Screening of Hydrological Data:

Tests for Stationarity and Relative Consistency

Screening of Hydrological Data:

Tests for Stationarity and Relative Consistency

E.R. Dahmen
M.J. Hall

Publication 49

International Institute for Land Reclamation and Improvement/ILRI
P.O.Box 45,6700 AA Wageningen,The Netherlands,1990

Acknowledgements

The authors, E. R. Dahmen and M. J. Hall, are grateful to the following organizations for kind permission to use the data in this publication:

The Meteorological Department, Thailand
The Port Authority, Thailand
The National Directorate of Water, Mozambique
The Public Works Department, Indonesia
The National Rivers Authority, United Kingdom
The Department of Public Works and Water Control, The Netherlands
The Hydraulic Department, Cambodia

ISBN 90 70754 231

Printed in The Netherlands

Contents

Preface

The availability of mainframes and personal computers has increased phenomenally over the last two decades. As a result, methods of data analysis that were once applied only in exceptional circumstances are now routine. Prominent among these methods are the statistical fitting of frequency distributions, the modelling of rainfall-runoff relationships, and the calculation of water balances, all of which rely heavily on hydrological and hydrometeorological data. At the same time, modern chip technology has revolutionized data collection and enabled the direct logging of hydrometeorological parameters.

Nevertheless, while the collection and analysis of hydrological data are improving, the environment everywhere in the world is being subjected to more and more obtrusive alterations, which can introduce non-homogeneity into data series that span the period of change. Similarly, the modernization of measuring equipment can cause inconsistency to appear in a data series. Therefore, it is ironic that, now, when hydrological data can be transmitted directly from the on-site equipment to the office computer system, increased vigilance is demanded of the engineer to ensure that they are not contaminated by extraneous influences.

A shortage of hydrological data hampers the planning and design of many water development schemes. Fortunately, thanks to noteworthy efforts like the widespread setting up of hydrometeorological stations during the International Hydrological Decade (1965 to 1974) and the subsequent International Hydrological Programme, more and more hydrological data are becoming available.

Our work during the past fifteen years in Southeast Asia, Africa, and South America has confirmed that the screening of hydrological data is a prerequisite to the successful design and implementation of water development schemes. Our experiences in teaching and training have prompted us to refine on the basic data-screening procedure, extend it, and present it in book form.

It took several years to elaborate the complete data-screening procedure. In this work, we received considerable assistance from participants in the post-graduate courses at the International Institute for Hydraulic and Environmental Engineering/IHE in Delft, in the International Course on Land Drainage at the International Institute for Land Reclamation and Improvement/ILRI in Wageningen, and in the courses at the Caribbean Institute for Meteorology and Hydrology in Barbados. It was these course participants who applied the procedure in their group exercises, testing it on numerous data sets, and thereby speeding up its verification.

Elements of the procedure have appeared in many previous works (e.g. in a publication of the World Meteorological Organization/WMO 1966). We make no apology for this apparent lack of originality. Our purpose was to bring together, in a common framework, the disparate details of a group of practical tests. We have not included specialized tests (e.g. those described by Buishand 1982 and Bernier 1977). Instead, we use the very basic tests that are also used in industrial quality control.

While using these tests to screen hydrological data, we found that we could also use them to perform a significance test on breaks in double-mass lines. Accordingly, the data-screening procedure offers an alternative to the analysis of variance, which is commonly advocated for this purpose. The advantages of using the same computational framework for testing absolute and relative consistency and homogeneity speak for themselves.

The easiest way to perform the data-screening procedure is with a dedicated computer program, so, with this book, we have included a floppy disc that contains such a program. It was developed on Acorn BBC and Cambridge computers in BBC-Basic and will run on MS-DOS-compatible machines with a CTA or EGA graphics adaptor. The Acorn BBC version of the program can be purchased from ILRI. If necessary, the engineer can perform the data-screening procedure on a desktop computer that has a simple spreadsheet, or with some squared paper and a pocket calculator capable of computing statistical functions, but we strongly recommend the enclosed program.

Reliable data are at the core of reliable hydrological studies and, therefore, vital to the management of land and water. The extra effort required to screen hydrological data before use is negligible and, we believe, well worth the time, as it will enhance the engineer's insight and understanding.

E.R. Dahmen, Delft
M.J. Hall, London
May 1989

Abstract

Hydrological data for water-management studies should be stationary, consistent, and homogeneous when they are used in frequency analyses or system simulations. A simple but efficient procedure for screening these data is to test annual or seasonal time series for absence of trend and stability of variance and mean. If required, this basic procedure can be extended to include tests for absence of persistence (with the first serial-correlation coefficient) and relative homogeneity and consistency (with double-mass analysis). Applied to proportionality factors, the basic procedure offers an alternative way to evaluate the significance of slope changes found in double-mass lines.

The procedure is illustrated by examples. All formulae needed to perform the tests are presented. Three annexes contain relevant statistical tables and additional problems (with solutions). The book comes with a computer program on floppy disc to allow the user to run the basic tests on a personal computer.

Keywords

Hydrological time-series analysis; stationarity; consistency; homogeneity; persistence; double-mass analysis; stationarity of proportionality factors.

1 Introduction

Engineering studies of water resources development and management depend heavily on hydrological data. These data should be stationary, consistent, and homogeneous when they are used for frequency analyses or to simulate a hydrological system. To determine whether the data meet these criteria, the engineer needs a simple but efficient screening procedure. Such a procedure is described in this book.

A time series of hydrological data is strictly stationary if its statistical properties (e.g. its mean, variance, and higher-order moments) are unaffected by the choice of time origin. (By 'unaffected', we mean that estimates of these properties agree within the range of expected statistical variability.) The basic data-screening procedure presented here is based upon split-record tests for stability of the variance and mean of such a time series. Although stability of these two properties indicates only a weak form of stationarity, this is enough to identify a non-stationary time series (Figure 1.1), or to select those parts of a time series that are acceptable for use.

A time series of hydrological data may exhibit jumps and trends owing to what Yevjevich and Jeng (1969) call inconsistency and non-homogeneity. Inconsistency is a change in the amount of systematic error associated with the recording of data. It can arise from the use of different instruments and methods of observation. Non-homogeneity is a change in the statistical properties of the time series. Its causes can be either natural or man-made. These include alterations to land use, relocation of the observation station, and implementation of flow diversions.

The tests for stability of variance and mean verify not only the stationarity of a time series, but also its consistency and homogeneity. In the basic data-screening procedure, these two tests are reinforced by a third one, for absence of trend. Because

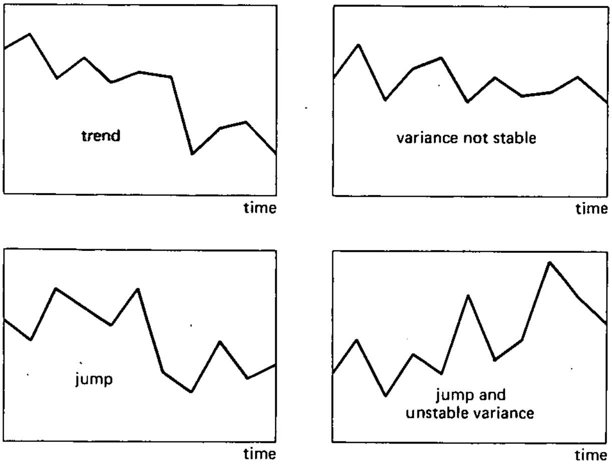

Figure 1.1 Four non-stationary time series

all three tests are performed on individual time series that are not compared with similar series, their results indicate the presence (or absence) of *absolute* consistency and homogeneity.

Although we have applied the basic data-screening procedure to time series of hydrological data that are summated over a year or a season, we assume that, if the data are acceptable at this level of aggregation, they will be equally acceptable at lower levels that cover, say, a month or a day. Nevertheless, the independence (and acceptability) of a time series depends on both the level of aggregation and the separation in time of the data points. Of these two, separation in time is the easier to verify.

For example, separation in time of the data points, so that successive hydrological events are not associated with related weather systems, is an obvious prerequisite to a successful frequency analysis, at whatever level of aggregation. A time series of annual rainfall totals or flow volumes is generally regarded as statistically independent. Groundwater carry-over and lake storage, however, can introduce persistence into a time series of flow volumes. Because of this, we have made it possible to extend the basic data-screening procedure to include a test for absence of persistence, based on the first serial-correlation coefficient.

A plot of progressive departures from the mean can help the engineer to pinpoint moments of change more accurately. Accordingly, we give an example of how to compute these departures and interpret the resulting plot.

After ascertaining the *absolute* consistency and homogeneity of the data series, one can use double-mass analysis to test its *relative* consistency and homogeneity. The basic data-screening procedure, when applied to a time series of proportionality factors, before and after a suspected break point in a double-mass line, is a good alternative to analysis of variance.

We give practical examples of how to use the data-screening procedure. We give complete equations to perform all the computations, but we discuss no statistical theory. Tables of relevant parts of Student's t-distribution and Fisher's F-distribution, for the customary level of significance of 5 per cent (two-tailed), are in Appendices 1 and 2. If another significance level is preferred, one should consult the tables in a statistical handbook (e.g. Spiegel 1961). Another possibility is to compute the significance level from the values of F and t, as Lackritz (1984) does, and as we have done in the computer program that accompanies this book.

We have not yet thoroughly investigated the power of the tests, i.e. their ability to reject a false test hypothesis. In many cases, their power will be weak. When the differences in test values (e.g. values from the t-test for stability of mean) are small, it will usually be of minor practical importance if a test fails to reject the test hypothesis (Hald 1952). But it frequently happens that the variance of a time series is large while the number of data is small. The variance of a hydrological time series will generally be greater than that of a controlled industrial process, and it will usually not be possible to continue sampling until there are enough data to increase the power of the test.

To minimize the problem, we recommend using no fewer than ten observations, or at least five in each sub-set. This is in line with the empirical rule in double-mass analysis, which states that one should disregard persistent changes that last less than five years. Our recommendation stems from our opinion of how the engineer should use the basic data-screening procedure, namely to identify time series that are obviously non-stationary. Even then, the engineer will still have to interpret the results of the tests, especially if they reject a test hypothesis. We emphasize here, therefore, that only a physical explanation of changes in variance and mean can justify the rejection of data that have probably been collected at great expense and under conditions that cannot be duplicated.

Rainfall records are extremely important. If they are consistent, they are independent of the works of man, thus providing an index for evaluating changes in, for example, stream flow. This is useful, as a change in runoff caused by a change in rainfall is not as troublesome as a change in runoff when there has been no change in rainfall (after Searcy and Hardison 1960). Accordingly, our examples deal mainly with time series of rainfall data, although the data-screening procedure can be applied equally well to time series of other data.

Most engineers prefer long time series of hydrological data. The longer the time series, however, the greater the chance that it is neither stationary, consistent, nor homogeneous. The latter part of a long time series can present a better data set if it is reasonable to expect that similar conditions will prevail in future.

We do not advocate using the common techniques of moving averages and double-mass analysis to screen data. Moving averages can introduce cycles into a time series that are difficult to analyze (World Meteorological Organization/WMO 1966). If a time series with fewer fluctuations is preferable, one can plot three-year or five-year averages in addition to the original time series. Double-mass analysis assumes proportionality between two variables. As it can verify only *relative* consistency and homogeneity, it cannot verify stationarity. Moreover, it requires more than one data set for the comparison, a luxury that is not always available.

2 The Data-Screening Procedure

The data screening procedure consists of four principal steps. These are:
- Do a rough screening of the data and compute or verify the totals for the hydrological year or season (Section 3.1);
- Plot these totals according to the chosen time step (e.g. month, year, season) and note any trends or discontinuities (Section 3.2);
- Test the time series for absence of trend with Spearman's rank-correlation method (Section 3.3);
- Apply the F-test for stability of variance and the t-test for stability of mean to split, non-overlapping, sub-sets of the time series (Section 3.4).

These steps form what we call the 'basic procedure'. If necessary, one can expand the basic procedure to include two additional steps. These are:
- Test the time series for absence of persistence by computing the first serial-correlation coefficient (Section 4);
- Test the time series for relative consistency and homogeneity with double-mass analysis (Section 6).

Together, the two sets of steps form the complete data-screening procedure, which is illustrated by the flowchart in Figure 2.1.

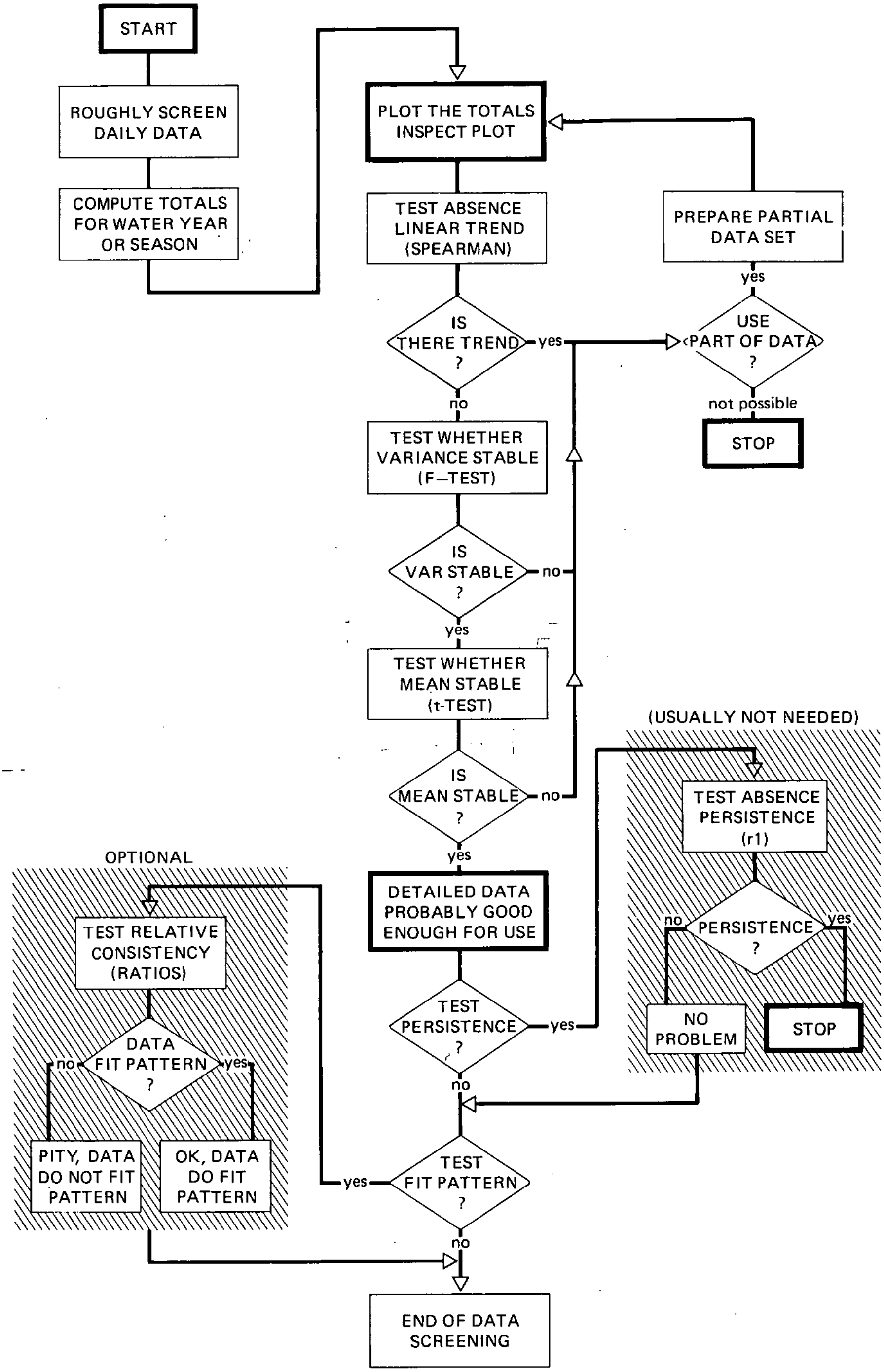

Figure 2.1 The data-screening procedure

3 The Basic Procedure

3.1 Rough Screening of the Data

The basic procedure begins with an initial, rough screening of the data. For rainfall totals, we advise tabulating daily observations by region (but observations from several collection stations should be available!). This will allow visual detection of whether the observations have been consistently or accidentally credited to the wrong day, whether they show gross errors (e.g. from weekly readings instead of daily ones), or whether they contain misplaced decimal points (Stol 1965). An analysis of the frequency distribution of one-day rainfall might also be useful. Other observations (e.g. of water levels) have their specific sources of error. One should be aware of these and the methods of detecting them.

Verifying the completeness of the data and checking the observer's arithmetic when computing totals is a useful exercise. One should particularly keep in mind the very real difference between 'no observation' and 'observation = 0'; both may have been entered as '–' (dash). Estimates of missing observations should be clearly marked as such.

In most cases, it is convenient – and perfectly acceptable – to use yearly totals as long as by 'year' one means 'water year' (hydrological year). This definition removes any risk of the seasons' being split over two years. Nevertheless, it can sometimes be better to analyze a specific period of a year (e.g. the wet or dry season, or even a particular month) if that period is a critical one in the envisaged water development scheme.

3.2 Plotting the Data

After doing a rough screening of the data, one plots them on arithmetic or semi-logarithmic graph paper. Figure 3.1 shows a time series of the yearly rainfall totals at the Bangkok Meteorological Department in Thailand from 1952 to 1985. Note that it does not show any obvious trends or discontinuities.

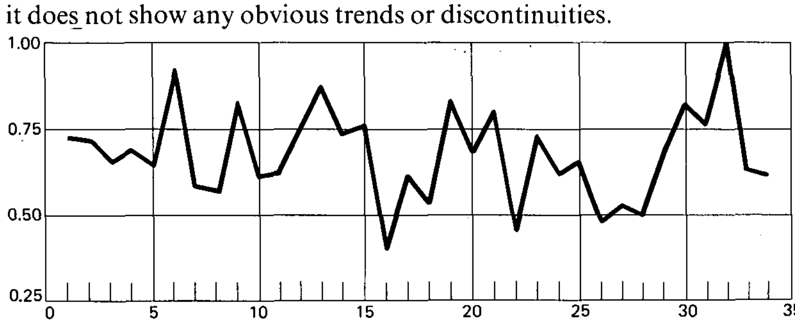

Figure 3.1 Time series of the yearly rainfall totals (in mm) at the Bangkok Meteorological Department from 1952 to 1985 (water years). The maximum observation is plotted at 1.0; other observations are scaled in relation to the maximum

3.3 Test for Absence of Trend

3.3.1 Spearman's Rank-Correlation Method

After plotting a time series, one must be sure that there is no correlation between the order in which the data have been collected and the increase (or decrease) in magnitude of those data. It is common practice to test the whole time series for absence of trend. Although one can choose to test only specific periods of the time series if these show signs of a possible trend, we advise against testing periods that are too short (ten to fifteen years). To verify absence of trend, we recommend using Spearman's rank-correlation method. It is simple and distribution-free, i.e. it does not require the assumption of an underlying statistical distribution. Yet another advantage is its nearly uniform power for linear and non-linear trends (WMO 1966). The method is based on the Spearman rank-correlation coefficient, R_{sp}, which is defined as:

$$R_{sp} = 1 - \frac{6 * \sum_{i=1}^{n} (D_i * D_i)}{n*(n*n-1)} \tag{3.1}$$

where n is the total number of data, D is difference, and i is the chronological order number. The difference between rankings is computed with:

$$D_i = Kx_i - Ky_i \tag{3.2}$$

where Kx_i is the rank of the variable, x, which is the chronological order number of the observations. The series of observations, y_i, is transformed to its rank equivalent, Ky_i, by assigning the chronological order number of an observation in the original series to the corresponding order number in the ranked series, y. If there are ties, i.e. two or more ranked observations, y, with the same value, the convention is to take Kx as the average rank. One can test the null hypothesis, $H_0{:}R_{sp} = 0$ (there is no trend), against the alternate hypothesis, $H_1{:}R_{sp} <> 0$ (there is a trend), with the test statistic:

$$t_t = R_{sp}\left[\frac{n-2}{1-R_{sp}*R_{sp}}\right]^{0.5} \tag{3.3}$$

where t_t has Student's t-distribution with $\nu = n-2$ degrees of freedom. Student's t-distribution is symmetrical around $t = 0$. Appendix 1 contains a table of the percentile points of the t-distribution for a significance level of 5 per cent (two-tailed).

(Incomplete tables, i.e. those listing only positive t-values and upper significance levels, are the rule in most statistical textbooks. One should therefore keep in mind that $t\{\nu,p\} = -t\{\nu,1-p\}$ when using such tables.) At a significance level of 5 per cent (two-tailed), the two-sided critical region, U, of t_t is bounded by:

$$\{-\infty, t\{\nu,2.5\%\}\} \cup \{t\{\nu,97.5\%\}, +\infty\}$$

and the null hypothesis is accepted if t_t is not contained in the critical region. In other words, the time series has no trend if:

$$t\{\nu,2.5\%\} < t_t < t\{\nu,97.5\%\} \tag{3.4}$$

If the time series does have a trend, the data cannot be used for frequency analyses or modelling. Removal of the trend is justified only if the physical processes underlying it are fully understood, which is rarely the case.

3.3.2 Application to Rainfall Data

Let us apply Spearman's rank-correlation method to the time series of rainfall data from Bangkok. We have introduced a tie (*) in the time series by increasing the amount of rainfall recorded for 1980 by 1 mm (Table 3.1). We entered the values in column Ky_i after locating the value of the ranked rainfall, y_i, in the column 'Rainfall' and copying the corresponding value of x (= index i). Because of the introduced tie, we averaged Kx_{19} and Kx_{20} to 19.5.

Table 3.1 Trend analysis of the yearly rainfall totals (in mm) at the Bangkok Meteorological Department from 1952 to 1985 (water years), with an introduced tie (*)

i = x	Rain-fall	y = Ranked Rainfall	Kx_i	Ky_i	D_i	D_i^2
1	1566	911	1.0	16.0	–15.0	225.00
2	1561	1011	2.0	22.0	–20.0	400.00
3	1414	1054	3.0	26.0	–23.0	529.00
4	1496	1094	4.0	28.0	–24.0	576.00
5	1411	1152	5.0	27.0	–22.0	484.00
6	1984	1195	6.0	18.0	–12.0	144.00
7	1279	1251	7.0	8.0	–1.0	1.00
8	1251	1279	8.0	7.0	1.0	1.00
9	1799	1330	9.0	10.0	–1.0	1.00
10	1330	1339	10.0	34.0	–24.0	576.00
11	1364	1344	11.0	17.0	–6.0	36.00
12	1617	1360	12.0	24.0	–12.0	144.00
13	1868	1364	13.0	11.0	2.0	4.00
14	1597	1392	14.0	33.0	–19.0	361.00
15	1642	1411	15.0	5.0	10.0	100.00
16	911	1414	16.0	3.0	13.0	169.00
17	1344	1428	17.0	25.0	–8.0	64.00
18	1195	1484	18.0	20.0	–2.0	4.00
19	1781	1496	19.5	4.0	15.5	240.25
20	1484	1496	19.5	29.0	–9.5	90.25
21	1716	1561	21.0	2.0	19.0	361.00
22	1011	1566	22.0	1.0	21.0	441.00
23	1579	1579	23.0	23.0	0.0	0.00
24	1360	1597	24.0	14.0	10.0	100.00
25	1428	1617	25.0	12.0	13.0	169.00
26	1054	1642	26.0	15.0	11.0	121.00
27	1152	1664	27.0	31.0	–4.0	16.00
28	1094	1716	28.0	21.0	7.0	49.00
29	*1496	1768	29.0	30.0	–1.0	1.00
30	1768	1781	30.0	19.0	11.0	121.00
31	1664	1799	31.0	9.0	22.0	484.00
32	2142	1868	32.0	13.0	19.0	361.00
33	1392	1984	33.0	6.0	27.0	729.00
34	1339	2142	34.0	32.0	2.0	4.00
Number of observations: 34						+ 7106.50
					R_{sp} =	–0.085791
					t_t =	–0.487

The table of percentile points for the t-distribution (Appendix 1) gives the critical values of t_t at the 5-per-cent level of significance for 34 – 2 = 32 degrees of freedom as:

$$t\{32,2.5\%\} = -2.02, \text{ and } t\{32,97.5\%\} = 2.02$$

Checking this result against the condition expressed in Equation 3.4:

$$-2.02 < ?\ -0.487\ ? < 2.02$$

one finds that the condition is satisfied. Thus, there is no trend. It is easy to verify that the original time series (without the introduced tie) had no trend, either, as $\Sigma D_i^2 = 7132.00$, $R_{sp} = -0.089687$, and $t_t = -0.509$.

3.3.3 Application to a Non-Stationary Time Series

Let us now apply the Spearman rank-correlation method to a non-stationary time series. Figure 3.2 shows a time series of the yearly rainfall totals at a problem station over twenty-two water years. A negative trend is clearly visible.

The values of R_{sp} and t_t are given in Table 3.2. There are no ties. The table of percentile points for the t-distribution (Appendix 1) shows that the critical values of t_t at the 5-per-cent level of significance for 22 – 2 = 20 degrees of freedom are:

$$t\{20,2.5\%\} = -2.09, \text{ and } t\{20,97.5\%\} = 2.09$$

Checking the computed t_t against the condition expressed in Equation 3.4:

$$-2.09 < ?\ -4.594\ ? < 2.09$$

one sees that the condition is not satisfied. Thus, there is a trend, and the time series is not stationary. If necessary, one can locate the exact break point in the time series by plotting the cumulative departures from the mean (Section 5) or using double-mass analysis (Section 6). Screening of the earlier data will then reveal whether they are suitable for further use.

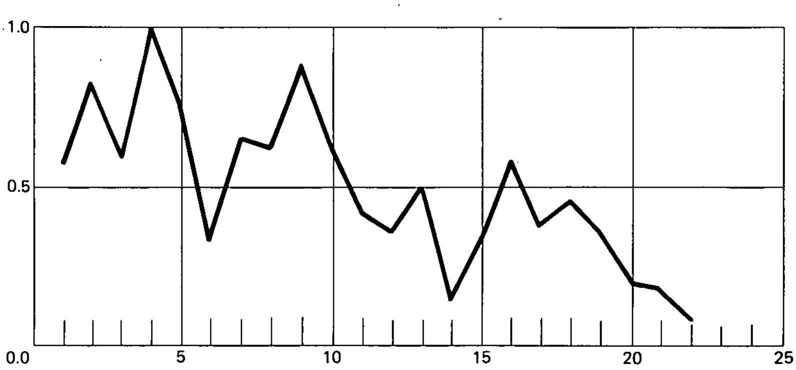

Figure 3.2 Time series of the yearly rainfall totals (in mm) at a problem station over twenty-two water years. The maximum observation is plotted at 1.0; other observations are scaled in relation to the maximum

Table 3.2 Trend analysis of the yearly rainfall totals (in mm) at a problem station over twenty-two water years

i = x	Rain-fall	y = Ranked Rainfall	Kx_i	Ky_i	D_i	D^2_i
1	1519	217	1.0	22.0	−21.0	441.00
2	2165	410	2.0	14.0	−12.0	144.00
3	1578	482	3.0	21.0	−18.0	324.00
4	2603	544	4.0	20.0	−16.0	256.00
5	1983	893	5.0	6.0	−1.0	1.00
6	893	907	6.0	15.0	−9.0	81.00
7	1703	944	7.0	19.0	−12.0	144.00
8	1656	955	8.0	12.0	−4.0	16.00
9	2307	1004	9.0	17.0	−8.0	64.00
10	1623	1094	10.0	11.0	−1.0	1.00
11	1094	1192	11.0	18.0	−7.0	49.00
12	955	1296	12.0	13.0	−1.0	1.00
13	1296	1519	13.0	1.0	12.0	144.00
14	410	1532	14.0	16.0	−2.0	4.00
15	907	1578	15.0	3.0	12.0	144.00
16	1532	1623	16.0	10.0	6.0	36.00
17	1004	1656	17.0	8.0	9.0	81.00
18	1192	1703	18.0	7.0	11.0	121.00
19	944	1983	19.0	5.0	14.0	196.00
20	544	2165	20.0	2.0	18.0	324.00
21	482	2307	21.0	9.0	12.0	144.00
22	217	2603	22.0	4.0	18.0	324.00

Number of observations: 22   + 3040.00

$R_{sp} = -0.716544$

$t_t = -4.594$

(The F-test for stability of variance and the t-test for stability of mean (Section 3.4) confirm the negative trend in the time series. The variances of Sub-Sets 1 to 11 and 12 to 20 of the time series are statistically similar: $F_t = 1.538$, $\nu_1 = 10$, and $\nu_2 = 10$, where F has the Fisher-distribution. Their means, however, are different at the 5-per-cent level of significance: $t_t = 4.492$ and $\nu = 20$. In addition, computation of the serial-correlation coefficient (Section 4) reveals persistence in the yearly rainfall totals, a highly unlikely phenomenon.)

The causes of the trend were, in fact, an ever-widening hole in the rain gauge, which the observer apparently did not note. For this reason, we do not give the location of the station or the years of observation.

3.4 Tests for Stability of Variance and Mean

3.4.1 The F-Test for Stability of Variance

In addition to testing the time series for absence of trend, one must test it for stability of variance and mean. The test for stability of variance is done first. There are two reasons for this sequence: firstly, instability of the variance implies that the time series is not stationary and, thus, not suitable for further use; secondly, the test for stability of mean is much simpler if one can use a pooled estimate of the variances of the two sub-sets. (This is permissible, however, only if the variances of the two sub-sets are statistically similar.)

The test statistic is the ratio of the variances of two split, non-overlapping, sub-sets of the time series. The distribution of the variance-ratio of samples from a normal distribution is known as the F, or Fisher, distribution. Even if the samples are not from a normal distribution, the F-test will give an acceptable indication of stability of variance.

Thus, the test statistic reads:

$$F_t = \frac{\sigma_1^2}{\sigma_2^2} = \frac{s_1^2}{s_2^2} \tag{3.5}$$

where s^2 is variance. Note that, to compute F_t, it is irrelevant whether one uses the sample standard deviation, s, or the population standard deviation, σ.

We give here two convenient formulae for computing the sample standard deviation, s, namely:

$$s = \left[\frac{\sum_{i=1}^{n}(x_i^2) - (\sum_{i=1}^{n}(x_i))^2/n}{n-1}\right]^{0.5}$$

and

$$s = \left[\frac{\sum_{i=1}^{n}(x_i^2) - n*\overline{x}^2}{n-1}\right]^{0.5}$$

where x_i is the observation, n is the total number of data in the sample, and $\overline{x}$ is the mean of the data.

The null hypothesis for the test, $H_0{:}s_1^2 = s_2^2$, is the equality of the variances; the alternate hypothesis is $H_1{:}s_1^2 <> s_2^2$. The rejection region, U, is bounded by:

$$\{0,F\{v_1,v_2,2.5\%\}\} U \{F\{v_1,v_2,97.5\%\},+\infty\} \tag{3.6}$$

where $v_1 = n_1-1$ is the number of degrees of freedom for the numerator, $v_2 = n_2-1$ is the number of degrees of freedom for the denominator, and n_1 and n_2 are the number of data in each sub-set. In other words, the variance of the time series is stable, and one can use the sample standard deviation, s, as an estimate of the population standard deviation, σ, if:

$$F\{v_1,v_2,2.5\%\} < F_t < F\{v_1,v_2,97.5\%\}$$

The F-distribution is not symmetrical for v_1 and v_2. One should therefore enter tables properly, usually by taking v_1 horizontally and v_2 vertically. (See Appendix 2 for a table of the F-distribution $F\{v_1,v_2,p\}$ for the 5-per-cent level of significance (two-tailed).)

(Many tables of the F-distribution in statistical textbooks are incomplete. They present only values of F that are greater than 1, i.e. only the values of higher probability. If the computed test statistic F_t is less than 1, it is still possible to use those tables by changing F_t to $1/F_t$. If one does this, however, one will also have to interchange the values of v_1 and v_2. The F-test thus appears to some as a one-tailed test because only the upper part of the distribution is used. It is, however, not correct to enter such tables at the 95.0-percentile row if the test is performed at the 5-per-cent level of significance. The 97.5-percentile row must be used for the two-tailed test, even when only the upper part of the table is available. If such is the case, the variance of the time series is stable if the value of the test statistic F_t complies with two conditions. These are:

$$F_t > 1 \tag{3.7a}$$

and

$$F_t < F\{v_1,v_2,97.5\%\} \tag{3.7b}$$

This method is tricky, and we do not use it here.)

One now divides the time series into two or three equal, or approximately equal, non-overlapping sub-sets and computes the variance of each with the square of the sample standard deviations, s. If the time series or the plot of cumulative departures from the mean contain a suspect period, one can delineate a sub-set to span that period and then compare it with one or more non-suspect periods.

3.4.2 The t-Test for Stability of Mean

The t-test for stability of mean involves computing and then comparing the means of two or three non-overlapping sub-sets of the time series (the same subsets from the F-test for stability of variance). A suitable statistic for testing the null hypothesis, $H_0 : \bar{x}_1 = \bar{x}_2$, against the alternate hypothesis, $H_1 : \bar{x}_1 <> \bar{x}_2$, is:

$$t_t = \frac{\bar{x}_1 - \bar{x}_2}{\left[\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2} * \left(\frac{1}{n_1}+\frac{1}{n_2}\right)\right]^{0.5}} \tag{3.8}$$

where n is the number of data in the sub-set, $\bar{x}$ the mean of the sub-set, and s^2 its variance. The test statistic t_t is valid for small samples with unknown variances. These variances can, however, differ only because of sampling variability if the t-test is applied in this form. This means that the variances of the sub-sets should not differ statistically: hence the requirement that the time series must be tested for stability of variance before it is tested for stability of mean. In samples from a normal distribution, t_t has a Student t-distribution. The requirement for normality is much less stringent for the t-test than for the F-test. One can apply the t-test to data that belong to any frequency distribution, but the length of the sub-sets should be equal if the distribution is skewed. One can avoid problems from a possibly skewed, underlying distribution by making the lengths of the sub-sets equal, or approximately so. For t_t, the two-sided critical region, U, is:

$$\{-\infty,t\{v,2.5\%\}\} \cup \{t\{v,97.5\%\},+\infty\}$$

with $v = n_1-1 + n_2-1$ degrees of freedom, i.e. the total number of data minus 2. If t_t is not in the critical region, the null hypothesis, $H_0:\bar{x}_1 = \bar{x}_2$, is accepted instead of the alternate hypothesis, $H_1:\bar{x}_1 <> \bar{x}_2$. In other words, the mean of the time series is considered stable if:

$$t\{\nu,2.5\%\} < t_t < t\{\nu,97.5\} \qquad (3.9)$$

3.4.3 Application to Rainfall Data

Let us apply the F-test for stability of variance and the t-test for stability of mean to the time series of rainfall data from Bangkok. We have computed the values of F_t and t_t for two sub-sets (Table 3.3) and three sub-sets (Table 3.4). Table 3.5 presents these values and their critical regions for various combinations of the sub-sets. The critical values come from the tables in Appendices 1 and 2.

The values of F_t fall outside the critical region in every case, so the pooled estimates of the variances can be used to do the t-test for stability of mean according to Equations 3.8 and 3.9. As the values of t_t also fall outside the critical region in every case, the variance and mean of the time series are stable at the 5-per-cent level of significance.

Table 3.3 Computation of F_t and t_t for two sub-sets of the yearly rainfall totals (in mm) at the Bangkok Meteorological Department from 1952 to 1985 (water years)

	Sub-Set I (Water Years 1-17)		Sub-Set II (Water Years 18-34)	
i	x_i	x_i^2	x_i	x_i^2
1	1566	2452356	1195	1428025
2	1561	2436721	1781	3171961
3	1414	1999396	1484	2202256
4	1496	2238016	1716	2944656
5	1411	1990921	1011	1022121
6	1984	3936256	1579	2493241
7	1279	1635841	1360	1849600
8	1251	1565001	1428	2039184
9	1799	3236401	1054	1110916
10	1330	1768900	1152	1327104
11	1364	1860496	1094	1196836
12	1617	2614689	1495	2235025
13	1868	3489424	1768	3125824
14	1597	2550409	1664	2768896
15	1642	2696164	2142	4588164
16	911	829921	1392	1937664
17	1344	1806336	1339	1792921
	+	+	+	+
Total	25434	39107248	24654	37234394

Number of observations:		17		17
$\bar{x}$:		1496.12		1450.24
s:		256.78		304.17
s^2:		65936.99		92518.32
F_t:	0.713		ν_1:	16
			ν_2:	16
t_t:	0.475		ν:	32

Table 3.4 Computation of F_t and t_t for three sub-sets of the yearly rainfall totals (in mm) at the Bangkok Meteorological Department from 1952 to 1985 (water years)

	Sub-Set I (Water Years 1-11)		Sub-Set II (Water Years 12-22)		Sub-Set III (Water Years 23-34)	
i	x_i	x_i^2	x_i	x_i^2	x_i	x_i^2
1	1566	2452356	1617	2614689	1579	2493241
2	1561	2436721	1868	3489424	1360	1849600
3	1414	1999396	1597	2550409	1428	2039184
4	1496	2238016	1642	2696164	1054	1110916
5	1411	1990921	911	829921	1152	1327104
6	1984	3936256	1344	1806336	1094	1196836
7	1279	1635841	1195	1428025	1495	2235025
8	1251	1565001	1781	3171961	1768	3125824
9	1799	3236401	1484	2202256	1664	2768896
10	1330	1768900	1716	2944656	2142	4588164
11	1364	1860496	1011	1022121	1392	1937664
12					1339	1792921
	+	+	+	+	+	+
Total	16455	25120305	16166	24755962	17467	26465375
Number of observations:	11		11		12	
$\bar{x}$:	1495.91		1469.64		1455.58	
s:	224.75		315.88		307.59	
s^2:	50512.09		99782.05		94609.17	
Sub-Sets:	1-11/12-22:		1-11/23-34:		12-22/23-34:	
ν_1:	10		10		10	
ν_2:	10		11		11	
F_t:	0.506		0.534		1.055	
ν:	20		21		21	
t_t:	0.225		0.356		0.108	

Table 3.5 Results of the computations of F_t and t_t for various combinations of sub-sets of the yearly rainfall totals (in mm) at the Bangkok Meteorological Department from 1952 to 1985 (water years)

Sub-Set (Water Years)	Sub-Set (Water Years)	ν_1,ν_2	$F_{2.5\%}$ / F_t / $F_{97.5\%}$	ν	$t_{2.5\%}$ / t_t / $t_{97.5\%}$
			0.362		-2.02
1-17	18-34	16,16	0.713	32	0.475
			2.76		2.02
			0.269		-2.09
1-11	12-22	10,10	0.506	20	0.225
			3.72		2.09
			0.273		-2.06
1-11	23-34	10,11	0.534	21	0.356
			3.53		2.06
			0.273		-2.06
12-22	23-34	10,11	1.055	21	0.108
			3.53		2.06

To summarize, then, a rough screening of the Bangkok rainfall data (not described here) and plotting the data as a time series revealed no major discrepancies. There was no trend, and the variance and mean were stable. Therefore the time series is stationary in the sense used for this data screening, and there is no immediate objection to using the data, even at lower levels of aggregation, i.e. those covering a day, a week, ten days, a month, and so on.

3.4.4 Application to Water Level Data

We shall now apply the F-test for stability of variance and the t-test for stability of mean to a time series of water level data. Figure 3.3 shows a time series of the maximum yearly levels of the Chao Phraya River from 1955 to 1974. The data were collected at Bang Sai, Thailand, in the lower catchment. In the middle of the observation period, a major storage reservoir and power station were built in the upper catchment. Although the reservoir controls only some of the floods from the upper catchment, there are indications that it might have affected the water levels downstream. Rough screening of the data revealed no obvious errors; the time series of the data has no trend ($t_t = -1.75, \nu = 18$).

Let us divide the time series into two sub-sets of ten years each (the periods before and after completion of the dam). The values of F_t and t_t for the sub-sets are given in Tables 3.6 and 3.7. They show that while the variances are stable, the means are not, as t_t is in the critical region. Therefore, the difference between the means of the sub-sets (2.84 m for Sub-Set I and 2.15 m for Sub-Set II) is real at the 5-per-cent level of significance. The time series shows a negative jump after completion of the dam: hence it is not stationary.

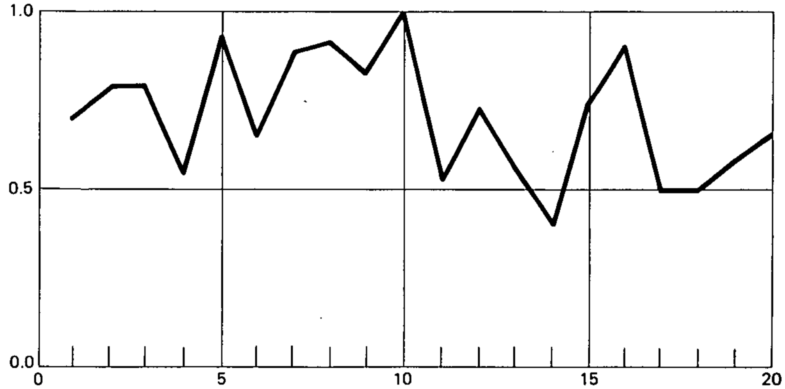

Figure 3.3 Time series of maximum yearly levels (in m) of the Chao Phraya River at Bang Sai, Thailand, from 1955 to 1974 (water years). The maximum observation is plotted at 1.0; other observations are scaled in relation to the maximum

Table 3.6 Computation of F_t and t_t for two sub-sets of maximum yearly water levels (in m) of the Chao Phraya River at Bang Sai, Thailand, from 1955 to 1974 (water years)

	Sub-Set I (Water Years 1-10)		Sub-Set II (Water Years 11-20	
i	x_i	x_i^2	x_i	x_i^2
1	2.49	6.2001	1.88	3.5344
2	2.80	7.8400	2.54	6.4516
3	2.78	7.7284	1.98	3.9204
4	1.95	3.8025	1.42	2.0164
5	3.29	10.8241	2.63	6.9169
6	2.30	5.2900	3.16	9.9856
7	3.14	9.8596	1.78	3.1684
8	3.20	10.2400	1.76	3.0976
9	2.92	8.5264	2.04	4.1616
10	3.51	12.3201	2.31	5.3361
	+	+	+	+
Total	28.38	82.6312	21.50	48.5890
Number of observations:	10		10	
$\overline{x}$:	2.8380		2.1500	
s:	0.4818		0.5125	
s^2:	0.2321		0.2627	
F_t:	0.884	ν_1:	9	
		ν_2:	9	
t_t:	3.093	ν:	18	

Table 3.7 Results of the computation of F_t and t_t for two sub-sets of maximum yearly water levels of the Chao Phraya River at Bang Sai, Thailand, from 1955 to 1974

Sub-Set (Water Years)	Sub-Set (Water Years)	ν_1,ν_2	$F_{2.5\%}$ F_t $F_{97.5\%}$	ν	$t_{2.5\%}$ t_t $t_{97.5\%}$
1-10	11-20	9,9	0.248 0.884 4.03	18	–2.10 3.90 2.10

4 Test for Absence of Persistence

4.1 The Serial-Correlation Coefficient

We stated in Section 1 that time series of yearly and seasonal totals are usually independent. Notable exceptions are time series of data from rivers with a considerable carry-over of groundwater flow from one year to the next and those of data from rivers whose catchments include large lakes. In these cases, one will want to test the time series for independence.

The serial-correlation coefficient can help to verify the independence of a time series. If a time series is completely random, the population auto-correlation function will be zero for all lags other than zero (when its value is unity, because all data sets are perfectly correlated with themselves), and the sample serial-correlation coefficients will deviate slightly from zero only because of sampling effects. For our purposes, it is usually sufficient to compute the lag 1 serial-correlation coefficient, i.e. the correlation between adjacent observations in a time series. Here, we define the lag 1 serial-correlation coefficient, r_1, according to Box and Jenkins (1970). This reads:

$$r_1 = \frac{\sum_{i=1}^{n-1} (x_i - \overline{x}) * (x_{i+1} - \overline{x})}{\sum_{i=1}^{n} (x_i - \overline{x})} \qquad (4.1)$$

where x_i is an observation, x_{i+1} is the following observation, $\overline{x}$ is the mean of the time series, and n is the number of data.

After computing r_1, one can test the hypothesis $H_0{:}r_1 = 0$ (that there is no correlation between two consecutive observations) against the alternate hypothesis, $H_1{:}r_1 <> 0$. Anderson (1942) defines the critical region, U, at the 5-per-cent level of significance as:

$$\{-1,(-1-1.96(n-2)^{0.5})/(n-1)\} \text{ U } \{(-1+1.96(n-2)^{0.5})/(n-1),+1\} \qquad (4.2)$$

4.2 Application to Rainfall Data

Let us now apply the test for absence of persistence to the time series of rainfall totals from Bangkok. Normally, rainfall data do not have to be checked for persistence, but we prefer to use here the same data that we used for the other examples. Table 4.1 gives the value of the lag 1 serial-correlation coefficient, r_1, as:

$$r_1 = 99811.67/2553178.94 = 0.0391$$

Equation 4.2 gives the upper confidence limit, UCL, for r_1 as:

$$UCL(r_1) = (-1+1.96(34-2)^{0.5})/(34-1) = 0.306$$

and the lower confidence limit, LCL, as:

$$LCL(r_1) = (-1-1.96(34-2)^{0.5})/(34-1) = -0.366$$

To accept the hypothesis $H_0{:}r_1 = 0$, the value of r_1 should fall between the UCL and the LCL.

Applying this condition to the time series, we see that the condition:

$$-0.366 <?\ 0.0391\ ?< 0.306$$

is satisfied. Thus, no correlation exists between successive observations. The data are independent, and there is no persistence in the time series.

Table 4.1 Computation of the lag 1 serial-correlation coefficient for the yearly rainfall totals (in mm) at the Bangkok Meteorological Department from 1952 to 1985 (water years)

(1) i	(2) x_i	(3) x_i-x_{av}	(4) $(3)_i$*$(3)_{i+1}$	(5) $(3)_i$*$(3)_i$
1952	1566	92.82	8152.09	8616.21
1953	1561	87.82	−5197.09	7712.97
1954	1414	−59.18	−1350.62	3501.85
1955	1496	22.82	−1419.09	520.91
1956	1411	−62.18	−31761.20	3865.91
1957	1984	510.82	−99189.91	260940.68
1958	1279	−194.18	43141.44	37704.50
1959	1251	−222.18	−72390.32	49362.38
1960	1799	325.82	−46650.26	106160.97
1961	1330	−143.18	15631.50	20499.50
1962	1364	−109.18	−15702.15	11919.50
1963	1617	143.82	56784.91	20685.21
1964	1868	394.82	48888.44	155885.62
1965	1597	123.82	20904.33	15332.27
1966	1642	168.82	−94908.62	28501.38
1967	911	−562.18	72619.97	316042.38
1968	1344	−129.18	35933.85	16686.56
1969	1195	−278.18	−85629.26	77382.15
1970	1781	307.82	3331.74	94755.33
1971	1484	10.82	2628.21	117.15
1972	1716	242.82	−112227.32	58963.27
1973	1011	−462.18	−48909.15	213607.09
1974	1579	105.82	−11976.73	11198.62
1975	1360	−113.18	5112.91	12808.91
1976	1428	−45.18	18936.91	2040.91
1977	1054	−419.18	134629.62	175708.91
1978	1152	−321.18	121782.56	103154.33
1979	1094	−379.18	−8274.97	143774.80
1980	1495	21.82	6434.09	476.27
1981	1768	294.82	56259.27	86920.91
1982	1664	190.82	127627.27	36413.62
1983	2142	668.82	−54292.73	447324.91
1984	1392	−81.18	10891.97	6589.62
1985	1339	−134.18		18003.33
	+		+	+
	50088		99811.67	2553178.94

$\bar{x} = 1473.18$ $\qquad r_1 = 0.0391$

5 Cumulative Departures from the Mean

In the early days of data verification, engineers paid much attention to the mean and changes in the mean of a time series. Until 1937, plotting and analyzing the cumulative departures from the mean were the principle methods of verifying the consistency and homogeneity of hydrometeorological data. In that year, C.F. Merriam proposed adding to these methods what we now call double-mass analysis (Section 6.2), and using the plot of cumulative departures from the mean from a single observation station to do a kind of rough screening.

Although the basic data-screening procedure uses statistics to screen individual time series of hydrological data, plotting the cumulative departures from the mean can be very helpful in testing such a time series for stability of mean. In the following section, we shall show how a plot of cumulative departures from the mean can also be used to facilitate the test for relative consistency.

Computing the cumulative departures from the mean is very straightforward. The assumption is that all previous observations result in a zero cumulative departure from the mean. Starting, then, from zero, one obtains the first cumulative departure by subtracting the mean from the first observation of the time series. For the second cumulative departure, one subtracts the mean from the second observation of the time series and adds this value to the first departure. The computing continues in this way until the last departure, which, of course, should be zero again. Table 5.1 gives the cumulative departures from the mean of a stationary time series (x1) and one suspected of being non-stationary (x2).

Table 5.1 Cumulative departures from the mean of two time series

	Time Series I		Time Series II	
i	$x1_i$	$\Sigma(x1_i-\overline{x}1)$	$x2_i$	$\Sigma(x2_i-\overline{x}2)$
0		0.00		0.00
1	1566	70.09	1387	85.64
2	1561	135.18	1450	234.27
3	1414	53.27	1584	516.91
4	1496	53.36	1423	638.55
5	1411	−31.55	1775	1112.18
6	1984	456.55	1187	997.82
7	1279	239.64	1225	921.45
8	1251	−5.27	950	570.09
9	1799	297.82	1221	489.73
10	1330	131.91	987	175.36
11	1364	0.00	1126	0.00
	$\overline{x}1$ = 1495.91		$\overline{x}2$ = 1301.36	

Plots of the cumulative departures from the mean, $x_i - \bar{x}$, are illustrated in Figure 5.1. Time Series I shows no consistent trend. It was computed from the yearly rainfall totals at the Bangkok Meteorological Department from 1952 to 1962. As there are no real differences in the mean of Time Series I, the plot shows only trends of short duration, and the cumulative departures from the mean fluctuate randomly. Time Series II, however, shows a definite positive trend in the first five years and a negative trend in the next six years. This plot can be interpreted easily because maximum and minimum totals correspond with breaks in the data's statistical properties of the mean.

It is possible to ascertain the significance of a break only by testing the time series for stability of variance and mean. With the F-test for stability of variance and the t-test for stability of mean, it is easy to confirm that the variance of Sub-Sets 1 to

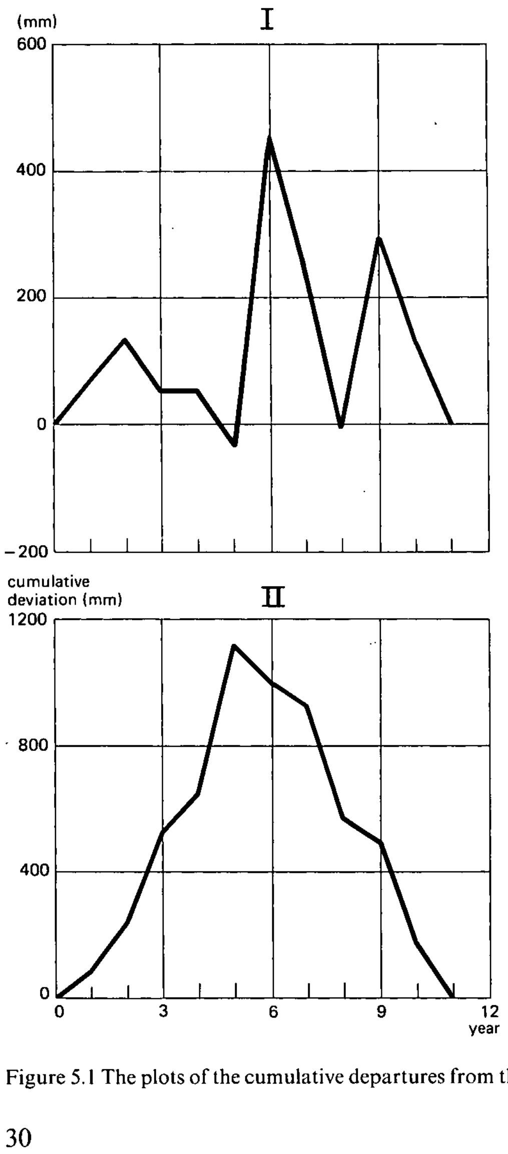

Figure 5.1 The plots of the cumulative departures from the mean of the two time series in Table 5.1

5 and 6 to 11 of Time Series II is stable (F_t = 1.748, with ν_1 = 4 and ν_2 = 5), but that the difference in the mean of these Sub-Sets is real (t_t = 4.853, with ν = 9). Time Series II is therefore not stationary, and the break in the mean after Year 5 is real.

Bernier (1977) and others have proposed various procedures for locating break points in plots of cumulative departures from the mean. Unfortunately, these require computations that are often very elaborate. For our purposes, applying the F-test for stability of variance and the t-test for stability of mean to several points around the expected break point will usually suffice (see also Section 6.3.2). This has the additional advantage of confirming the stability of the variance, which is an equally important prerequisite to the use of the data.

6 Tests for Relative Consistency and Homogeneity

6.1 Introduction

In this section, 'consistency' will mean 'consistency and homogeneity'. The tests we describe here (like those we described previously) cannot differentiate between inconsistency and non-homogeneity. This would require an analysis of the process that generated the data, including the observational practices (which are causes of inconsistency) and the physical changes (which are causes of non-homogeneity).

A time series of hydrological data is relatively consistent if the periodic data are proportional to an appropriate simultaneous time series (Chang and Lee 1974). In other words, *relative* consistency means that the hydrological data at a certain observation station are generated by the same mechanism that generated similar (e.g. rainfall/rainfall) or related (e.g rainfall/runoff) data at other stations. It is common practice to verify relative consistency with double-mass analysis.

The relative consistency of time series from different stations is often irrelevant, and the data in these series can very well be suitable for independent use if they are absolutely consistent and homogeneous. Nevertheless, it is still a good idea to verify whether the areas covered by certain stations are in the same hydrological region. If a time series is relatively consistent, it is suitable for correcting and filling in data, but we recommend reading the description of double-mass analysis in Section 6.2 before using it in this way.

Relative consistency is a true consistency only if there is physical evidence to support this. Examples of a possibly true and relatively consistent relation are the proportionality between rainfall and rainfall, computed runoff and observed runoff, and rainfall and sediment concentration in a river.

To determine relative consistency, one compares the observations from a certain station with the mean of observations from several nearby stations. This mean is called the 'base' or 'pattern'. It is difficult to say how many stations the pattern should comprise. The more stations, the smaller the chance that inconsistent data from a particular one will influence the validity of the average of the pattern. Ten is the accepted minimum number of stations, but there may not be that many in the area. If there are fewer than ten stations, the data from each one must be checked very carefully before being included in the pattern.

In conventional double-mass analysis, this checking requires removing from the pattern the data from a certain station and comparing them with the remaining data. If these data are consistent with the general totals in the area, they are re-incorporated into the pattern. Double-mass analysis cannot, however, detect similar changes that occurred at the stations simultaneously. For example, if, at the same time, all the stations in a region started to record data that were, say, 50 percent too great, the double-mass curve would not show a significant break. Consequently, double-mass analysis is not suitable for testing the stationarity of a time series. We recommend using the basic data-screening procedure for this purpose instead.

In some cases, the time series from certain stations are inconsistent with the pattern but consistent with each other. One should then group these stations in a pattern of their own and accept that there is a regional anomaly in the general pattern. Mapping the location of these stations will make it easier to decide how to group them.

6.2 Double-Mass Analysis

Double-mass analysis assumes a linear relation between time series of hydrological data. As this assumption may not be valid at all rates of accumulation, it must be verified. Rainfall data are usually proportional to totals at nearby stations in the same hydrological area.

The term 'double-mass curve' is commonly used in the literature. We shall use the term 'double-mass line' instead, to stress the assumed linear relation between the data sets. Non-linear relations fall outside the scope of this book.

A linear relation between two variables that include the pair $x = 0$ and $y = 0$ can be expressed as:

$$y = b * x \tag{6.1}$$

where b is a proportionality factor.

If y_i is the time series to be tested, x_i the time series of the pattern, and $i = 0, ..., n$ (the number of data pairs and the index of the time steps), then the plot of $Y_i = \Sigma(y_i)$ (the mass of y) against $X_i = \Sigma(x_i)$ (the mass of x) will result in a broken line through the origin, with an average slope $b_{av} = Y_n/X_n$. The line passes through the origin because the sum of the data at time zero is zero for both X and Y. Defining the average slope as the slope of the line through the points 0,0 and Y_n, X_n will give a good enough estimate of the true mean of the proportionality factors.

The plotted points will never fall exactly on the average line. If there is a trend away from the line during a certain period, then an opposite trend will necessarily materialize during a following period to realize the average slope for the whole period. Analyzing persistent trends away from the average slope, one sees that break points between two periods with apparently different slopes indicate the moment at which the linear relationship changes between the means of two parts of the time series. This is a break that, if significant, indicates a real change.

Double-mass analysis is used not only to verify the relative consistency of a time series, but also to find correction factors for errors and fill in gaps. This application is limited to monthly and yearly totals, as it normally does not work with daily ones. Furthermore, at its best, double-mass analysis preserves the mean and not the standard deviation of the time series, unless a proportional error has been made (e.g. measuring rainfall in a measuring jar that is not calibrated for the sampling area).

It is generally acknowledged that C.F. Merriam was the first to use double-mass analysis to test a time series for relative consistency. In a paper published in 1937, Merriam compares two tests for relative consistency, namely the plotting of cumulative departures from the mean and the cumulative plotting of one time series against another, i.e. double-mass analysis. He concludes that double-mass analysis is useful in screening time series of rainfall and runoff data if it indicates absence of change in proportionality, i.e. absence of change in the slope of the line. A major drawback was the initial lack of objective criteria to judge whether an apparent change in proportionality was a real change.

Weiss and Wilson (1953), in their paper on evaluating the significance of slope changes, stress that the probability of an abrupt change occurring purely by chance is important. They describe a statistical method to determine whether there is a significant difference between the mean slope of the periods before and after the break. Their method is not widely used, possibly because it requires a special protractor and nomograph.

Searcy and Hardison (1960) use analysis of variance in a statistical test of the significance of an apparent break in the slope. In their example, they use shortcuts to apply the F-test for stability of variance to the problem, thus obscuring the test (as they themselves admit). This may be the reason why so few engineers use the test, even if they are conversant with analysis-of-variance computations and their interpretation.

Hansel and Schäfer (1970) and Dyck (1980) present a straightforward analysis of variance for determining the significance of slope changes in double-mass lines. Unfortunately, these publications have generally gone unnoticed, possibly because they are not in English.

Hence the statements in the literature like this one by Bernier (1977), who writes that 'the great shortcoming [of double-mass analysis] is the lack of appropriate statistical techniques for the determination of the significance of apparent breaks'. Most hydrological handbooks (and, for that matter, lecture notes on hydrology) shift the problem to the reader, who finds remarks like: 'a break in the double-mass line is a real break and indicates a real change if the break is significant,' without any definition of what this significance could be.

In a test of the significance of changes in the proportionality factor, analysis of variance supposes a normal distribution of the data. The computations can be done easily with a programmable calculator (many of which have pre-programmed algorithms) or a computer with a statistical package. We have chosen not to discuss these procedures here. Instead, we shall show how the basic data-screening procedure is an elegant alternative, suitable for almost all applications (Section 6.3).

6.2.1 A Simple Example of Double-Mass Analysis

We shall use the data in Table 6.1 to give a simple example of double-mass analysis. The cumulative sums, X_i and Y_i, are plotted against each other and a line (or lines) of best fit are drawn freehand. A simplification is to plot Y_i against the year numbers, which is certainly permissible if successive values of X_i are not too different (top plot, Figure 6.1). The plot is difficult to interpret, even with scales more appropriate than the ones used here. Searcy and Hardison (1960) recommend plotting the cumulative differences between the sums at the test station against the corresponding values of the average line, i. e. plotting $Y_i - b_{av} * X_i$ against X_i, or against the time-step index (bottom plot, Figure 6.1).

Table 6.1 Sample data for double-mass analysis

(1) i	(2) x_i = Pattern	(3) y_i = Test	(4) $\Sigma(x_i) = X_i$	(5) $\Sigma(y_i) = Y_i$	(6) $Y_i + -b_{av}*X_i$	(7) $a_i = y_i/x_i$
0	0	0	0	0	0	
1	1362	1243	1362	1243	16	0.9126
2	1111	990	2473	2233	5	0.8911
3	1337	1310	3810	3543	111	0.9798
4	1392	1255	5202	4798	112	0.9016
5	1914	1784	7116	6582	172	0.9321
6	1252	1232	8368	7814	276	0.9840
7	1309	1189	9677	9003	286	0.9083
8	1283	1102	10960	10105	232	0.8589
9	1260	979	12220	11084	76	0.7770
10	1643	1421	13863	12505	18	0.8649
11	1415	1240	15278	13745	–17	0.8763
12	1450	1236	16728	14981	–87	0.8524
13	1141	1115	17869	16096	0	0.9772
			$b_{av} = 0.90078$			

In this plot of cumulative differences, also called a residual-mass plot, the maximum and minimum values correspond to break points in the original double-mass line, making interpretation easier (c.f. the plot in Section 5). In our example, an obvious – and possibly significant – break point is at Year 7. There are other methods of identifying possible break points in double-mass lines (e.g. Singh 1968, Chang and Lee 1974), but the one we describe here is simple and efficient.

The average ratio or slope of the data in Table 6.1, from Year 1 to Year 7, is:

$$b_1 = (9003-0)/(9677-0) = 0.9304$$

and from Year 8 to Year 13 it is:

$$b_2 = (16096-9003)/(17869-9677) = 0.8658$$

The question is now: are b_1 and b_2 statistically different? For a conventional answer, i.e. one arrived at by analysis of variance, we refer the reader to the literature (Weiss and Wilson 1953, Searcy and Hardison 1960, Hansel and Schäfer 1970, and Dyck 1980).

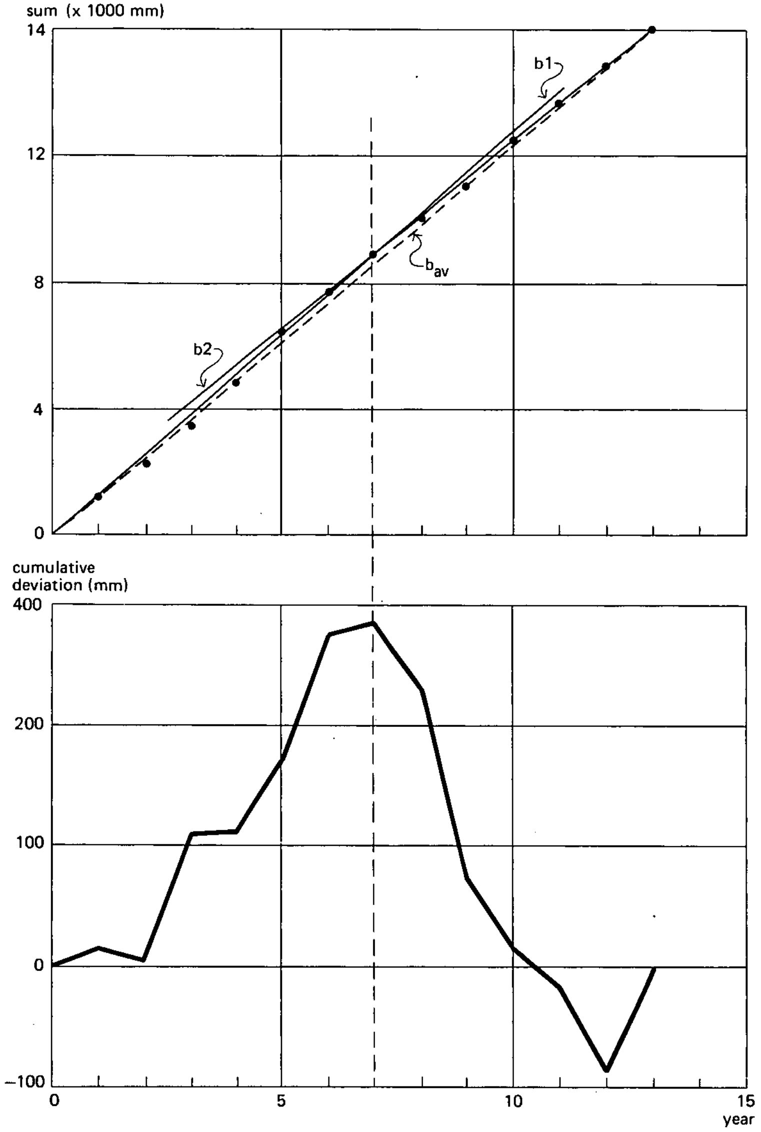

Figure 6.1 Two ways of plotting the data in Table 6.1. Top: Simplified double-mass plot (Y_i vs i). Bottom: Plot of cumulative differences (residual-mass plot, $Y_i - b_{av} * X_i$ vs i)

6.3 Analysis of Proportionality Factors

The basic data-screening procedure is a good alternative to double-mass analysis as a test for relative consistency. One applies it to the time series of proportionality factors from a test station and a pattern, $a_i = y_i/x_i$. Testing the proportionality factors (ratios) of two periods for stability of mean is equivalent to testing the significance of changes in slopes of periods before and after an apparent break point in a double-mass line. The additional advantage of the basic data-screening procedure is that the stability of variance is tested too, indicating any data corrections that might have influenced the variance.

A disadvantage is the omission of data if one adheres to the requirement that the number of data in the sub-sets should be equal, or nearly so (Section 3.4.2). Usually, to identify significant changes, even short series of data before and after the apparent break point are sufficient to arrive at a conclusion. In exceptional cases, one can verify normal distribution with, for example, d'Agostino's method (1971). If the data are normally distributed, then, of course, one can use sub-sets with an unequal number of data.

6.3.1 A Simple Example of Analysis of Proportionality Factors

We shall now give an example of the alternative test, using the data in Table 6.1. The proportionality factors (ratios) are in Column 7. The time series of the proportionality factors, a_i, is shown in Figure 6.2.

The plot shows no obvious negative trend. The test for absence of trend reveals no negative trend, either. Even so, previous computations and the slopes in the simplified double-mass plot in Figure 6.1 indicate that the second half of the time series has a smaller mean than the first.

The computations of F_t and t_t for the time series are shown in Table 6.2. The results of the computations are shown in Table 6.3. The critical values of F_t and t_t are from Appendices 1 and 2. Table 6.3 shows that the variance and mean of the sub-sets are statistically similar. This means that the maximum value of cumulative departures from the mean in Water Year 7 does not correspond to a real break in the line, and that the proportionality between the data from the test station and those from the pattern does not change thereafter.

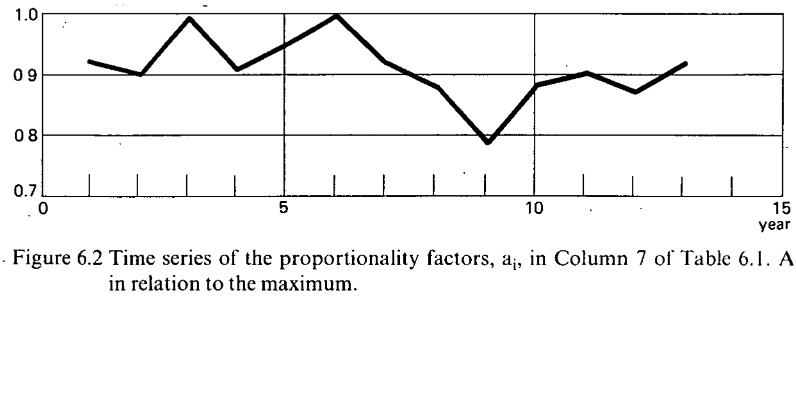

Figure 6.2 Time series of the proportionality factors, a_i, in Column 7 of Table 6.1. All values are scaled in relation to the maximum.

Table 6.2 Computation of F_t and t_t for two sub-sets of proportionality factors, a_i

	Sub-Set I (Water Years 1-17)		Sub-Set II (Water Years 8-13)	
i	x_i	x_i^2	x_i	x_i^2
1	0.9126	0.83284	0.8589	0.73775
2	0.8911	0.79406	0.7770	0.60370
3	0.9798	0.96001	0.8649	0.74802
4	0.9016	0.81288	0.8763	0.76795
5	0.9321	0.86881	0.8524	0.72661
6	0.9840	0.96826	0.9772	0.95495
7	0.9083	0.82501		
	6.5095 +	6.0619 +	5.2067 +	4.5390 +

Number of observations:	7		6
$\bar{x}$:	0.9299286		0.8677903
s:	0.0376250		0.0642113
s^2:	0.0014156		0.0041231
F_t:	0.343	v_1:	6
		v_2:	5
t_t:	2.171	v:	11

Table 6.3 Results of the computation of F_t and t_t for two sub-sets of proportionality factors, a_i

Sub-Set I (Water Years)	Sub-Set II (Water Years)	v_1,v_1	$F_{2.5\%}$ / F_t / $F_{2.5\%}$	v	$t_{2.5\%}$ / t_t / $t_{97.5\%}$
			0.169		−2.20
1-7	8-13	6,5	0.343	11	2.17
			6.98		2.20

If one disregards the observation for Water Year 13 (or does not have it yet) and tests for changes in the slope of Water Years 1 to 6 and 7 to 12, one will find a significant change in the proportionality factors: $F_t = 0.838$, with $v_1 = 5$ and $v_2 = 5$ is acceptable at the 5-per-cent level of significance; $t_t = 3.204$, with $v = 10$, indicating that the factors changed after Water Year 6. Trend analysis also indicates that there is a change ($t_t = -2.81$, with $v = 10$).

6.3.2 Application to Runoff Data

Now let us use the analysis of proportionality factors with the basic data-screening procedure to test the relative consistency of a time series of runoff data. We have taken the data for this example from Table 2 of a paper by Searcy and Hardison

(1960). The data cover the water years from 1921 to 1945. We shall compare the yearly runoff from Stream A with that from a pattern. The data and computations are presented in Table 6.4. To adhere to the cumulative sums of Searcy and Hardison, we have corrected the assumed printing error in the observation for Stream A in 1936 (Water Year 16). According to the double-mass analysis performed by Searcy and Hardison, there is a break point in 1938 (Water Year 18), which their analysis of variance confirms.

Table 6.4 Test of the relative consistency of runoff data* (in inches) collected from 1921 to 1945 (water years)

(1) Index No.	(2) x_i = Pattern	(3) y_i = Stream A	(4) $\Sigma(x_i)$ $= X_i$	(5) $\Sigma(y_i)$ $= Y_i$	(6) $Y_i + -b_{av}*X_i$	(7) $a_i = y_i/x_i$
0	0.00	0.00	0.00	0.00	0.00	
1	19.61	19.73	19.61	19.73	–2.86	1.0061
2	12.29	15.80	31.90	35.53	–1.22	1.2856
3	8.12	17.52	40.02	53.05	6.94	2.1576
4	14.39	16.58	54.41	69.63	6.94	1.1522
5	3.53	5.33	57.94	74.96	8.21	1.5099
6	13.80	16.45	71.74	91.41	8.76	1.1920
7	24.03	30.67	95.77	122.08	11.74	1.2763
8	12.40	21.22	108.17	143.30	18.68	1.7113
9	19.70	21.96	127.87	165.26	17.94	1.1147
10	18.10	19.34	145.97	184.60	16.43	1.0685
11	5.13	9.87	151.10	194.47	20.39	1.9240
12	18.30	24.81	169.40	219.28	24.12	1.3557
13	12.20	15.53	181.60	234.81	25.59	1.2730
14	7.94	9.35	189.54	244.16	25.79	1.1776
15	25.58	32.75	215.12	276.91	29.07	1.2803
16	4.06	7.75	219.18	284.66	32.14	1.9089
17	13.76	19.72	232.94	304.38	36.01	1.4331
18	28.64	28.33	261.58	332.71	31.35	0.9892
19	10.41	15.04	271.99	347.75	34.39	1.4448
20	10.68	13.65	282.67	361.40	35.74	1.2781
21	30.15	17.42	312.82	378.82	18.42	0.5778
22	21.60	17.82	334.42	396.64	11.36	0.8250
23	8.96	9.41	343.38	406.05	10.45	1.0502
24	20.01	21.13	363.39	427.18	8.52	1.0560
25	40.25	37.85	403.64	465.03	0.00	0.9404
			$b_{av} = 1.15209$			

* After Searcy and Hardison (1960).

Figure 6.3 illustrates two plots: one of the proportionality factors, a_i, of the runoff from Stream A and the pattern (Column 7 of Table 6.4), and one of the cumulative departures from the mean slope (Column 6 of Table 6.4). The maximum departure occurs in 1937 (Water Year 17), the year prior to the one that Searcy and Hardison identified as the break point. This discrepancy could have come from their graphical

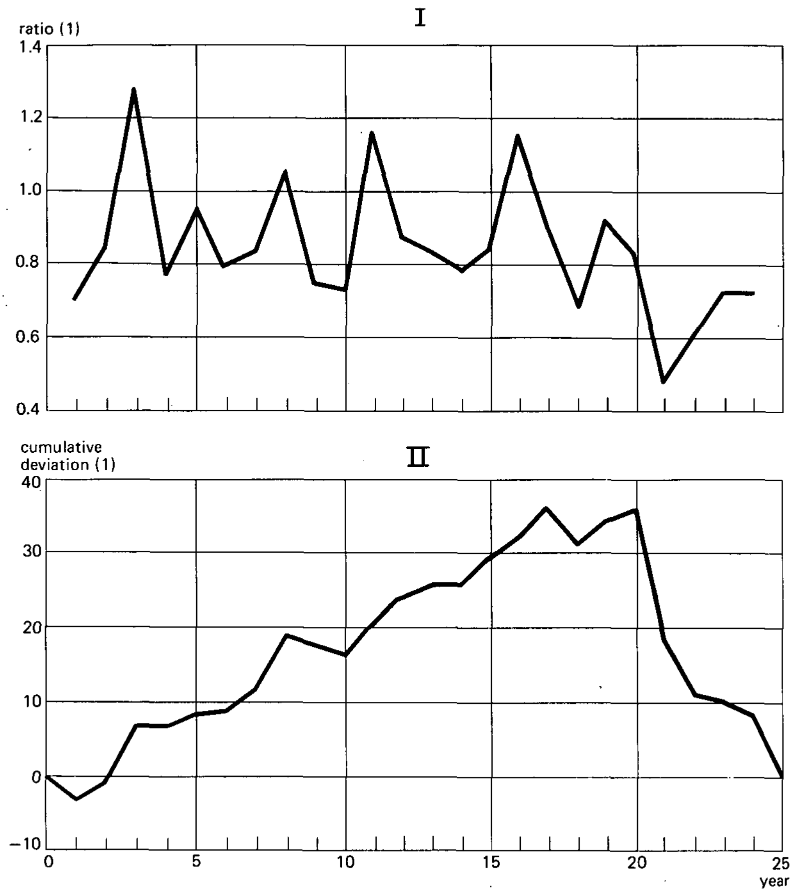

Figure 6.3 Cumulative departures from the mean slope (Plot I), and the proportionality factors, a_i, of the runoff from Stream A and a pattern (Plot II), as computed from the data in Table 6.4.

interpretation of the double-mass plot (a difficulty mentioned in Section 6.2), or from their prior knowledge of when the problem appeared.

Spearman's rank-correlation method confirms the trend in the time series of proportionality factors ($t_t = -2.13$ and $\nu = 23$). The computations of F_t and t_t for various combinations of sub-sets are shown in Table 6.5. The number of data of each sub-set in the column 'Sub-Set I' has been varied for Water Years 16 to 19. To keep each sub-set approximately equal, we have included the maximum number of data in Sub-Set I as a function of the number of data in Sub-Set II. None of the computed values of F_t falls within the critical region for all combinations of sub-sets. The variance is therefore stable, and one can apply the t-test for stability of mean. The mean of all the combinations, however, is not stable, proving that there is a break in the double-mass line in one of the water years from 1937 to 1940.

Table 6.5 Results of the computation of F_t and t_t for four combinations of sub-sets of proportionality factors

Sub-Set I (Water Years)	Sub-Set II (Water Years)	v_1,v_1	$F_{2.5\%}$ F_t $F_{97.5\%}$	v	$t_{2.5\%}$ t_t $t_{97.5\%}$
7-16	17-25	9,8	0.244 1.272 4.36	17	−2.10 2.46 2.10
9-17	18-25	8,7	0.221 1.441 4.90	15	−2.12 2.61 2.12
11-18	19-25	7,6	0.195 1.372 5.70	13	−2.14 2.43 2.14
13-19	20-25	6,5	0.169 1.468 6.98	11	−2.20 2.72 2.20

The basic data-screening procedure, when applied to the individual time series, shows that the time series from Stream A and the pattern are stationary and suitable for further use. A correlation with the rainfall on the catchments, through the intermediary of computed runoff, could explain the change in proportionality around water year 1938. Another explanation is that the change was only temporary.

7 References

Anderson, R.L. 1942. Distribution of the Serial Correlation Coefficient. In: *Annals of Mathematical Statistics* 13:1-13.

Bernier, J. 1977. Etude de la stationnarité des séries hydrométéorologiques. In: *Houille Blanche* 4:313-319.

Box, G.P. and G.M. Jenkins 1976. *Time-Series Analysis, Forecasting and Control.* Holden-Day, San Francisco.

Buishand, T.A. 1982. Some Methods for Testing the Homogeneity of Rainfall Records. In: *Journal of Hydrology* 58:11-27.

Chang, M. and R. Lee 1974. Objective Double-Mass Analysis. In: *Water Resources Research* 10(6):1123-1126.

d'Agostino, R.B. 1971. An Omnibus Test of Normality for Moderate and Large Size Samples. In: *Biometrika* 58(2):341-348.

Dyck, S. *u.a.* 1980. *Angewandte Hydrologie. Teil 1: Berechnung und Regelung des Durchflusses der Flüsse.* W. Ernst & Sohn, Berlin-München.

Hald, A. 1952. *Statistical Theory with Engineering Applications.* John Wiley & Sons, New York.

Hansel, N. and U. Schäfer 1970. Die Doppelsummenanalyse. In: *Wasserwirtschaft-Wassertechnik* 20(5):145-150.

Lackritz, J.R. 1984. Exact p-Values for F- and t-Tests. In: *The American Statistician* 38:312-314.

Merriam, C.F. 1937. A Comprehensive Study of the Rainfall on the Susquehanna Valley. In: *Transcripts. American Geophysical Union. Reports and Papers, Hydrology* 18:471-476.

Searcy, J.K. and C.H. Hardison 1960. *Double-Mass Curves.* Geological Water Supply Paper 1541-B. U.S. Geological Survey, Washington, D.C.

Singh, R. 1968. Double-Mass Analysis on the Computer. In: *Proceedings. American Society of Civil Engineers. Journal of Hydraulics* 1:139-142.

Spiegel, M.R. 1961. *Theory and Problems of Statistics.* Schaum Publishing Co., New York.

Stol, Ph.Th. 1965. On Quantative Methods for Analysing Hydrologic Data. In: *Proceedings of the Seminar of ECAFE Regions 1965/66 on the Use and Interpretation of Hydrologic Data.* United Nations' Water Resources Series 34. United Nations, New York.

Weiss, L.L. and W.T. Wilson 1953. Evaluation of Significance of Slope Changes in Double-Mass Curves. In: *Transcripts of the American Geophysical Union* 34(6):893-896.

World Meteorological Organization/WMO 1966. *Climatic Change.* J.M. Mitchell (Editor). WMO No. 195. TP. 100. WMO, Geneva.

Yevjevich, V. and R.I. Jeng 1969. *Properties of Non-Homogeneous Hydrologic Time Series.* Hydrology Paper 32. Colorado State University Press, Fort Collins.

Further Reading

Crow, E.L., F.A. Davis, and M.W. Maxfield 1960. *Statistics Manual.* Dover Publications, New York.

Haan, C.T. 1977. *Statistical Methods in Hydrology.* Iowa State University Press, Ames.

Roche, M. 1963. *Hydrologie de surface.* ORSTOM/Gauthier Villars, Paris.

Appendix 1

Percentile Points of the t-Distribution $t\{\upsilon,p\}$ for the 5-Per-Cent Level of Significance (Two-Tailed)

Appendix 2

Percentile Points of the F-Distribution $F\{\upsilon_1,\upsilon_2,p\}$ for the 5-Per-Cent Level of Significance (Two-Tailed)

Appendix 3

Additional Problems

Appendix 4

Answers to the Additional Problems

Appendix 1

Percentile Points of the t-Distribution t{ν,p for the 5-Per-Cent Level of Significance (Two-Tailed)

$p = P(t <= t_p)$:		0.025	0.975
ν:	4	−2.78	2.78
	5	−2.57	2.57
	6	−2.54	2.54
	7	−2.36	2.36
	8	−2.31	2.31
	9	−2.26	2.26
	10	−2.23	2.23
	11	−2.20	2.20
	12	−2.18	2.18
	14	−2.14	2.14
	16	−2.12	2.12
	18	−2.10	2.10
	20	−2.09	2.09
	24	−2.06	2.06
	30	−2.04	2.04
	40	−2.02	2.02
	60	−2.00	2.00
	100	−1.98	1.98
	160	−1.97	1.97
	∞	−1.96	1.96

Note: It is customary to take the next higher ν-value if the required number of degrees of freedom is not listed in a table. It is evident that this practice will produce a more severe test.

Appendix 2

Percentile Points of the F-Distribution $F\{\upsilon_1,\upsilon_2,p\}$ for the 5-Per-Cent Level of Significance (Two-Tailed)

$p = P(F <= F_p)$		ν_1: 4	5	6	7	8	9	10	11	12	14	16
0.025	ν_2: 5	.107	.140	.169								
0.975		7.39	7.15	6.98								
0.025	6		.143	.172	.195							
0.975			5.99	5.82	5.70							
0.025	7			.176	.200	.221						
0.975				5.12	4.99	4.90						
0.025	8				.204	.226	.244					
0.975					4.53	4.43	4.36					
0.025	9					.230	.248	.265				
0.975						4.10	4.03	3.96				
0.025	10						.252	.269	.284			
0.975							3.78	3.72	3.66			
0.025	11							.273	.288	.301		
0.975								3.53	3.47	3.43		
0.025	12								.292	.305	.328	
0.975									3.32	3.28	3.21	
0.025	14									.312	.336	.355
0.975										3.05	2.98	2.92

Appendix 2 (continued)

$p =$ $P(F <= F_p)$		ν_1:14	16	18	20	24	30	40	60	100	160	∞
0.025	ν_2:16	.342	.362	.379								
0.975		2.82	2.76	2.71								
0.025	18		.368	.385	.400							
0.975			2.64	2.60	2.56							
0.025	20			.391	.406	.430						
0.975				2.50	2.46	2.41						
0.025	24				.415	.441	.468					
0.975					2.33	2.27	2.21					
0.025	30					.453	.482	.515				
0.975						2.14	2.07	2.01				
0.025	40						.498	.533	.573			
0.975							1.94	1.88	1.80			
0.025	60							.555	.600	.642		
0.975								1.74	1.67	1.60		
0.025	100								.625	.674	.706	
0.975									1.56	1.48	1.44	
0.025	160									.696	.733	
0.975										1.42	1.36	
0.025	∞											1.00
0.975												1.00

Appendix 3

Additional Problems

Problem 1

There are three hydrometric stations near the town of Malema in the Malema River Basin in Mozambique: one in the main river (the Malema), one in a small tributary (the Mutivasse) and one in another tributary (the Nataleia). Concurrent records of runoff in millions of cubic metres (MCM) are available for the period from October 1959 to September 1981 (Table A3.1)

Table A3.1 Annual runoff (in MCM) at three hydrometric stations in the Malema River Basin, Mozambique, from 1959 to 1980 (water years)

	Hydrometric Station		
Water Year	Malema	Mutivasse	Nataleia
1959	445.7	13.6	39.8
1960	937.0	40.7	129.2
1961	1071.9	48.3	137.5
1962	1024.2	66.4	184.2
1963	532.0	42.4	95.1
1964	1073.2	57.4	145.1
1965	741.3	52.3	102.1
1966	767.2	27.7	82.5
1967	768.4	34.8	103.4
1968	832.8	36.0	96.6
1969	511.4	31.9	99.3
1970	865.7	49.8	170.3
1971	456.6	20.0	49.0
1972	447.6	32.4	105.0
1973	1350.0	68.5	237.0
1974	472.7	17.2	58.8
1975	1153.7	66.8	524.2
1976	503.6	31.3	122.9
1977	805.3	61.7	279.2
1978	753.4	62.1	195.8
1979	593.6	45.0	154.9
1980	533.7	38.8	126.9

a. Screen the data in the three time series according to the basic data-screening procedure.
b. Screen the proportionality factors of the three time series (three combinations).
c: Screen the proportionality factors of the average runoff from the Malema and Mutivasse Rivers vs those from the Nataleia River.
d. Discuss your results.

Problem 2

Table A3.2 gives the maximum one-day rainfall at a rainfall station over ten water years. Rough screening of the data has not been done.

Table A3.2 Maximum one-day rainfall (in mm) at a rainfall station over ten water years

1	2	3	4	5	6	7	8	9	10
121.7	307.3	180.1	118.4	185.9	276.7	430.8	434.1	395.1	189.2

a. Screen the data according to the basic data-screening procedure.
b. Discuss your results.

Problem 3

Table A3.3 gives the maximum one-day rainfall at the Bangkok Meteorological Department over thirty-four water years.

Table A3.3 Maximum one-day rainfall (in mm) at the Bangkok Meteorological Department from 1952 to 1985 (water years)

Water Year	One-Day Rainfall	Water Year	One-Day Rainfall
1952	111.0	1969	81.2
1953	84.1	1970	98.4
1954	53.8	1971	97.8
1955	108.8	1972	141.0
1956	69.6	1973	84.0
1957	105.1	1974	109.4
1958	73.5	1975	68.3
1959	82.9	1976	84.8
1960	123.2	1977	52.9
1961	75.8	1978	63.5
1962	69.8	1979	167.3
1963	90.4	1980	84.1
1964	114.7	1981	105.7
1965	93.3	1982	72.7
1966	124.2	1983	99.1
1967	54.1	1984	85.6
1968	153.7	1985	107.3

a. Screen the data according to the basic data-screening procedure.
b. Can the data series be used for frequency analysis?

Problem 4

Table A3.4 gives the mean annual runoff from the Bogowonto River at Bener, Indonesia, from September 1961 to August 1974.

Table A3.4 Mean annual runoff (in m^3/s) from the Bogowonto River at Bener, Indonesia, over thirteen water years

1	2	3	4	5	6	7	8	9	10	11	12	13
8.19	8.60	7.24	10.68	8.46	9.29	11.02	11.55	8.95	8.65	8.94	9.83	10.45

a. Screen the data according to the basic data-screening procedure.
b. Discuss your results.

Problem 5

Table A3.5 gives the annual rainfall at two rainfall stations (Sampan and Kedung Wringin) in the Wadaslintang area of Java, Indonesia, from September 1967 to August 1983 (water years).
The project hydrologist has suggested that the last six years of the Kedung Wringin data be corrected to the values given in the last column of the table.

Table A3.5 Annual rainfall (in mm) at two rainfall stations in the Wadaslintang area, Indonesia, from 1967 to 1982 (water years)

	Rainfall Station		
	Sampan	Kedung Wringin	
Water Year		uncorrected	corrected
1967	5524	3798	3798
1968	3086	2854	2854
1969	3778	2877	2877
1970	3343	2979	2979
1971	2951	2881	2881
1972	3650	3589	3589
1973	3447	2882	2882
1974	4120	4965	4965
1975	3693	3127	3127
1976	2679	2175	2175
1977	2361	8154	3561
1978	3670	6505	2841
1979	2071	5207	2274
1980	3988	7852	3429
1981	3320	6705	2928
1982	2963	5136	2243

a. Screen the data according to the basic data-screening procedure.
b. Perform a conventional double-mass analysis on the Sampan data and the uncorrected Kedung Wringin data.
c. Screen the proportionality factors of the Sampan data against the uncorrected and corrected Kedung Wringin data.
d. Discuss your results.

Problem 6

Table A3.6 gives the mean annual runoff of the Derwent River at the Yorkshire Bridge, U.K., for fifty-six calendar years.

Table A3.6 Mean annual runoff (in m^3/s) of the Derwent River at the Yorkshire Bridge from 1906 to 1961 (calendar years)*

Cal. Year	Mean Ann. Runoff	Cal. Year	Mean Ann. Runoff	Cal. Year	Mean Ann. Runoff	Cal. Year	Mean Ann. Runoff
1906	0.925	1920	1.088	1934	0.847	1948	0.970
1907	1.025	1921	0.752	1935	1.208	1949	0.926
1908	0.802	1922	0.986	1936	1.156	1950	0.988
1909	1.031	1923	1.090	1937	0.967	1951	1.272
1910	1.089	1924	0.872	1938	1.009	1952	1.034
1911	0.768	1925	0.779	1939	1.145	1953	0.792
1912	1.194	1926	0.894	1940	0.925	1954	1.478
1913	0.846	1927	1.002	1941	1.128	1955	0.786
1914	0.954	1928	0.932	1942	0.896	1956	1.187
1915	0.976	1929	0.835	1943	0.893	1957	1.019
1916	1.080	1930	1.155	1944	1.209	1958	1.219
1917	0.845	1931	1.161	1945	0.869	1959	0.710
1918	1.132	1932	1.005	1946	1.214	1960	1.266
1919	1.035	1933	0.695	1947	0.927	1961	1.010

* The mean annual runoff for the whole period was 3.07 m^3/s. Data are given as multiples of the mean annual discharge.

a. Screen the data according to the basic data-screening procedure.
b. Discuss your results.

Problem 7

Table A3.7 gives the mean annual runoff of the Rhine River at Lobith, The Netherlands, over seventy-two calendar years.

Table A3.7 Mean annual runoff (in m^3/s) of the Rhine River at Lobith, The Netherlands, from 1901 to 1972 (calendar years)*

Cal. Year	Mean Ann. Runoff	Cal. Year	Mean Ann. Runoff	Cal. Year	Mean Ann. Runoff	Cal. Year	Mean Ann. Runoff
1901	1.012	1919	1.053	1937	1.265	1955	1.133
1902	0.987	1920	1.199	1938	0.810	1956	1.000
1903	0.835	1921	0.504	1939	1.136	1957	1.010
1904	0.916	1922	1.082	1940	1.407	1958	1.093
1905	0.879	1923	1.161	1941	1.301	1959	0.812
1906	1.019	1924	1.239	1942	0.936	1960	0.815
1907	0.885	1925	0.902	1943	0.685	1961	1.149
1908	0.909	1926	1.238	1944	0.797	1962	1.039
1909	0.762	1927	1.217	1945	1.227	1963	0.747
1910	1.272	1928	0.894	1946	0.960	1964	0.669
1911	0.935	1929	0.744	1947	0.701	1965	1.258
1912	0.966	1930	0.980	1948	1.137	1966	1.442
1913	1.018	1931	1.431	1949	0.543	1967	1.224
1914	1.327	1932	0.908	1950	0.694	1968	1.303
1915	1.047	1933	0.801	1951	1.098	1969	1.049
1916	1.197	1934	0.606	1952	0.950	1970	1.364
1917	1.065	1935	0.988	1953	1.067	1971	0.737
1918	0.904	1936	1.202	1954	0.742	1972	0.605

* The mean annual runoff for the whole period was 2195.0 m^3/s. Data are given as multiples of the mean annual discharge.

a. Screen the data according to the basic data-screening procedure.
b. Discuss your results.

Problem 8

Table A3.8 presents the available data on the runoff of the Prek Thnot River at Anlong Touk, Cambodia, over twenty water years. The time series is not continuous. Determine whether the time series is stationary and, hence, suitable for frequency analyses and reservoir operation studies.

Table A3.8 Mean annual runoff (in MCM) of the Prek Thnot River at Anlong Touk, Cambodia, over twenty water years

Water Year	Mean Annual Runoff	Water Year	Mean Annual Runoff
1904	2579	1930	1109
1905	1080	1931	1510
1906	1337	1932	1663
1921	2086	1933	1040
1922	3969	1960	949
1924	2189	1961	1204
1926	1336	1962	1808
1927	1439	1963	616
1928	776	1964	1406
1929	1151	1965	1600

Appendix 4

Answers to the Additional Problems

We have performed all the tests at the recommended confidence level of 5 per cent. We have not described rough screening of the data. There are no data plots; our advice is to plot the data before performing the tests. A summary of the test results is in Table A4.1. Additional notes on the data series and all of the test results are at the end of the Appendix.

Table A4.1 Summary of the results of the data-screening tests

Problem Number		Two Sub-Sets 1 & 2		<--- Three Sub-Sets ---> 1 & 2		1 & 3		2 & 3		1st Ser. Correl. Coeff.
	Trend	F	t	F	t	F	t	F	t	
1 Malema	no	v	v	v	v	v	v	v	v	v
Mutivasse	no	v	v	v	v	v	v	v	v	v
Nataleia	no	*	–	v	v	*	–	*	–	v
Mutiva./Nataleia	neg	v	*	v	v	v	*	v	*	*
Malema/Mutiva.	neg	v	v	v	v	v	v	v	v	v
Malema/Nataleia	neg	v	v	v	*	v	*	v	*	v
(Mal.+Mut.)/Nataleia	neg	v	*	v	v	v	v	v	v	v
2 Max. 1-day rainfall	no	v	*	–	–	–	–	–	–	–
3 Bangkok 1-day rainfall	no	v	v	v	v	v	v	v	v	–
4 Bogowonto runoff	no	v	v	–	–	–	–	–	–	v
5 Sampan rainfall	no	v	v	v	v	v	v	v	v	–
K. Wringin rainfall	pos	*	–	*	–	*	–	v		–
K. Wringin corrected	no	v	v	*	–	v	v	v	v	–
Samp./K.Wr. ratio	neg	v	*	v	v	v	*	v	*	–
Samp./K.Wr. corrected	no	v	v	v	v	v	v	v	v	–
6 Derwent runoff	no	v	v	v	v	v	v	v	v	*
7 Rhine runoff	no	v	v	v	v	v	v	v	v	v
8 Prek Thnot runoff	no	*	–	*	–	*	–	v	v	v

v passed the test
* did not pass the test
– test not performed

Some Additional Notes

Problem 1

The runoff data for the Malema and Mutivasse Rivers are stationary and relatively consistent. Even so, one will have to keep checking the proportionality factors to make sure that the negative trend does not continue. The water level observations and rating curves at the Nataleia station have to be checked. The data cannot be used in their current form.

Problem 2

The time series is not stationary. One will have to ascertain which of the two parts of the time series is correct. (Inspection of the original observation sheets revealed that, during the last five years of the observation period, the rain gauge was read at weekly intervals, and not at daily intervals as was done – and correctly so – during the first half of the observation period. The earlier period contains, consequently, a series of correct observations that can be used at lower levels of aggregation.)

Problem 3

There is no objection to applying a frequency analysis to the data.

Problem 4

The Bogowonto runoff data pass all tests successfully.

Problem 5

The Sampan data are stationary, the uncorrected data for Kedung Wringin are not. The corrected Kedung Wringin data seem to be acceptable, notwithstanding a small instability in the variance. A plot of the double mass indicates a possibly significant break around Water Year 10, which agrees with the result obtained with the basic data-screening procedure.

The significance of the apparent break in the double-mass curve after Water Year 10 was confirmed by applying the basic data-screening procedure to the uncorrected proportionality factors between the two stations. The correction factor of 1/2.29 on the totals of the last six years is satisfactory because the corrected ratios passed all the tests.

(The measuring jar at the rainfall station at Kedung Wringin broke in February 1977. The observer, unable to obtain a new measuring jar, found a similar glass pot and copied the graduations from the remains of the original jar onto his improvised measuring device. Fortunately, this was discovered while the improvised jar was still in use, when the hydrologist who noticed the break in the double-mass line went to see the observer, and it was possible to correct the daily rainfall observations from Kedung Wringin fairly accurately. There are many other examples of observers who take care to record continuous observations in the best possible manner.)

Problem 6

There is some carry-over from surface storage in the Derwent River catchment. The data series cannot be used for frequency analyses in its current form.

Problem 7

The Rhine River data are stationary and suitable for use.

Problem 8

The data on the Prek Thnot River are not stationary. Nevertheless, the results of the basic data-screening procedure indicate that the data covering the last fourteen water years (1926 to 1933 and 1960 to 1965) can be joined in one continuous time series that is stationary. Those data would then be suitable for frequency analyses and reservoir operation studies.

Currently available ILRI publications

No.	Publications	Author	ISBN No.
11	Reclamation of saft-affected soils in Iraq	P. J. Dieleman (ed.)	–
15	Planning of service centres in rural areas of developing countries	D. B. W. M. van Dusseldorp	–
16	Drainage principles and applications (in 4 volumes)	–	90 70260 123, –131, –62 X and –63 8
16S	Principos y aplicaciones del drenaje (en 4 volúmenes)	–	–
17	Land evaluation for rural purposes	R. Brinkman and A. J. Smyth	90 70260 859
19	On irrigation efficiencies	M. G. Bos and J. Nugteren	90 70260 875
20	Discharge measurements structures. (3rd revised edition)	M. G. Bos	90 70754 150
21	Optimum use of water resources	N. A. de Ridder and A. Erez	–
24	Drainage and reclamation of salt-affected soils	J. Martinez Beltrán	–
25	Proceedings of the International Drainage Workshop	J. Wesseling (ed.)	90 70260 549
26	Framework for regional planning in developing countries	J. M. van Staveren and D. B. W. M. van Dusseldorp	90 70260 832
27	Land reclamation and water management	–	90 70260 689
29	Numerical modelling of groundwater basins: A user-oriented manual	J. Boonstra and N. A. de Ridder	90 70260 697
30	Proceedings of the Symposium on Peat Lands Below Sea Level	H. de Bakker and M. W. van den Berg	90 70260 700
31	Proceedings of the Bangkok Symposium an Acid Sulphate Soils	H. Dost and N. Breeman (eds.)	90 70260 719
32	Monitoring and evaluation of agricultural change	Josette Murphy and Leendert H. Sprey	90 70260 743
33	Introduction to farm surveys	Josette Murphy and Leendert H. Sprey	90 70260 735
34	Evaluation permanente du développement agricole	Josette Murphy and Leendert H. Sprey	90 70260 891
35	Introduction aux enquêtes agricoles en Afrique	Josette Murphy and Leendert H. Sprey	90 70260 956
36	Proceedings of the International Workshop on Land Evaluation for Extensive Grazing (LEEG)	W. Siderius (ed.)	90 70260 948
37	Proceedings of the ISSS Symposium on 'Water and solute movement in heavy clay soils'	J. Bouma, P. A. C. Raats (ed.)	90 72060 972
38	Aforadores de caudal para canales abiertos.	M. G. Bos, J. A. Replogle and A. J. Clemmens	90 70260 921
39	Acid Sulphate Soils: A baseline for research and development	D. Dent	90 70260 980
40	Land evaluation for land-use planning and conservation in sloping areas	W. Siderius (ed.)	90 70260 999
41	Research on water management of rice fields in the Nile Delta, Egypt	S. EL. Guindy & I. A. Risseeuw; H. J. Nijland (Ed.)	90 70754 08 8
43	BASCAD, a mathematical model for level basin irrigation	J. Boonstra and M. Jurriëns	90 70754 12 6
44	Selected Papers of the Dakar Symposium on Acid Sulphate Soils	H. Dost (ed.)	90 70754 13 4
45	Health and Irrigation Volume 2	J. M. V. Oomen, J. de Wolf and W. R. Jobin	90 70754 177
45	Health and Irrigation Volume 1	J. M. V. Oomen, J. de Wolf and W. R. Jobin	90 70754 21 5

46	CRIWAR: A Simulation Program for Calculating the Crop Irrigation Water Requirement of a Cropped Area	J. Vos, P. Kabat, M. G. Bos, R. A. Feddes,	90 70754 22 3
47	Analysis and Evaluation of Pumping Test Data	G. P. Kruseman and, N. A. de Ridder	90 70754 20 7
48	SATEM: Selected Aquifer Test Evaluation Methods	J. Boonstra	90 70754 19 3
49	Screening of hydrological data: Tests for stationarity and relative consistency	E. R. Dahmen and M. J. Hall	90 70754 23 1

No.	Bulletins	Author	ISBN No.
1	The auger hole method	W. F. van Beers	90 70260 816
4	On the calcium carbonate content of young marine sediments	B. Verhoeven	–
6	Mud transport studies in coastal water from the Western Scheldt to the Danish frontier	A. J. de Groot	–
8	Some nomographs for the calculation of drain spacings	W. F. J. van Beers	–
9	The Managil South-Western Extension: An extension to the Gezira Scheme	D. J. Shaw	–
10	A viscous fluid model for demonstration of groundwater flow to parallel drains	F. Homma	90 70260 824
11S	Análisis y evaluación de los datos de ensayos por bombeo	G. P. Kruseman and N. A. de Ridder	–
11F	Interprétation et discussion des pompages d'essai	G. P. Kruseman and N. A. de Ridder	–
12	Gypsifereous Soils	J. G. van Alphen and F. de los Rios Romero	–
13	Groundwater hydraulics of extensive aquifers	J. H. Edelman	90 70260 794
No.	Bibliographies		
7	Agricultural extension in developing countries	C. A. de Vries	–
8	Bibliography on cotton irrigation	C. J. Brouwer and L. F. Abell	–
9	Annotated bibliography or surface irrigation methods	S. Raadsma, G. Schrale	–
10	Soil Survey interpretation	R. H. Brook	–
13	Abstract journals or irrigation, drainage and water resources engineering	L. F. Abell	–
18	Drainage: An annotated guide to books and journals	G. Naber	90 70260 93 X

Other publications

Paper International Symposium. Polders of the World (3 volumes)	90 70260 751 76 X and –778
Final Report Symposium Polders of the World	–
Proceedings Symposium Lowland. Development in Indonesia.	90 70754 07 X