2

The techniques

2.1 Aggregation and disaggregation

2.1.1 Definition of aggregation

"Aggregation" is defined in Webster as :

"the collection of units or parts into a mass or whole"

In this study, the word aggregation will have a rather restricted meaning. First of all, a fixed aggregation length T is chosen. An aggregated time series $X^{[T]}$ is defined as a discrete time series which gives the mean of the original process X over every interval of the form $[(n-1)T, nT]$.

Formally :

$$X^{[T]}(n) \stackrel{\text{def}}{=} \frac{1}{T} \int_{(n-1)T}^{nT} X(\tau) d\tau \quad n = 0, 1, 2, \dots \quad (2.1)$$

Figure 2.1 shows a time series considered to be characteristic for this study : discharge data on a daily basis for the Rhine.

The point of view taken in this study is that we built up the aggregations *recursively* : at each step we double the aggregation length.

It is then easy to derive from the definition a basic recursion formula :

$$\begin{aligned} X^{[T]}(n) &= \frac{1}{T} \int_{(n-1)T}^{nT} X(\tau) d\tau \\ &= \frac{1}{T} \left\{ \int_{(2n-2)(T/2)}^{(2n-1)(T/2)} X(\tau) d\tau + \int_{(2n-1)(T/2)}^{(2n)(T/2)} X(\tau) d\tau \right\} \\ &= \frac{X^{[T/2]}(2n) + X^{[T/2]}(2n-1)}{2} \end{aligned} \quad (2.2)$$

2.1.2 Definition of disaggregation

Disaggregation is the inverse of aggregation. It is thus a technique by which the original process is reconstructed out of the aggregated.

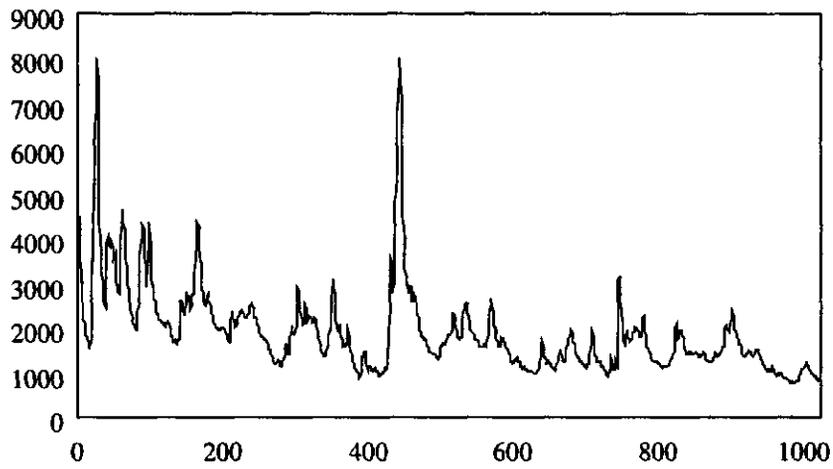


Figure 2.1: A typical time series : 1024 days of discharges (m^3s^{-1}) of the Rhine at Andernach

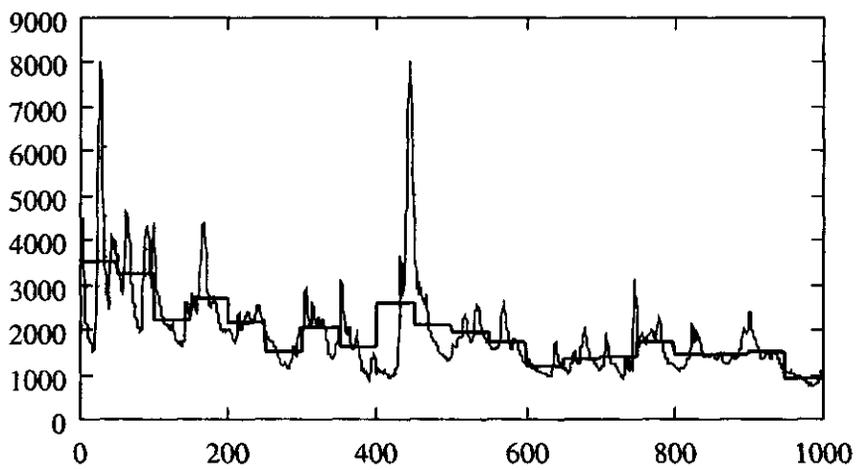


Figure 2.2: An aggregation of the Andernach data with $T = 50$

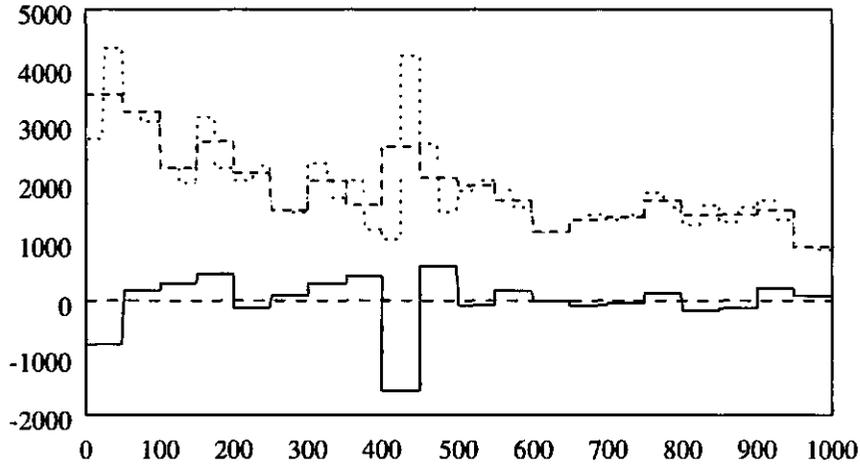


Figure 2.3: The process $X_{[50]}$ for the Andernach data, the dotted lines are $X^{[50]}$ and $X^{[25]}$

This disaggregation can be performed if one knows the so called *wavelet coefficients of level N* ¹

In general, these coefficients are defined by :

$$X_{[T]}(n) \stackrel{\text{def}}{=} \frac{1}{T} \left\{ \int_{(n-1)T}^{(n-1/2)T} X(\tau) d\tau - \int_{(n-1/2)T}^{nT} X(\tau) d\tau \right\} \quad n = 0, 1, 2, \dots \quad (2.3)$$

If, as will be done in this study, one concentrates on halving the aggregation length², recursive formulas can be written :

$$X_{[T]}(n) = \frac{X^{[T/2]}(2n-1) - X^{[T/2]}(2n)}{2} \quad (2.4)$$

With the help of this, the disaggregation step can be written as :

$$X^{[T/2]}(2n-1) = X^{[T]}(n) + X_{[T]}(n) \quad (2.5)$$

$$X^{[T/2]}(2n) = X^{[T]}(n) - X_{[T]}(n) \quad (2.6)$$

2.1.3 High and low filtering and computational aspects

One can consider aggregation as a kind of sampling of a *moving average* process :

$$X^{[T]}(t) = \frac{1}{T} \int_{t-T}^t X(\tau) d\tau \quad (2.7)$$

1. There is a "wavelet theory" that justifies the use of the name *wavelet* here. As this theory as such is not needed to understand this report, we refer the interested reader to the preceding report, referenced on page 7

2. Halving is a basic step in wavelet theory

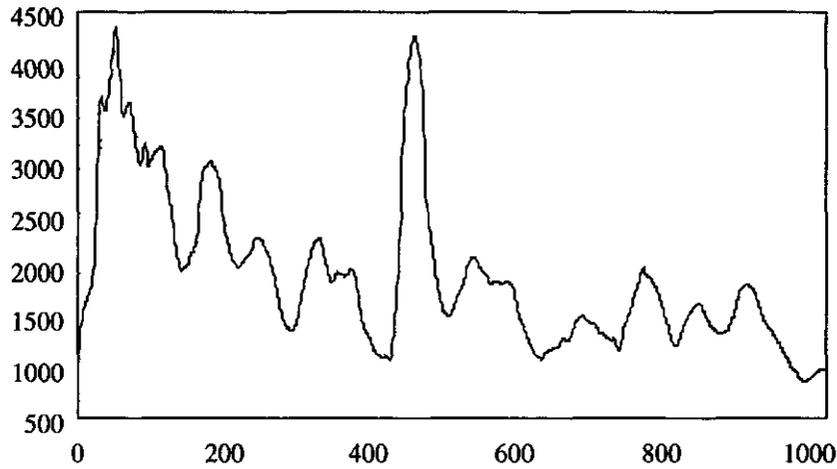


Figure 2.4: The $X^{\{50\}}$ -process for the Andernach data

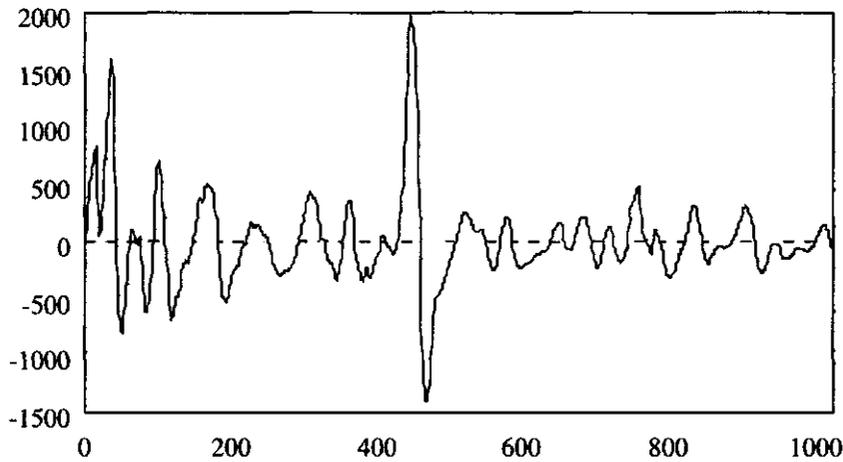


Figure 2.5: The $X_{\{50\}}$ -process for the Andernach data

Sampling means that one takes only the values of the moving average process at certain points :

$$X^{\{T\}}(n) = X^{\{T\}}(nT) \quad (2.8)$$

In a similar way, one can consider the wavelet coefficients as a sampling of the following continuous version :

$$X_{\{T\}}(n) = \frac{1}{T} \left\{ \int_{t-T}^{t-T/2} X(\tau) d\tau - \int_{t-T/2}^t X(\tau) d\tau \right\} \quad (2.9)$$

The moving average process is sometimes called *the low filtered process* and the continuous wavelet coefficient *the high filtered process*, because they amplify the low frequencies and the high frequencies respectively.

If one does not go down to the continuous underlying process, but stops at a certain aggregation level, one can write discrete analogs of the formulas above :

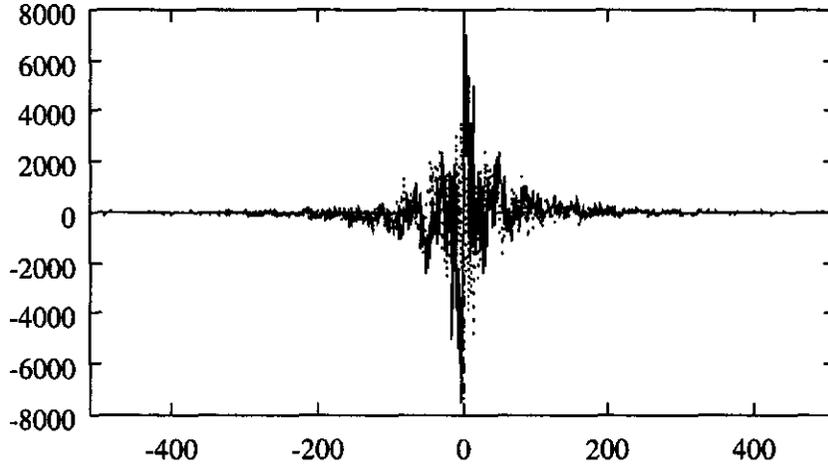


Figure 2.6: The fourier transform of the Andernach data, solid line is the real component, dotted is the imaginary component

$$X^{\{N*T\}}(k) = \frac{1}{N} \sum_{k-N+1}^k X^{\{T\}}(n) \quad (2.10)$$

$$X_{\{N*T\}}(k) = \frac{1}{N} \left\{ \sum_{k-N+1}^{k-N/2} X^{\{T\}}(n) - \sum_{k-N/2+1}^k X^{\{T\}}(n) \right\} \quad (2.11)$$

Both formulas can also be written as a convolution :

$$X^{\{N*T\}}(k) = \sum \psi^{\{N\}}(k-n) X^{\{T\}}(n) \quad (2.12)$$

$$X_{\{N*T\}}(k) = \sum \psi_{\{N\}}(k-n) X^{\{T\}}(n) \quad (2.13)$$

with :

$$\psi^{\{N\}}(k) = \frac{1}{N} \star \begin{cases} 1 & \text{if } 0 \leq k < N \\ 0 & \text{else} \end{cases} \quad (2.14)$$

$$\psi_{\{N\}}(k) = \frac{1}{N} \star \begin{cases} +1 & \text{if } N/2 \leq k < N \\ -1 & \text{if } 0 \leq k < N/2 \\ 0 & \text{else} \end{cases} \quad (2.15)$$

Convolutions can be calculated with the help of Fourier ³ transforms :

$$\mathcal{F}(X^{\{N*T\}}) = \mathcal{F}(\psi^{\{N\}})\mathcal{F}(X^{\{T\}}) \quad (2.16)$$

$$\mathcal{F}(X_{\{N*T\}}) = \mathcal{F}(\psi_{\{N\}})\mathcal{F}(X^{\{T\}}) \quad (2.17)$$

3. Ofcourse, Fourier techniques can also provide a much more fundamental insight in filtering, one can think of wavelet theory as a kind of time-frequency analysis, where Fourier analysis is only a frequency analysis. Wavelet coefficients are thus a analogon of Fourier coefficients. We refer for this to the technical appendix of the preceding rapport, referenced on page 1.3

In all practical situations, one only has a finite number of data ($n = 0, \dots, N$). This gives "starting up" problems, because the formulas 2.10 and 2.11 require values "before the first observation" (for $n < 0$). One way out of this problem is to put all data "in a circle", i.e. to take the last data values as the values preceding the first. When using this approach the Fourier formulas still hold.

If one has to do a lot of high and low filtering of the same data series, it pays to use the Fourier formulas above (note that one has to calculate $\mathcal{F}(X^{(T)})$ only once). If the number of data is a power of two, one has of course the FFT-algorithm ("Fast Fourier Transform") to calculate the necessary transforms.

2.2 The VRF and MVRF function

2.2.1 Statistical disaggregation

In many cases an exact disaggregation is impossible and a *statistical disaggregation* is the only possibility left.

For such a disaggregation one needs of course a good probabilistic description of the wavelet coefficients.

In all the analysis that follows, we assume the stochastic processes to be second order stationary.

The mean of aggregation processes and wavelet coefficients (and of the low and high filtered process) are trivial :

$$E [X^{[T]}(n)] = E [X^{(T)}(n)] = E [X(n)] = \mu \quad (2.18)$$

$$E [X_{[T]}(n)] = E [X_{\{T\}}(n)] = 0 \quad (2.19)$$

2.2.2 The variance reduction function

Because the mean is trivial, the variance is the first important characteristic to be studied.

The *variance reduction function* (or VRF) is defined as the function that gives for each aggregation length the variance of the corresponding aggregated (or low filtered) process.

Formally :

$$\text{VRF}(T) = \text{VAR} [X^{[T]}(n)] = \text{VAR} [X^{(T)}(n)] \quad (2.20)$$

This is a clearly decreasing function.

2.2.3 The marginal reduction function

The *marginal variance reduction function* (or MVRF) is defined as the function that gives for each aggregation length the variance of the corresponding wavelet coefficients (or high filtered process).

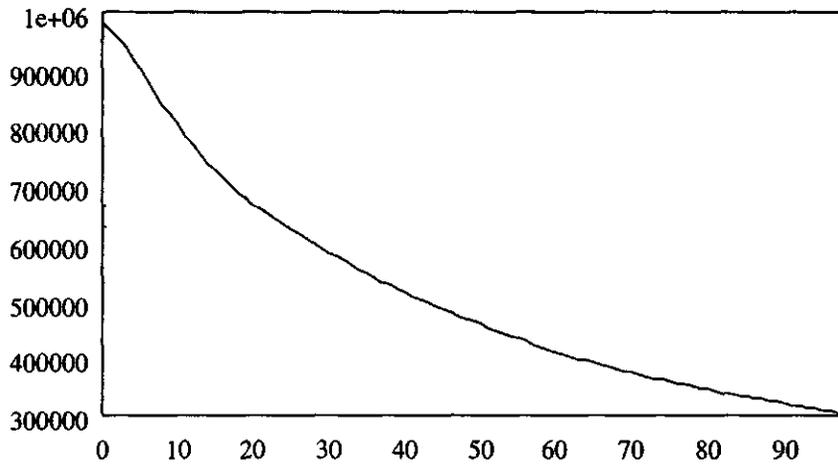


Figure 2.7: The VRF function of the Andernach data

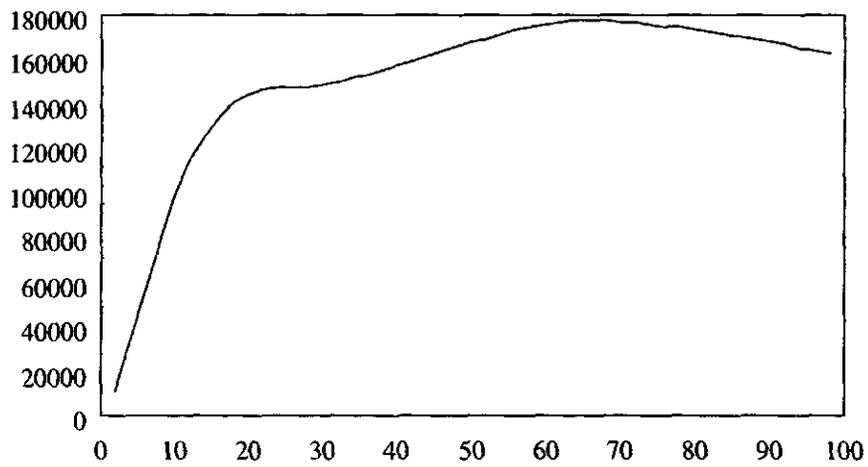


Figure 2.8: The MVRF function of the Andernach data

Formally :

$$\text{MVRF}(T) = \text{VAR} [X_{\{T\}(n)}] = \text{VAR} [X_{\{T\}(n)}] \quad (2.21)$$

There is also a second interpretation of the MVRF function :

$$\text{MVRF}(T) = \text{VRF}(T/2) - \text{VRF}(T) \quad (2.22)$$

A proof of this second interpretation is found in appendix A

So the point $\text{MVRF}(10) = 100000$ ⁴ in figure 2.8 may be interpreted in the following two ways :

1. It is the variance of the process $X_{[10]}$ that ones has to add/subtract (according to the formulas 2.5 and 2.6) to $X_{[10]}$ to obtain $X_{[5]}$.
2. It is the difference between the variances of the $X_{[5]}$ and $X_{[10]}$ process respectively.

It is the believe of the authors of this report, that the MVRF function plays a central role in the disaggregation analysis. This more so than the VRF function (although this is mathematically equivalent) because, as the examples in the following paragraph show, the form variations of the MVRF function are much more expressive than those of the VRF functions.

2.2.4 A few characteristic examples

A rainfall series

Figure 2.9 shows rainfall data measured on a very fine timescale : every value stands for the average of 10 seconds of rainfall. Rainfall is almost always measured in a cumulative way, and by that is always an aggregated process.

Figure 2.10 shows the VRF function of these rainfall data. This VRF function does not differ to much in character (for the naked eye at least) from the VRF function of the Andernach data (see figure 2.7).

The MVRF function of these rainfall data, as shown in figure 2.11 differs dramatically : the function does *not* go to zero for small T . Going from left to right, the Andernach VRF decreases continuously (after a certain point), the rainfall VRF sudenly starts to increase again.

This also shows, that if the rainfall VRF should only be given for $T \geq N * 10\text{sec}$, as illustrated by figure 2.12 one could have no idea about the variance that is hidden in the very small time intervals, and one would be tempted to extrapolate in a such a way that it would resemble the Andernach VRF.

Another consequence from this is that the process is certainly not continuous : the variance to be added when going to smaller and smaller scales does not decrease, as a should be the case for a continuous process.

4. To appreciate the large numbers on the y-axis : a variance of $\approx 100000 \text{ m}^6\text{s}^{-2}$ corresponds to a standard deviation of $\approx 300 \text{ m}^3\text{s}^{-1}$.

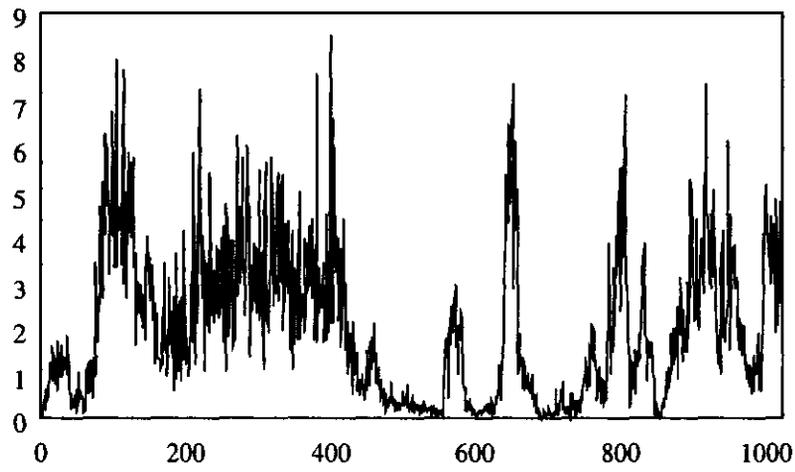


Figure 2.9: The rainfall series

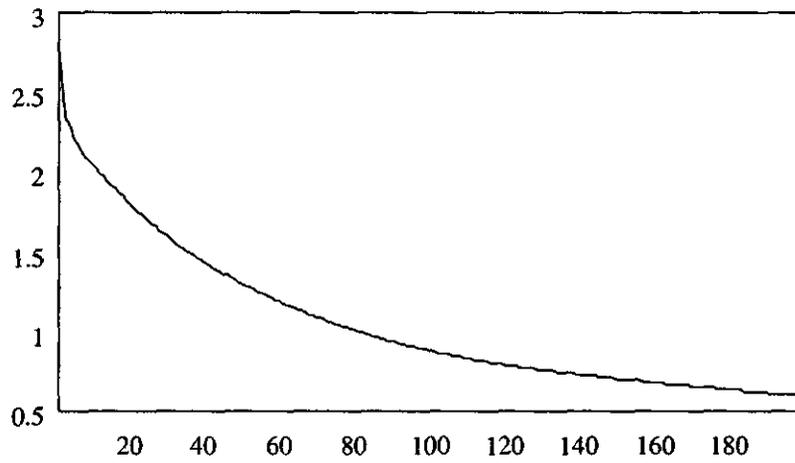


Figure 2.10: The VRF function of the rainfall series

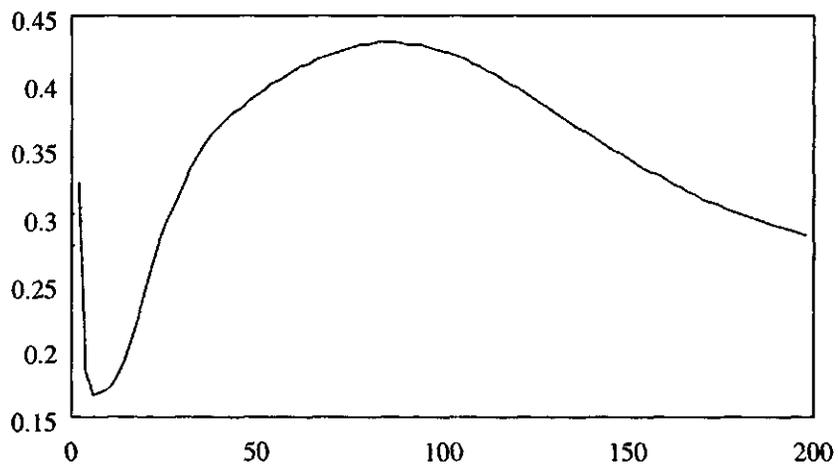


Figure 2.11: The MVRF function of the rainfall series

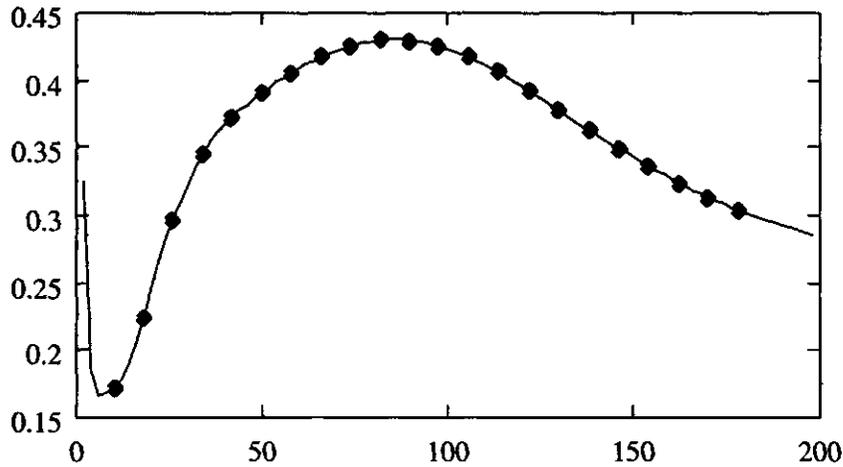


Figure 2.12: A "sampled" MVRF function

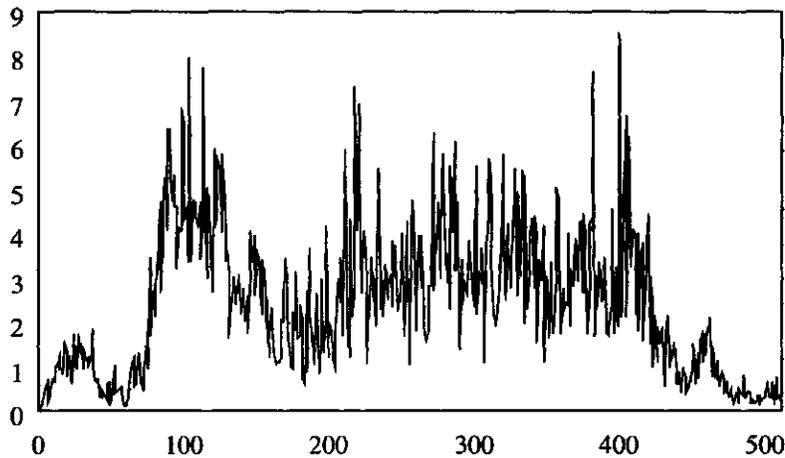


Figure 2.13: The first part of the rainfall series

A closer look on the first part Figure 2.13 shows only the first part of the rainfall data. This part is very irregular. Figure 2.14 shows the MVRF function of this first part. Clearly the steep rising when $T \rightarrow 0$ is more pronounced here.

A closer look on the second part Figure 2.15 shows only the second part of the rainfall data. This part is clearly more regular than the first part. Figure 2.16 shows the MVRF function of this first part. Clearly the steep rising when $T \rightarrow 0$ is almost not present here, and the MVRF function resembles more that of the Rhine data in figure 2.8.

Conclusions Clearly the form of the MVRF function found for the Rhine data (see figure 2.8) is not the only one possible.

If the time series consists of different parts, which have different behaviour, the MVRF of the combination shows a kind of mixture of both parts, as can be

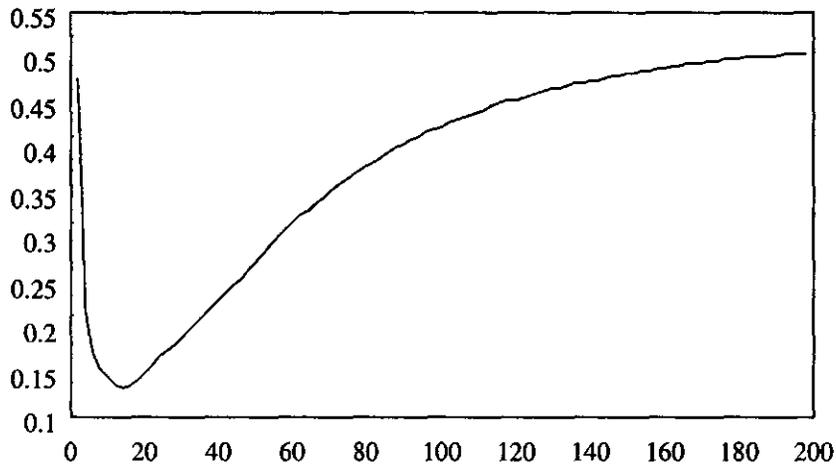


Figure 2.14: The MVRF function of the first part of the rainfall series

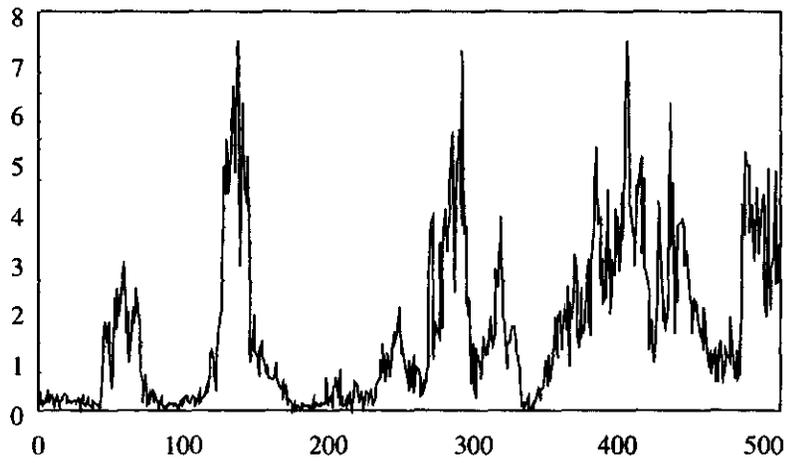


Figure 2.15: The second part of the rainfall series

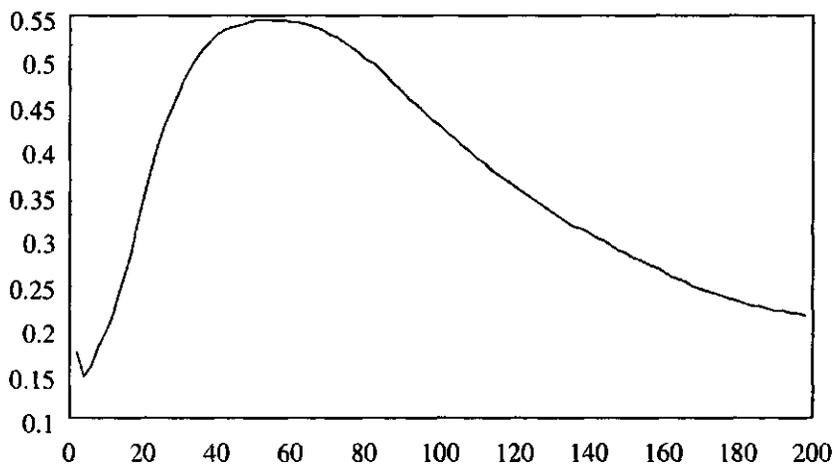


Figure 2.16: The MVRF function of the second part of the rainfall series

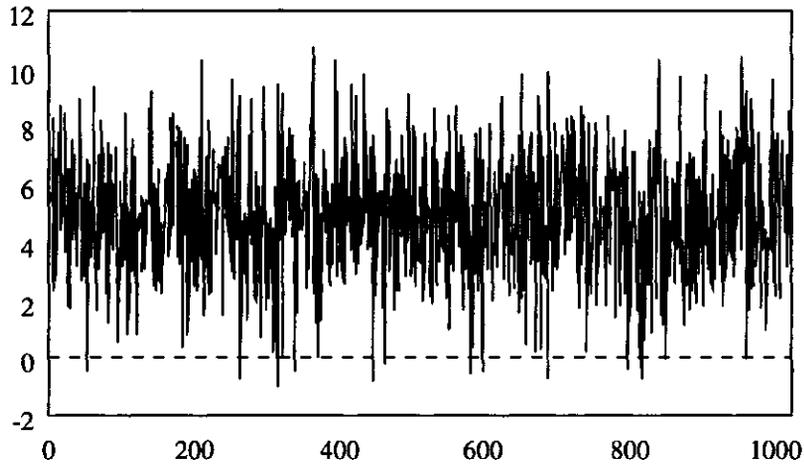


Figure 2.17: A white noise process

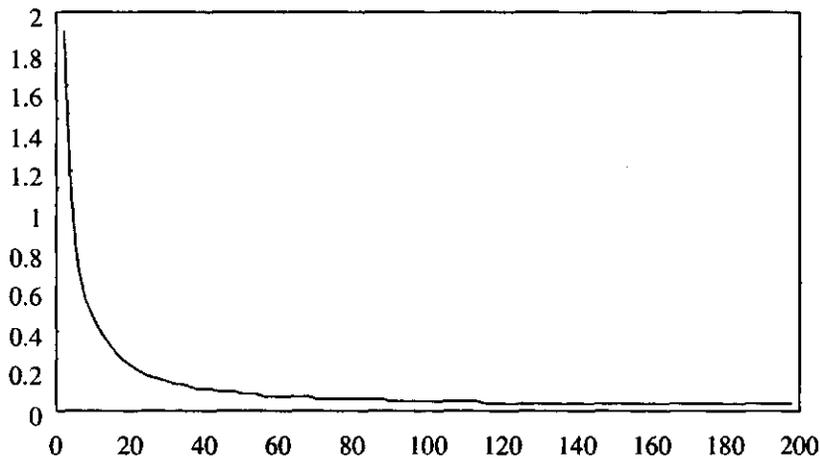


Figure 2.18: The MVRF function of the white noise process

seen from figures 2.11, 2.14 and 2.16.

White Noise

Figure 2.17 shows a white noise process : all the data are independent (Gaussian) variables. The MVRF function of this process –see figure 2.18– shows a very characteristic behaviour : the variance to be added to disaggregate keeps increasing when one goes down to smaller and smaller scales. One can show that in this case :

$$\text{MVRF}(T) \sim \frac{1}{T} \quad (2.23)$$

Comparing this to the rainfall data, one can conclude that the MVRF as depicted in figure 2.11 shows “white noise” behaviour near the origin, which is caused by the manifest white noise character of the first part of this data, as shown by figure 2.14.

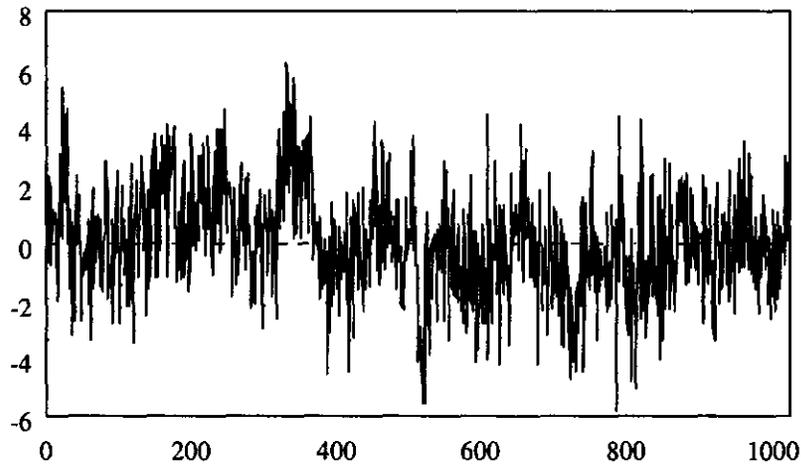


Figure 2.19: A fractal process

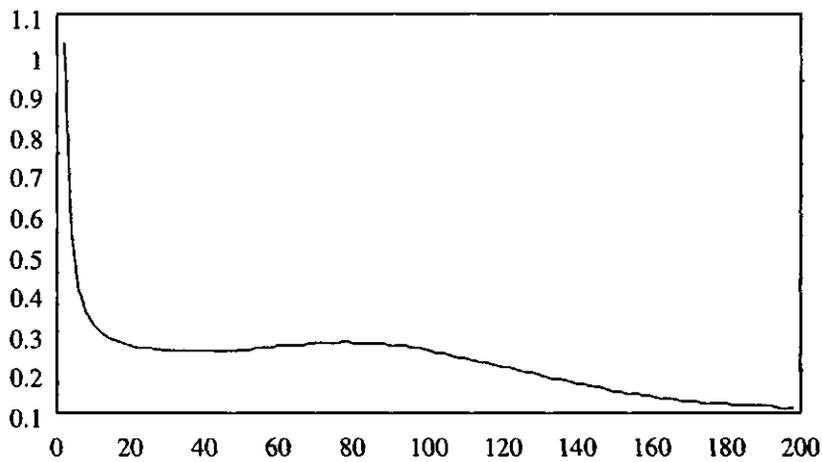


Figure 2.20: The MVRF function of the fractal process

Pure fractal behaviour

The standard description of (pure) fractal behaviour (for time series) :

*if one zooms in in time, one sees
-after rescaling the y-axis-
(statistically) the same time series*

can be translated for the variance by :

$$\text{MVRF}(T) \approx T^\alpha \tag{2.24}$$

White noise is a special case with $\alpha = -1$. Of course, other values for α can also be exploited.

Figures 2.19 and 2.20 show a (artificially generated) fractal process and its corresponding MVRF function.

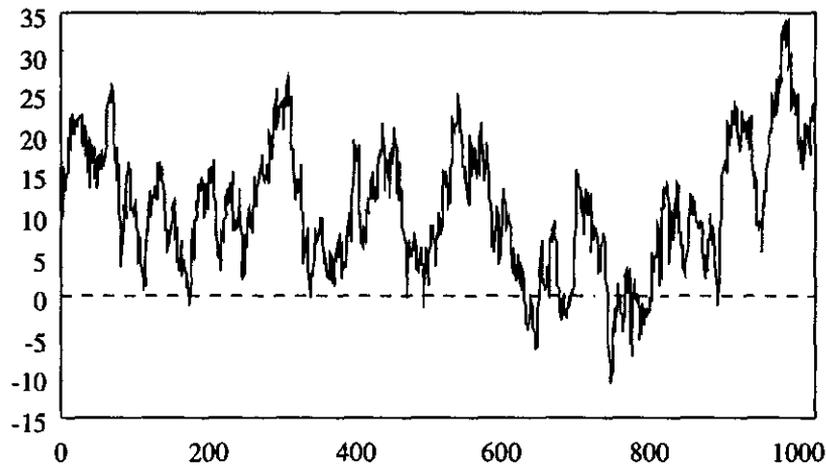


Figure 2.21: An AR(1) process with $\alpha = 0.98$

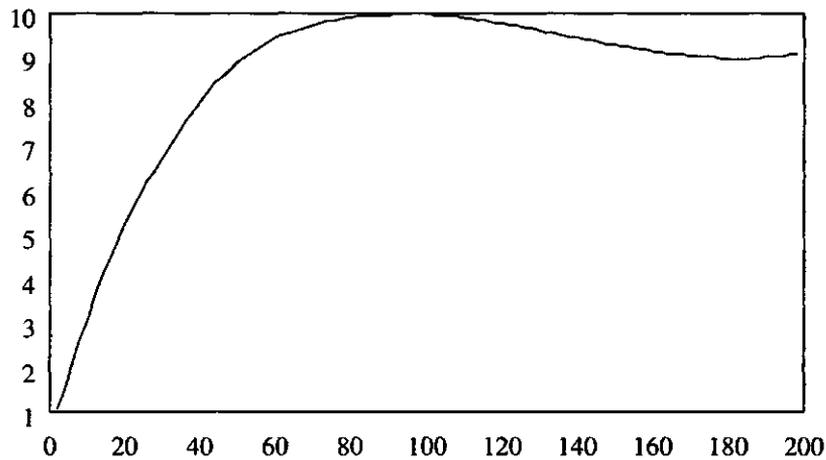


Figure 2.22: The MVRF function of the AR(1) process with $\alpha = 0.98$

Some AR examples

The easiest way to artificially model time series with dependencies is to use AR(1)-models, discretely :

$$X(n) = \alpha X(n-1) + \varepsilon(n) \quad (2.25)$$

where $\varepsilon(n)$ is white noise.

Figure 2.21 shows a time series generated with $\alpha = 0.98$. Figure 2.22 shows the corresponding MVRF function. Clearly this function has a form which is very similar to the Andernach data, see figure 2.8.

Figure 2.23 shows an AR(1) series generated with $\alpha = 0.5$. The MVRF function of this process, as shown in figure 2.24 is a kind of mixture of the preceding AR case and the white noise case, the latter being dominant.

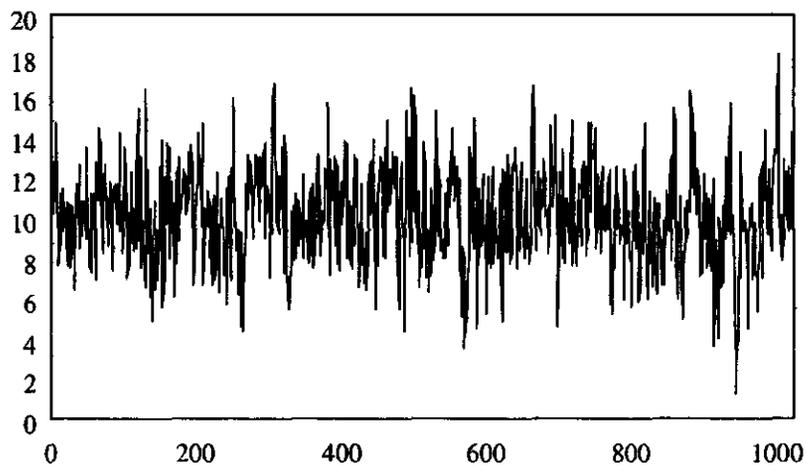


Figure 2.23: An AR(1) process with $\alpha = 0.5$

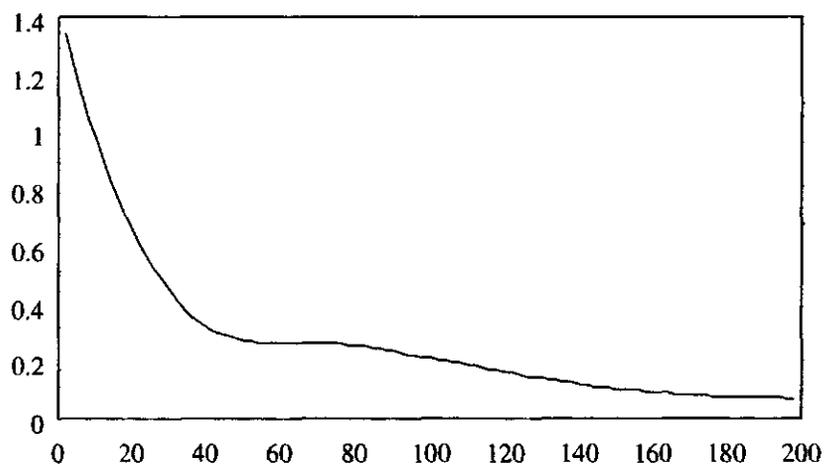


Figure 2.24: The MVRF function of the AR(1) process with $\alpha = 0.5$

2.3 Use of VRF and MVRF in disaggregation

It should be clear from the above, that the most important part of the MVRF function (and the VRF function) is the part near the origin. So a sentence as "these two MVRF's have the same form" in the following paragraphs should be interpreted as : "these two MVRF's have the same form near the origin".

2.3.1 Stationarity of VRF and MVRF

In all calculations above, the stationarity of the time series and in particular the stationarity of the VRF and MVRF functions was tacitly assumed.

In all practical situations, such a stationarity should be checked. In *theory* the stationarity of the VRF and MVRF functions are a direct consequence of the total stationarity of the underlying time series. However, the *total stationarity* of a data series is never (and cannot be) checked. One usually only checks stationarity of the mean and second order moments (variance and covariance). Then, an extra check on particular stationarity of the important functions VRF and MVRF should explicitly be made.

Because the VRF and MVRF functions (specially near the origin) are more sensitive to extremes than e.g. the mean, one usually needs longer series to test stationarity. Calculations of VRF and MVRF functions for 5 year intervals of discharges of the Rhine at Lobith did show rather large differences. Changing the interval length to 10 years however showed already a greater stationarity.

Besides trends, periodicity is another source of non-stationarity. For the hydrological series studied here, a different VRF and MVRF for the summer and for the winter seasons were calculated and treated separately afterwards. Specially the winter functions were convincingly stationary.

2.3.2 Stationarity under climate change

In the paragraph above, the stationarity of the VRF and MVRF of data series was discussed. Because the more far reaching aim of the present study is to investigate peak discharges under climatic change, one should also consider the stationarity of the VRF and MVRF function under this climatic change.

Such a question can only be answered if one has two time series for the same catchment, one before and one after climatic change. To produce such series, one needs models that can calculate discharge series under different climatic conditions in the same catchment *based on physical principles*.

A few preliminary calculations were performed for the Vecht catchment, where such a model is available. The results are shown in figures 3.13 and 3.14. These calculations show that the VRF and MVRF change under different climatic scenarios, but that *the form of the MVRF function is relative stable*.

It is the belief of the authors that the climate changes most dramatically the behaviour at larger time scales (e.g. due to change in snowmelt). Changes due to higher variability of the rainfall may change the height of the MVRF function, but -at least in the Vecht case- much less the form.

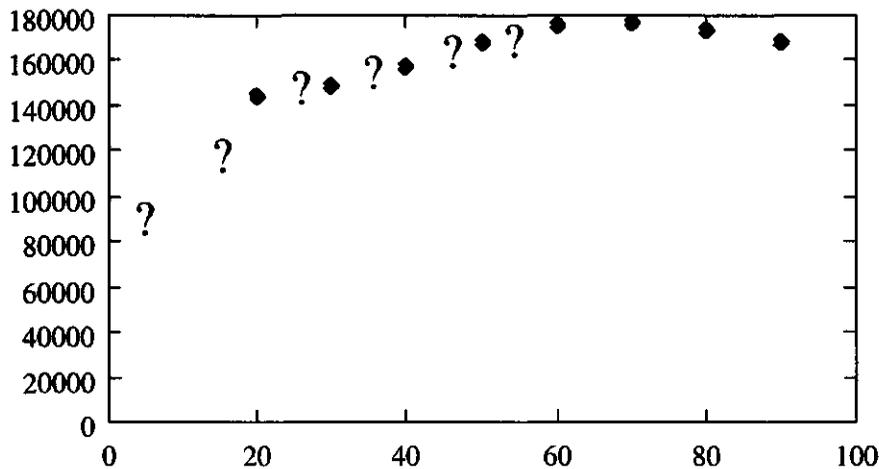


Figure 2.25: The sampled MVRF of the Andernach data

2.3.3 Extrapolating VRF and MVRF

It will seldom be the case that the VRF(T) and MVRF(T) function are available for the whole range of T 's of interest. In many cases one has only partial information.

The most common form of this partial information is that one has only data down to a certain aggregation level. E.g. one has only decade values and no values on a finer time scale.

This means that the VRF(T) and MVRF(T) function are only known on a finite grid. Formally, if T_{min} is the finest scale for which aggregated data are available, only VRF(nT_{min}) and MVRF(nT_{min}) for $n = 0, 1, 2, \dots$ can be calculated. Figure 2.12 shows such a situation.

In the case of the RHINEFLOW-2 model, an 10-day aggregation will be available. The MVRF calculated on basis of these data will give only values on the grid $T = 20, 30, 40, \dots$ Figure 2.25 shows the MVRF of the Andernach data (see figures 2.1 en 2.8) on the same grid.

This creates two kind of problems : and *interpolation* problem (derive values between two grid points) and *extrapolation* problem (derive values for values of T smaller then T_{min}).

For the extrapolation problem, one certainly needs also knowledge of how the function VRF and MVRF can behave. Without this, completely wrong conclusion can be drawn, as the example of the sampled rainfall MVRF function in figure 2.12 showed. This "knowledge of behaviour below T_{min} " should come from induction based on other, similar time series where information on a lower aggregation level is available.

Suppose e.g. that one has a model that calculates decade discharges under different climatic scenarios. In order to say something about the behaviour of the VRF and MVRF function below this level, one certainly needs some characteristic examples of time series below this decade level, e.g. on daily level.

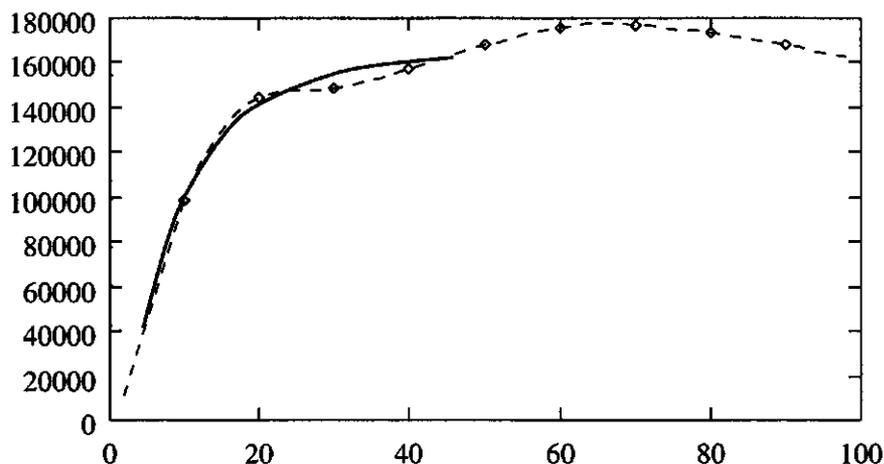


Figure 2.26: A possible extrapolation of the sampled MVRF of the sampled Andernach MVRF, the dashed line is the "true" MVRF

Hopefully, one can then derive from this enough "general behaviour" for the extrapolation of the VRF and MVRF function in those cases where only decade values are available.

2.3.4 Need for parametric models

A very concrete way to do that extrapolation would be by means of a *parametric model*.

The example of figure 2.25 clearly shows the need for such a model.

Suppose that one could show, based on many characteristic series, that a typical behaviour for small T of the MVRF can be described by a simple parametric model :

$$\text{MVRF}(T) \approx \varphi(T, \alpha_1, \alpha_2, \dots) \text{ for small } T \quad (2.26)$$

then one could use this parametric model in those cases where only information down to T_{min} is available to extrapolate from the calculated values the behaviour for much smaller T .

To make a much to simple example :

Suppose one has convinced one self that for discharges of catchments in the Rhine basin a typical behaviour of the MVRF function for small T (say down to a level of 12 hours) is of the form $\text{MVRF}(T) \approx T^\alpha + \beta$.

Suppose that now that one has for a certain catchment calculated decade values (e.g. under a different climate). Then one can calculate the MVRF function for values of $T = n[\text{days}]$. Using this, one can try to find an α and β such that $T^\alpha + \beta$ fits reasonable through the most leftwards points of this function, and use this parametric form to extrapolate the function down to a level of half a day.

Figure 2.26 shows a possible extrapolation of the sampled MVRF of the Andernach data.

The crucial point in the above scheme is that one should have a parametrisation that fits reasonably all series that are considered to be representative. In the present stage of the study, this has yet to be done.

The parametric models should be able to summarise “the form” of the MVRF function in a few parameters. Because it is to be expected that this form (at least near the origin) is rather stable, these parameters could be thought then to be insensible to climatic change and thus form the basis of the disaggregation in different climatic scenario's.

Chapter 3

Application for discharges in the Rhine catchment

In this section, some pictures will be represented for discharges in the Rhine catchments. Data were splitted up in two parts : the summer part (months May through October) and the winter part (months November through april).

3.1 Difference between summer and winter

Figures 3.1 and 3.2 show the Rhine at Stilli (Switzerland) respectively for the summer and the winter period. The difference in statistical behaviour of both series shows itself clearly in their MVRF functions. The winter MVRF has a form very similar to the Andernach case (see figure 2.8).

Because our main interest lies in the winter period, and because the summer MVRF's can differ significantly from the winter ones (and are in general less similar), most examples in the rest of the paragraph will be of the winter period.

3.2 Along the Mosel

Figure 3.3 shows the Saar in the upper part of its catchment (winters 1978-1990).

Figures 3.4 and 3.5 show the Saar at Fremersdorf, i.e. at the end of its catchment respectively for the summer and winter (years 1978-1990). Clearly again, the summer part has a "non-typical" form. The winter part is more regular, like in figure 2.8.

Figure 3.6 shows the Mosel at Cochem (winters 1978-1990). Here the behaviour of the MVRF is very regular.

3.3 Along the Rhine

Figure 3.7 shows the Rhine in Neuhausen -just downstream of the Bodensee-, figure 3.8 at Kaub and figure 3.9 Lobith (all for the winters 1978-1990).

The MVRF function of the Rhine at Neuhausen is a-typical (the Bodensee damps the fluctuations), the MVRF's at Kaub and Lobith show very typical behaviour.

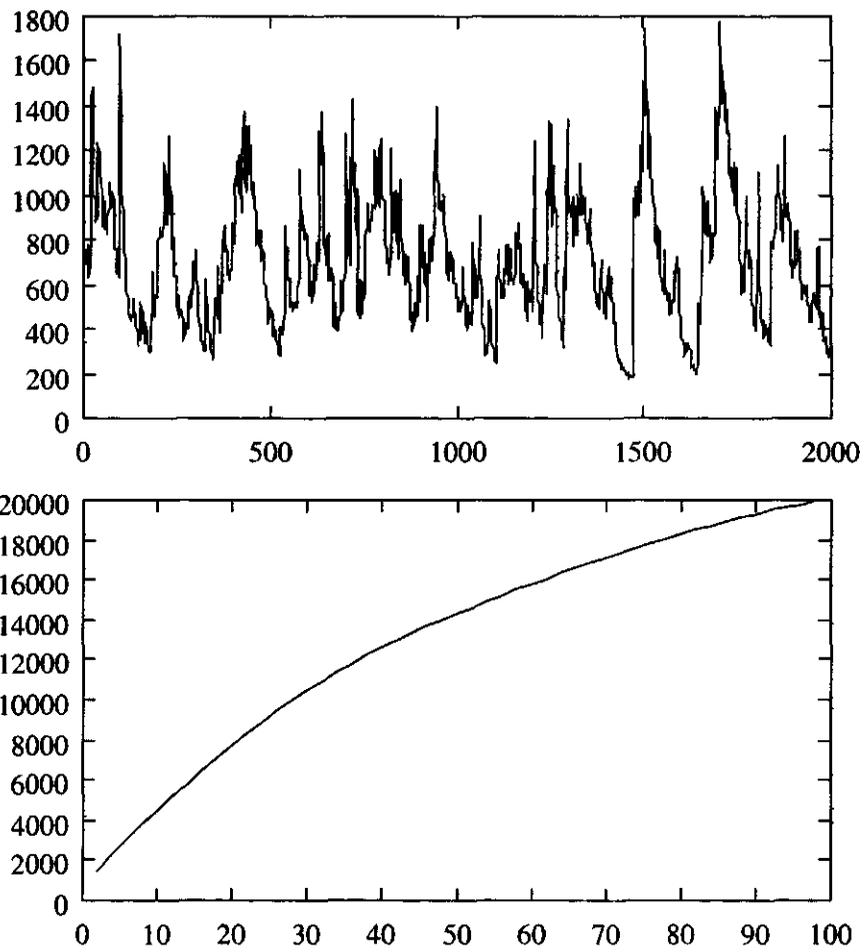


Figure 3.1: The Rhine at Stilli in the summer

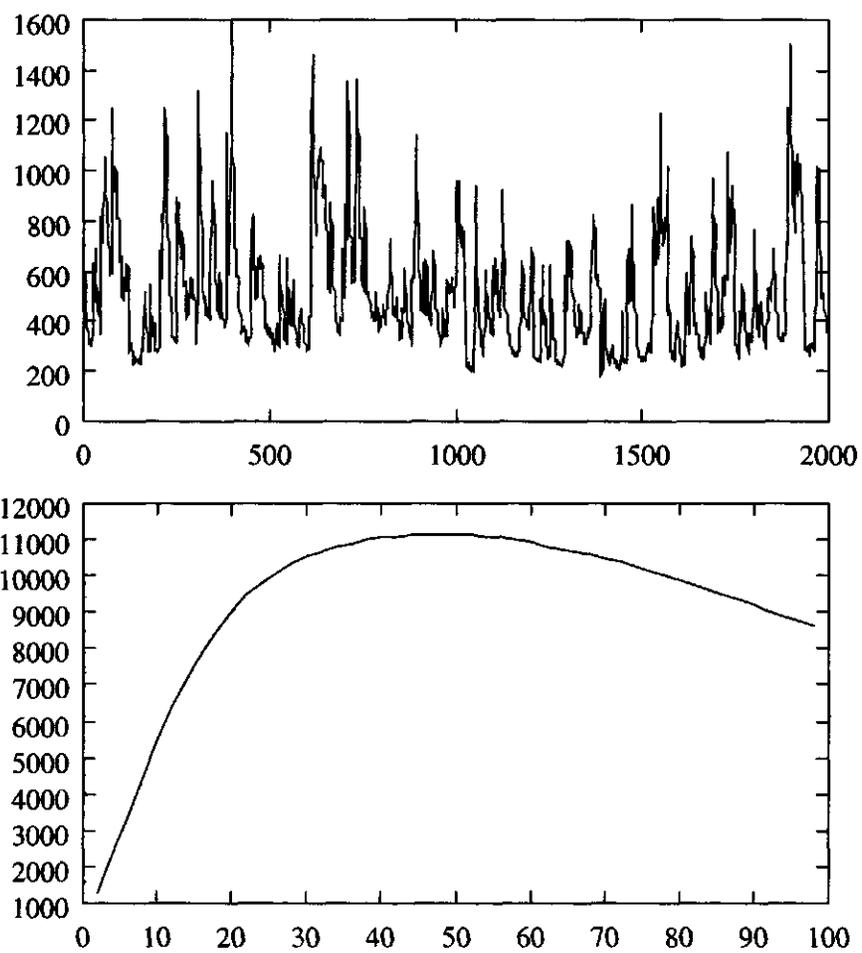


Figure 3.2: The winter part of the Rhine at Stilli

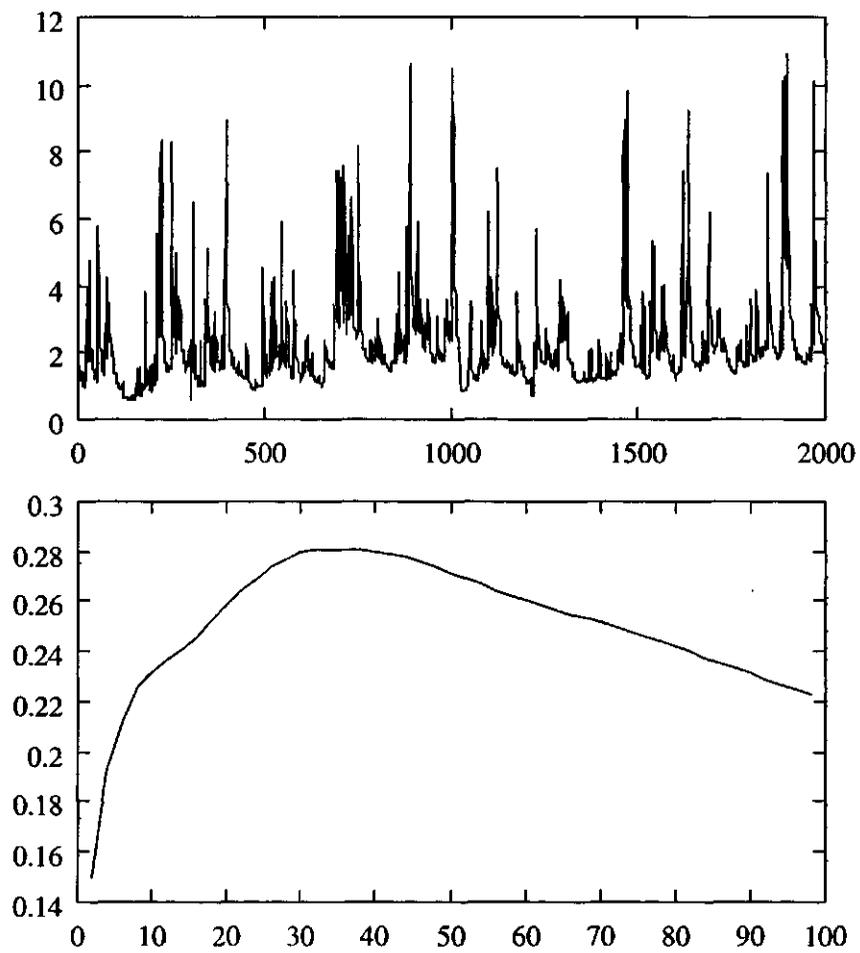


Figure 3.3: The Saar in Laneuveville in winter time

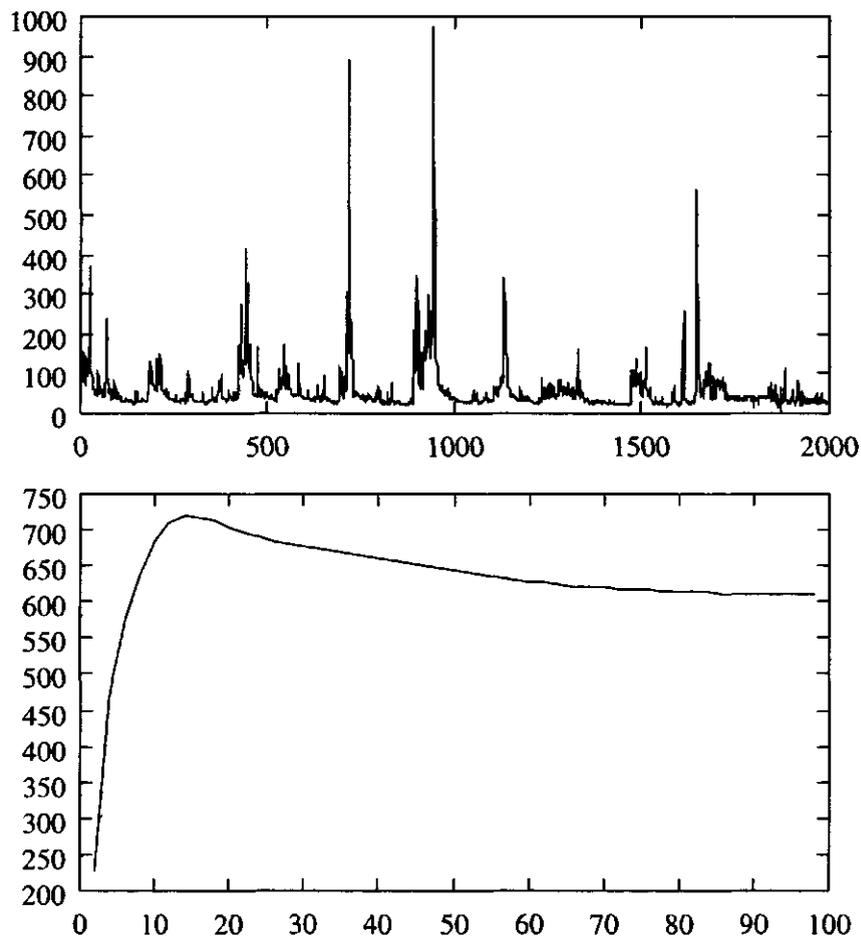


Figure 3.4: The Saar at Fremersdorf in the summer

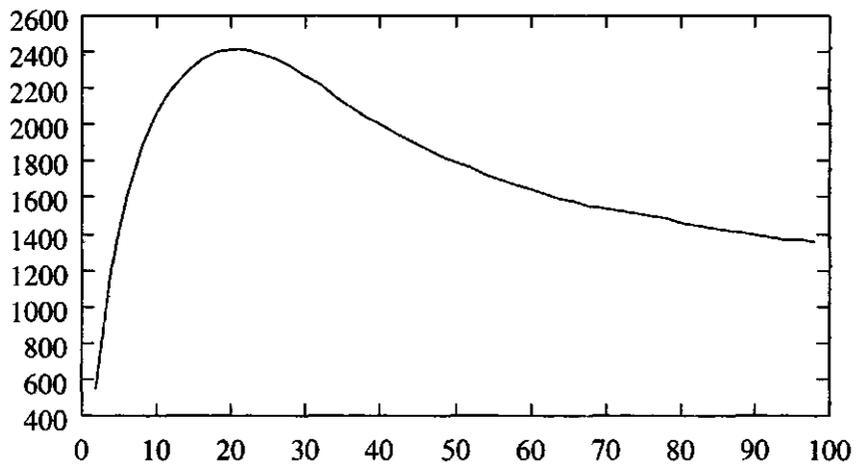
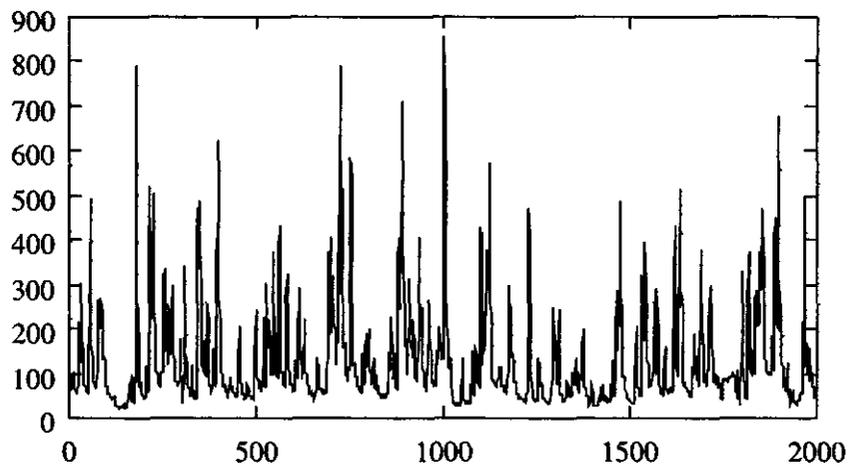


Figure 3.5: The Saar at Fremersdorf in the winter

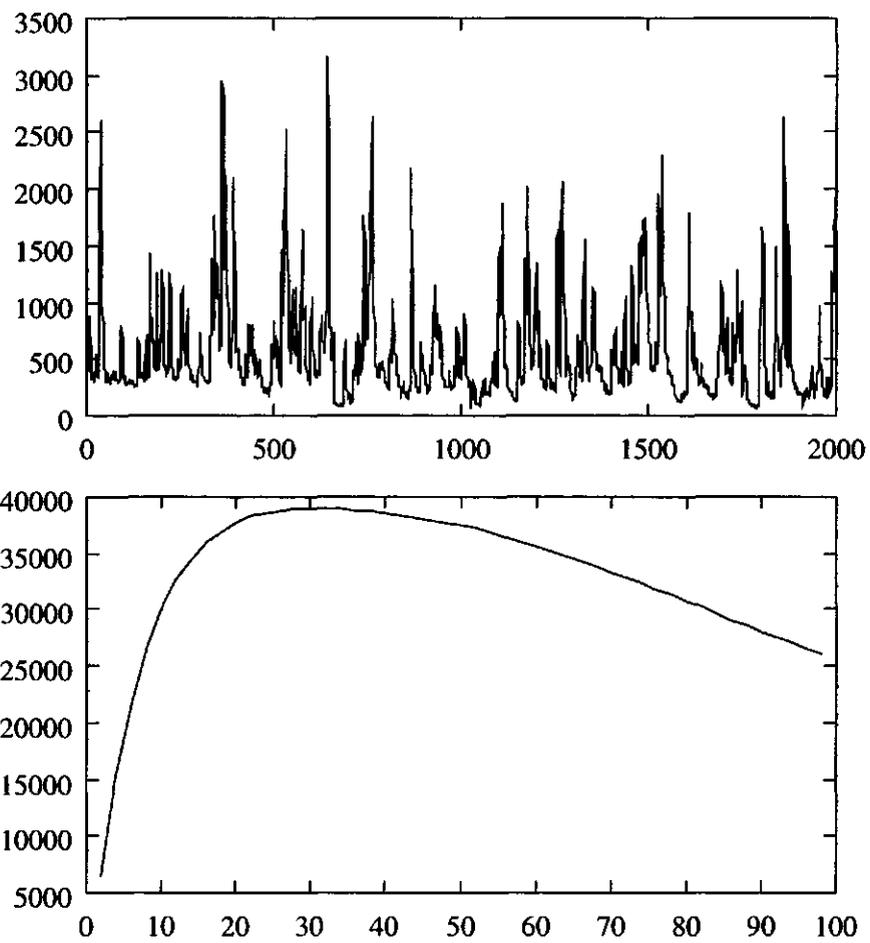


Figure 3.6: The Mosel at Cochem in the winter

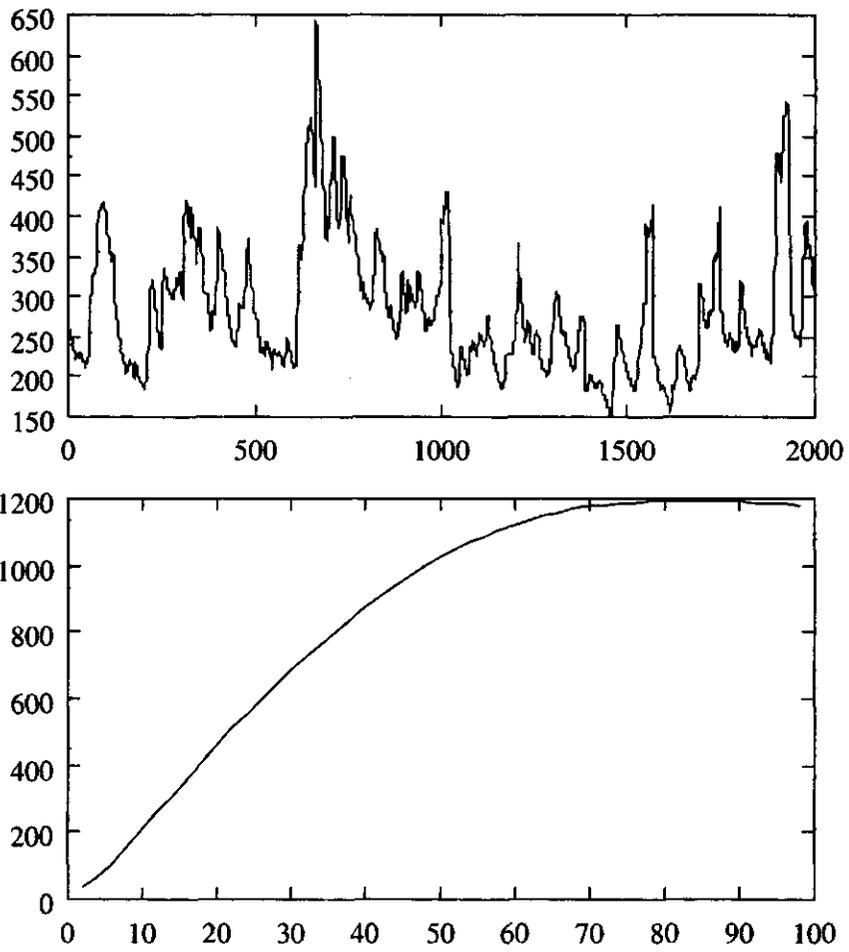


Figure 3.7: The Rhine at Neuhausen in the winter

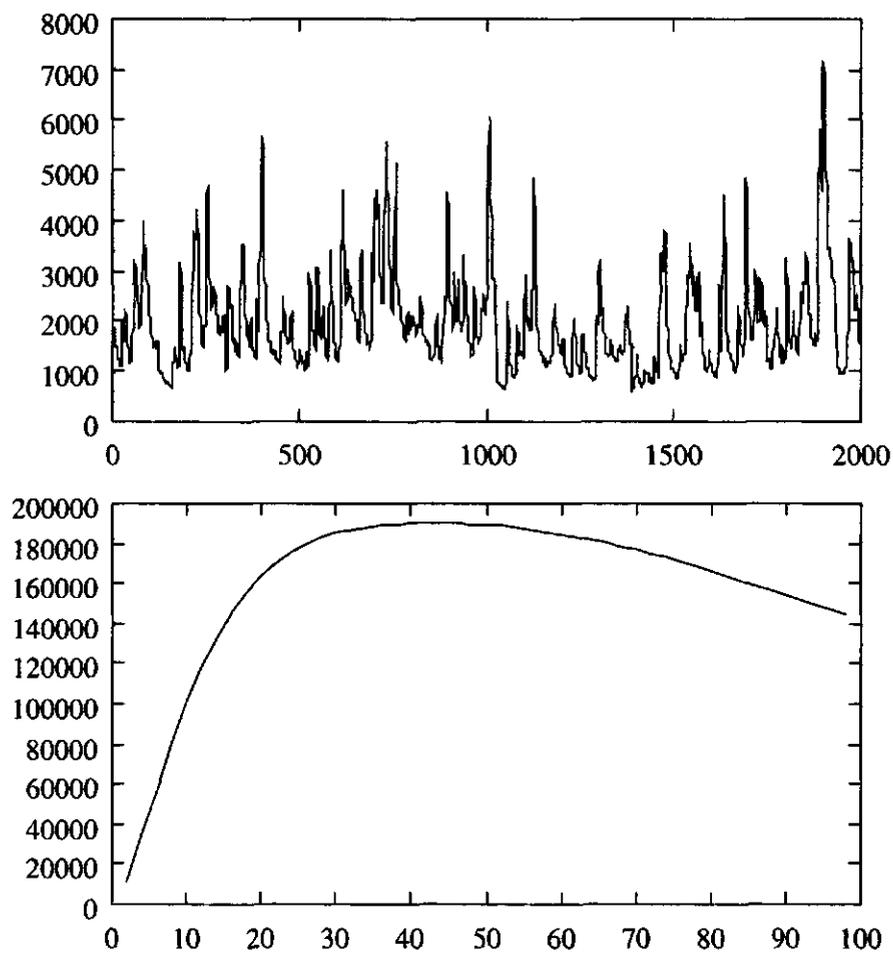


Figure 3.8: The Rhine at Kaub in the winter

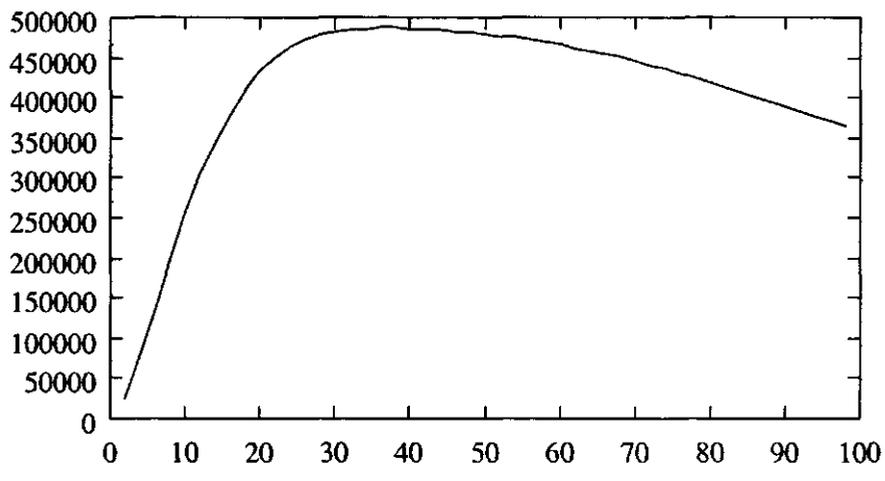
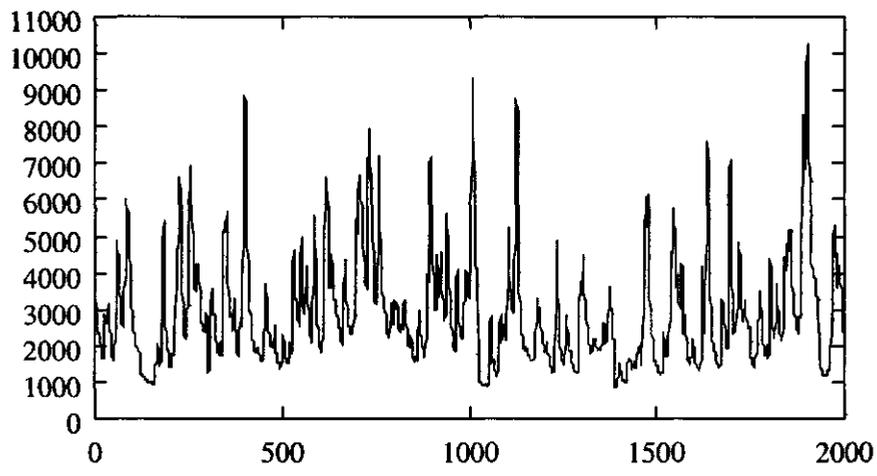


Figure 3.9: The Rhine at Lobith in the winter

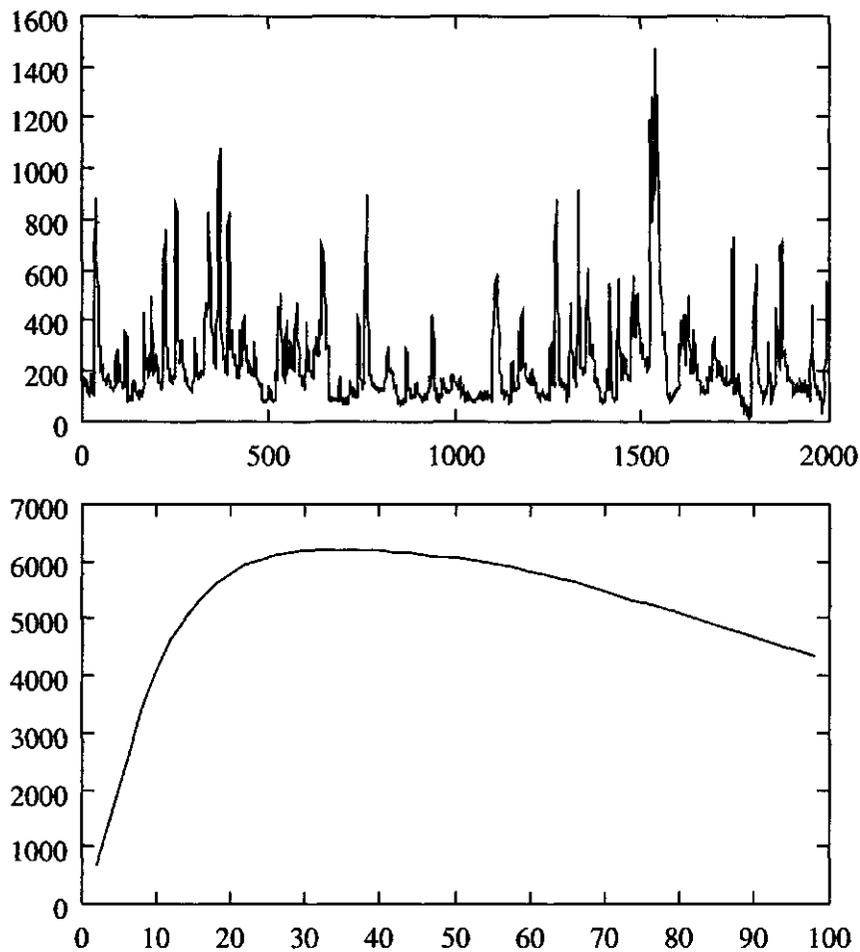


Figure 3.10: The Main at Obernau in the winter

3.4 Some tributaries of the Rhine

In this section, some other tributaries (other than the Mosel) are presented : the Main at Oberanau in figure 3.10, which shows typical winter behaviour, the Neckar at Gundelsheim in figure 3.11, which shows a non-so-typical behaviour (the Neckar at this place is under influence of a backwater curve of a downstream control), and the Lippe a Schermbeck in figure 3.12, again more typical (all for the winters 1980-1992).

3.5 The Vecht under different climates

For the Vecht catchment, a physical based model that can calculate discharges under different climate scenarios is available ¹.

Figure 3.13 shows the Vecht (winters 1978-1990) under the present climate.

1. Parmet & Bouma, 1995

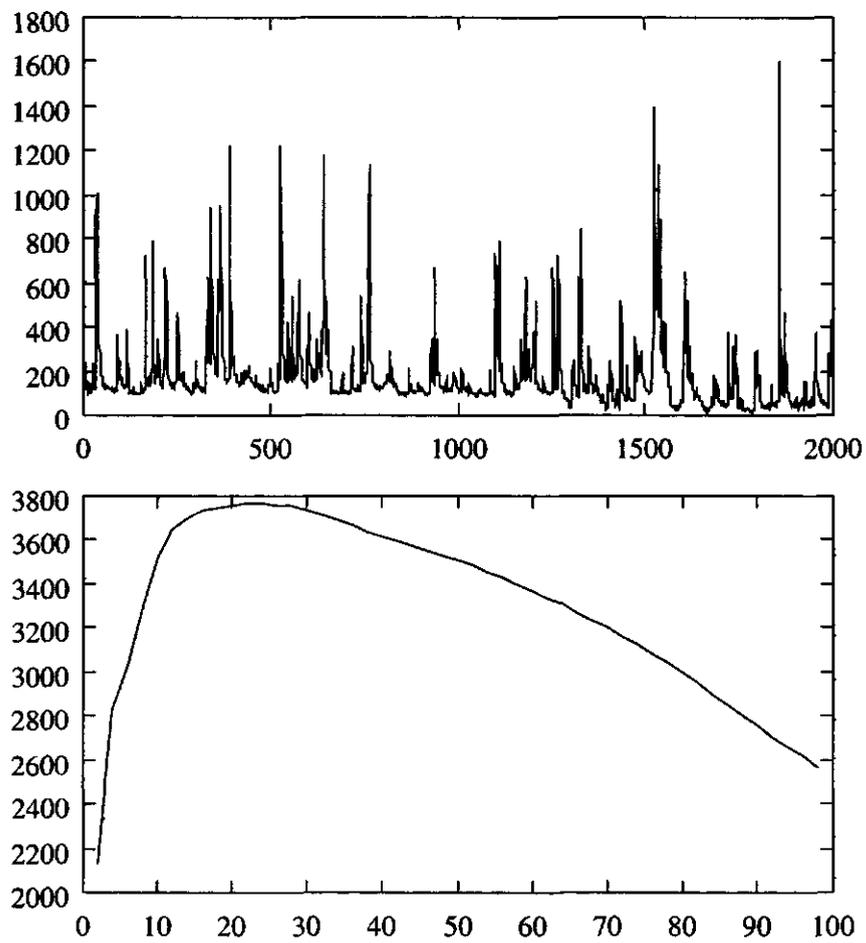


Figure 3.11: The Neckar at Gundelsheim in the winter

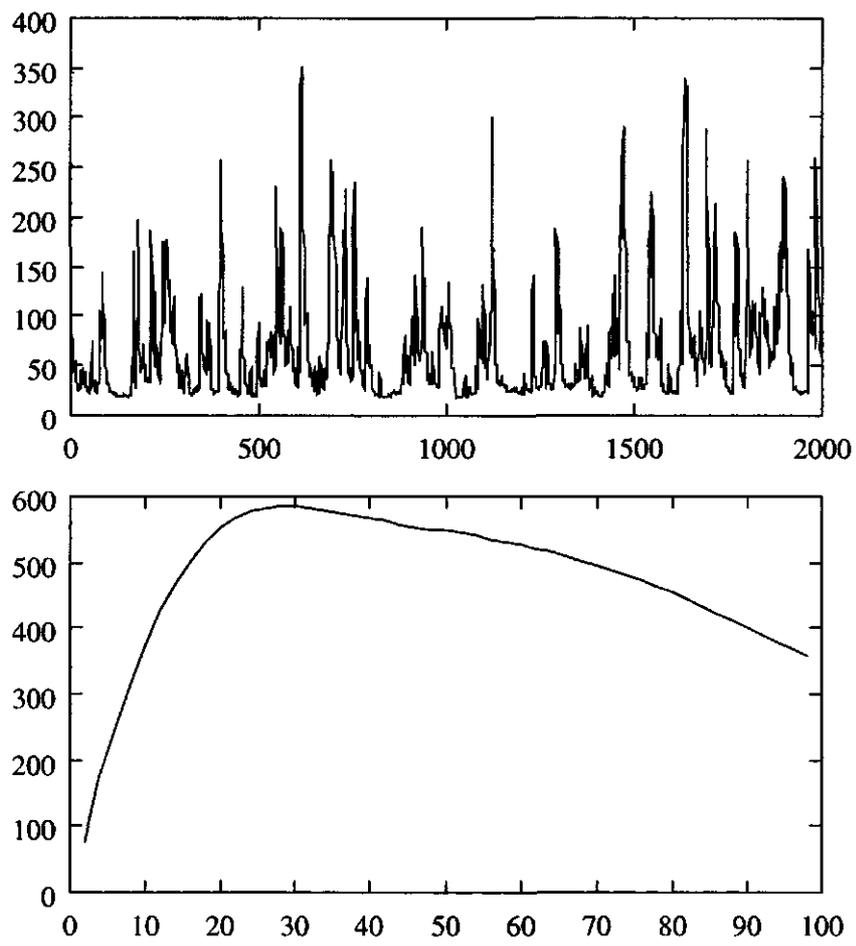


Figure 3.12: The Lippe at Schermbeck in the winter

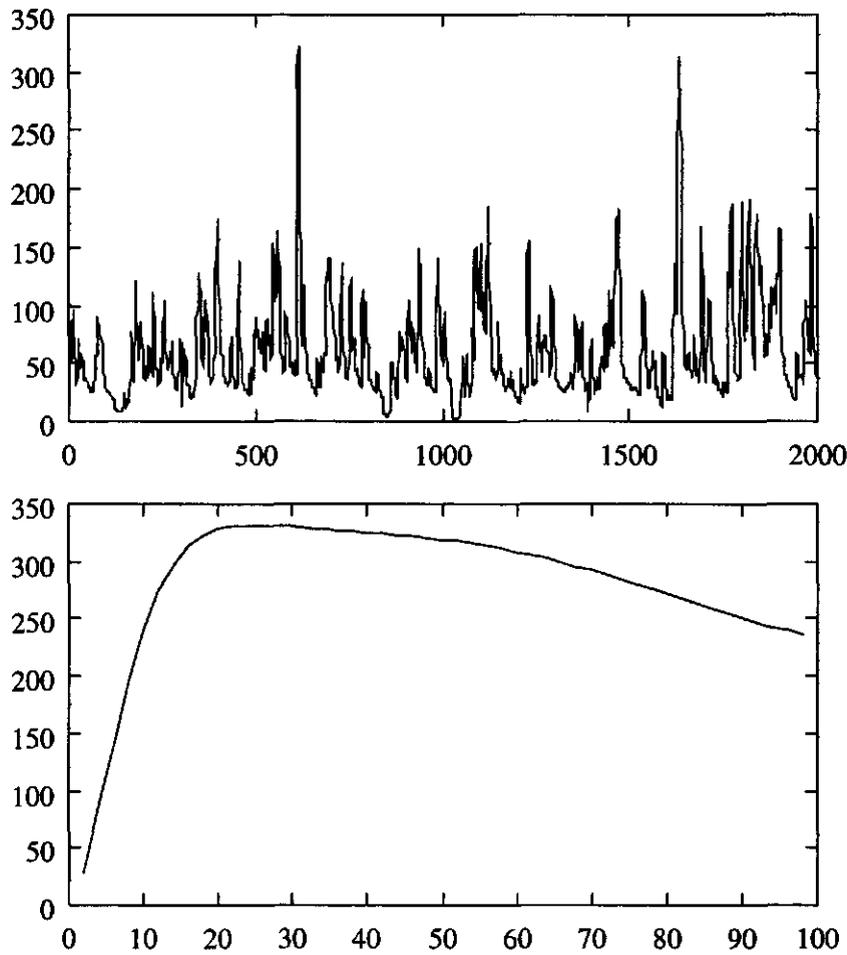


Figure 3.13: The Vecht in the winter

Figure 3.14 shows the same Vecht under the rather extreme “UKHI2100 scenario”². The range of the discharges almost doubled. The “form” of the two MVRF functions is very similar, although the ranges differ by a factor 2.

2. UKHI (Hulme et al, 1994). This scenario is characterised by a strong increase in winter precipitation in the winter half year (up to 45%), and a minor decrease of precipitation in the summer period (at most -5%). The annual temperature rise is about 4 degrees Celsius.

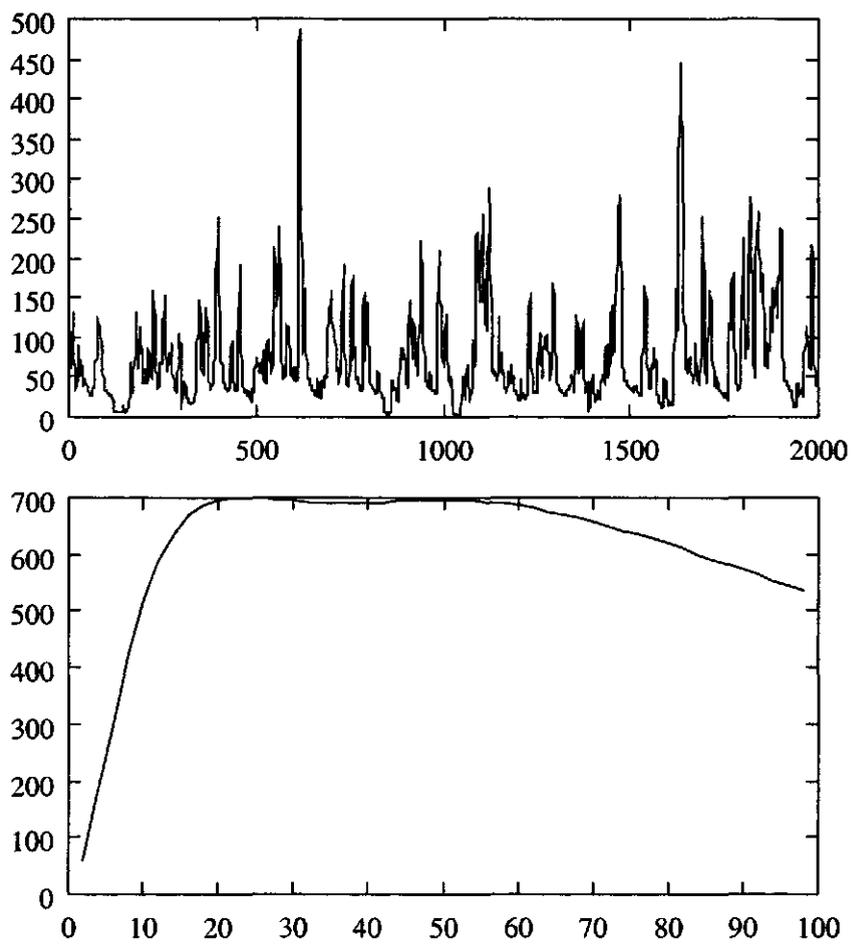


Figure 3.14: The Vecht in the winter after climate change

Chapter 4

After variance : covariance

4.1 The MACF and MCCF functions

After mean and variances, the covariances and the correlations –which can be directly calculated from the covariances– are the most important characteristics of a time series.

So, if (and **only if**) one has successfully disaggregated the variance, one can try to disaggregate the covariance.

For halving the disaggregation length, one needs the following covariances :

$$\text{MACF}(T, k) = \text{COV} \left(X_{[T]}(n), X_{[T]}(n+k) \right) \quad k = 0, 1, 2, \dots \quad (4.1)$$

$$\text{MCCF}(T, k) = \text{COV} \left(X_{[T]}(n), X^{[T]}(n+k) \right) \quad k = \dots, -1, 0, 1, \dots \quad (4.2)$$

The acronym MACF stands for Marginal Auto Covariance Function and MCCF for Marginal Cross Covariance function.

Technically, it is not difficult to calculate these functions. In general, the same procedure as for the MVRF should be followed :

- first by investigating a lot of characteristic examples, sufficient confidence on stationarity should be obtained
- a simple but convincing parametrisation should be formulated, tried and valitated

4.2 The "all independ" choice

One can, of course, make a trivial choice for the MACF and MCCF function :

$$\text{MACF}(T, k) = 0 \text{ for } k \neq 0 \quad (4.3)$$

$$\text{MCCF}(T, k) = 0 \text{ for all } k \quad (4.4)$$

If one accepts this, a disaggregation simulation is relatively easy. To go from T to $T/2$:

1. one generates independent random (gaussian) variables with mean 0 and variance $MVCF(T)$. These form the process $X_{[T]}$;
2. one uses the disaggregations formulas to calculate $X_{[T/2]}$.

If needed, this procedure can be repeated.

The figures 4.1 and 4.2 illustrate the recursive procedure described above. The end result is shown in figure 2.19. The independence makes that the "binary grid-points" remain visible. The processes generated in this way are thus not stationary. In this sense, they have a very unwanted behaviour. They may have however the right "extreme" behaviour, or some other correct statistical characteristic.

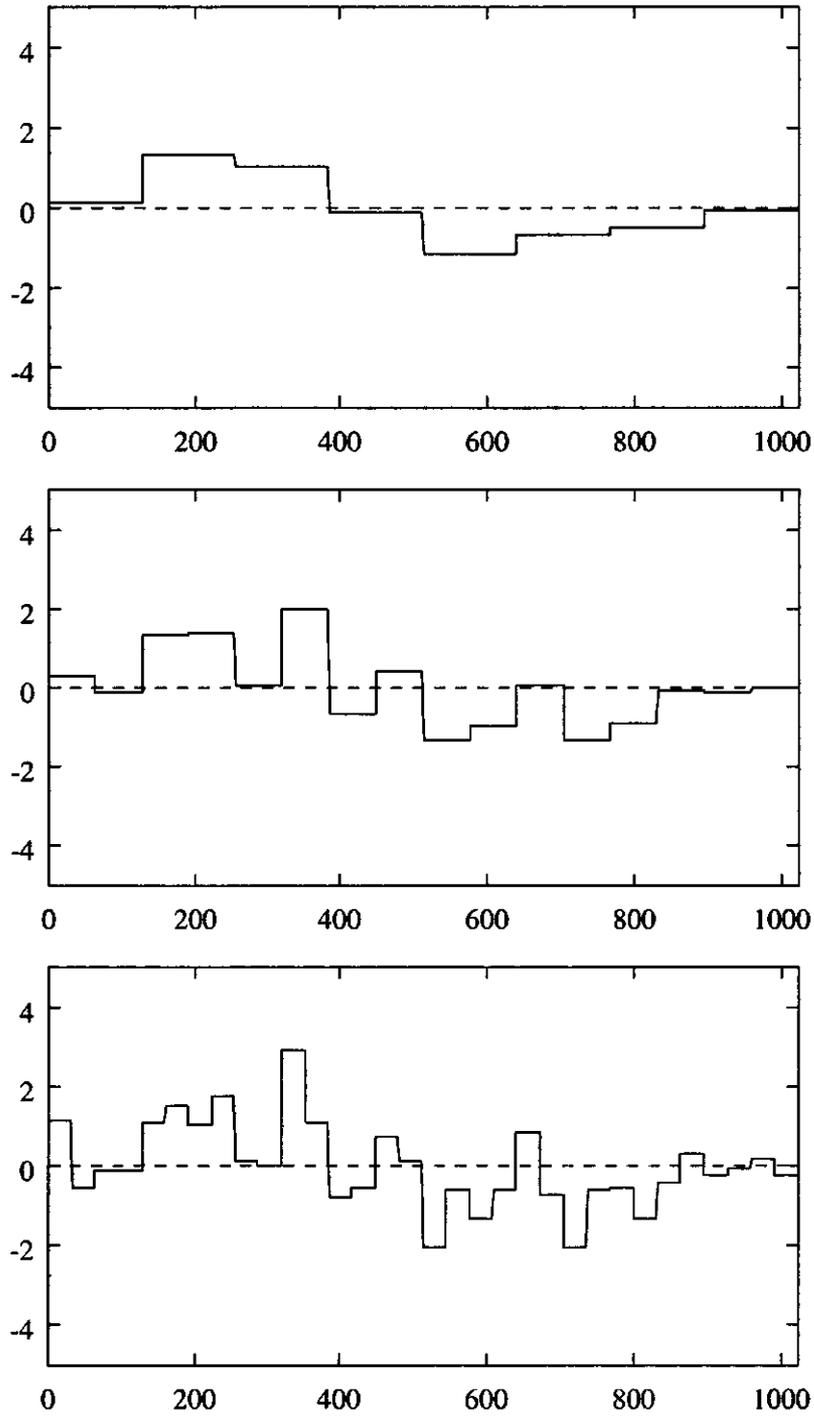


Figure 4.1: The first three stages of the generation process.

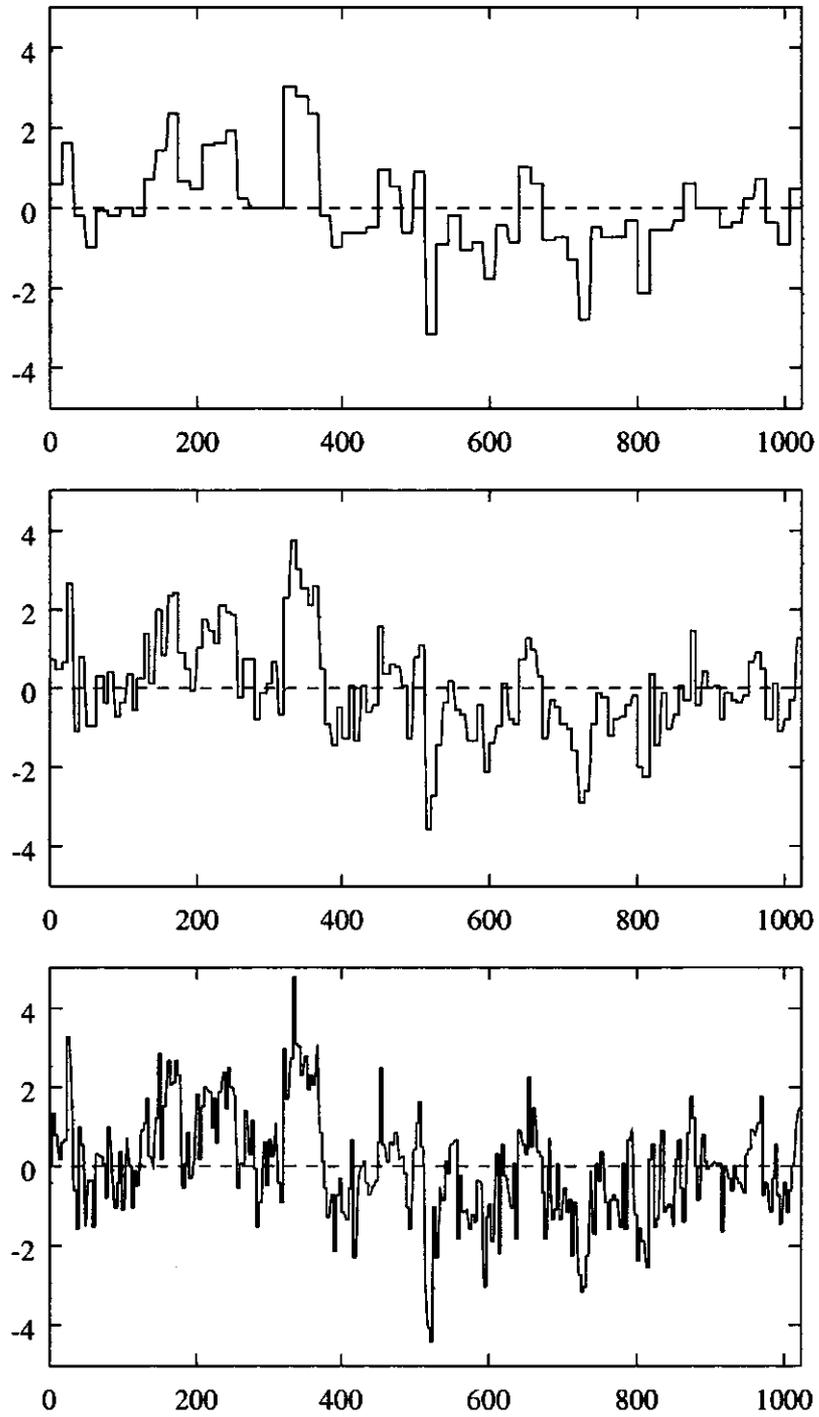


Figure 4.2: The next three stages of the generation process.

Chapter 5

Conclusions and continuation

5.1 Conclusions

1. The MVRF function is the basic first step in disaggregation analysis.
2. At this stage of the study, the necessary software to calculate the important statistics is available in a functional form.
3. A lot of relevant examples have been calculated.
4. The examples do not show White Noise (or Fractal Noise) behaviour.
5. The form of the MVRF for the winter is sufficient stable (for series longer than 10 years and for catchments larger than 1000 km²) to draw the conclusion that successful disaggregation of the variance function is possible for different catchments from the decade scale down to the day scale.
6. The form of the MVRF in summer is less stable, certainly for smaller catchments.
7. The strongest decrease of the winter MVRF's is below 10 days. The new Rhineflow model will provide us data down to a scale 10 days. This means that MVRF constructed with Rhineflow-2 data will stop at day=20. The winter-MVRF's show that Rhineflow-2 models pin down the total MVRF to an acceptable precision.
8. Smaller catchments are in general difficult to model (specially in the summer)
9. A few first examples show that the form of the MVRF is also stable under different climates. This should further be investigated
10. It is relatively easy to simulate disaggregation if the MVRF is known and one accepts independencies.

5.2 What to do next

5.2.1 Stability under climate change

More examples of series on small time scale (day) of the same catchment under different climatic conditions should be investigated. The hypothesis whether the form of the MVRF(T) for small T is stable should be tested.

5.2.2 Proposal for parametrisation

A parametrisation that is rich enough to fit all the examples considered to be representative, small enough to be easily fitted, should be developed.

At this stage of the study, such a parametrisation is not available.

If no better suggestion is found, parametric forms derived from simple AR (or ARMA) models could be tried.

5.2.3 Simulation with independencies

If the preceding two points are successfully finished, one can try to stochastically disaggregate (i.e. to simulate) by *neglecting* all dependencies, i.e. by assuming $MACF \equiv 0$ and $MCCF \equiv 0$. A first description of this technique can be found in paragraph 4.2.

5.2.4 Criteria for evaluation

Criteria for evaluation should be developed. Certainly at this stage, the mean and the variance are correctly simulated.

A criterium is a statistic that can be calculated on the simulated disaggregated series.

Examples of such statistics are :

1. high quantiles (e.g. 95% quantile);
2. duration of level exceedence
3. return periods for high levels calculated by classical extreme (Gumbel?) statistics

Appendix A

Proof of second interpretation of MVRF

The proof of the "second" interpretation of the MVRF function goes as follows. Using equations 2.2 and 2.4 one can write :

$$\begin{aligned} \mathbf{E} \left[X^{[T]}(n) X_{[T]}(n) \right] &= \frac{1}{4} \left\{ \mathbf{E} \left[\left(X^{[T/2]}(2n-1) \right)^2 \right] - \mathbf{E} \left[\left(X^{[T/2]}(2n) \right)^2 \right] \right\} \\ &= 0 \end{aligned}$$

this last equality following from stationarity.

With the help of this, one can show that $X^{[T]}(n)$ and $X_{[T]}(n)$ are uncorrelated :

$$\begin{aligned} \text{COV} \left[X^{[T]}(n) X_{[T]}(n) \right] &= \mathbf{E} \left[\left(X^{[T]}(n) - \mu \right) X_{[T]}(n) \right] \\ &= \mathbf{E} \left[X^{[T]}(n) X_{[T]}(n) \right] - \mu \mathbf{E} \left[X_{[T]}(n) \right] \\ &= 0 \end{aligned} \tag{A.1}$$

From this, the second interpretation of the MVRF is easily derived :

$$\begin{aligned} \text{VAR} \left[X_{[T/2]}(n) \right] &= \text{VAR} \left[X^{[T]}(n) + X_{[T]}(n) \right] \\ &= \text{VAR} \left[X^{[T]}(n) \right] + \text{VAR} \left[X_{[T]}(n) \right] + 2\text{COV} \left[X^{[T]}(n) X_{[T]}(n) \right] \\ &= \text{VAR} \left[X^{[T]}(n) \right] + \text{VAR} \left[X_{[T]}(n) \right] \end{aligned}$$

Appendix B

An example of a typical set of computations

All computations were done by mean of c-code. A simple interface was build and used to call these functions in different ways. A characteristic examples, showing typical behaviour and computations to be performed, follows. In this, the following functions are used :

cent centralises the data, e.g. substracts the mean

TtoF transforms (with FFT) a time series to its Fourier representation : form the time domain to the frequency domain

FotT performs an inverse Fourier transform (with FFT) : from frequency domain to the time domain;

newfil generates a filter (see equations 2.14 and 2.15)

convFF makes the product of two complex series (see formulas 2.16 and 2.17)

vf calculates the VRF and MVRF function

```
gnufilename = demo.gnu
```

```
newvar(datafilename) = demo.dat  
newvar(months)       = 1,12  
newvar(years)        = 1941,1943
```

the following lines read the data and plot them

```
newvar(data) = getQdat(datafilename,months,years)  
@plot(data,str(original data from,datafilename))
```

centralise

```
data = cent(data)
```

calculate and plot the Fourier transform of the data

```
newvar(Fdata) = TtoF(@complex(data))  
@plot(mir(Fdata),'Fourier transform, xmir(@len(Fdata)))
```

the filters :

```
newvar(fil) = newfil(@len(data),32,low)
```

```
@plot(fil, 'a low filter')
fil = newfil(@len(data),32,high)
@plot(fil, 'a high filter')
rm('fil')
```

to filter the data is done by using the Fourier transform of the filters :

```
low :
newvar(level) = 32
newvar(Ffil) = TtoF(newfil(@len(data),level,low))
newvar(fildata) = FtoT(convFF(Fdata,Ffil))
@plot(fildata,de laag gefilterde data)
```

```
high :
Ffil = TtoF(newfil(@len(data),level,high))
fildata = FtoT(convFF(Fdata,Ffil))
@plot(fildata,de hoog gefilterde data)
```

```
rm('fildata')
```

calculation of the VRF and MVRF functions :

```
newvar(datavf) = vf(data,100,low)
@plot(datavf, 'the VRF function')
```

```
datavf = vf(data,100,high)
@plot(datavf, 'the MVRF function')
rm('datavf')
```

```
closegnu
```

```
quit
```

Appendix C

The fractal generation code

```
for (i=0; i<numlevels; i++)
{
    x[i] = 0;
    level[i] = 2;
    halflevel[i] = 1;

    for (j=0; j<i; j++)
    {
        level[i] *= 2;
        prevlevel[i] *=2;
    }
}

for(i=0; i<numdata; i++)
{
    fprintf(fout,"%d ",i);
    result = 0.0;
    for (j=levels-1; j>= 0; j--)
    {
        if(i%level[j] == 0) x[j] = MVRF[j] * gauss();
        if(i%halflevel[j] == 0)    x[j] = -x[j];

        result += x[j];
    }
    printf("%f\n",totx);
}
```