The Standardized Precipitation Index

Theory

The Standardized Precipitation Index (SPI) is a tool which was developed primarily for defining and monitoring drought. It allows an analyst to determine the rarity of a drought at a given time scale (temporal resolution) of interest for any rainfall station with historic data. It can also be used to determine periods of anomalously wet events. The SPI is not a drought prediction tool.

Mathematically, the SPI is based on the cumulative probability of a given rainfall event occurring at a station. The historic rainfall data of the station is fitted to a gamma distribution, as the gamma distribution has been found to fit the precipitation distribution quite well. This is done through a process of maximum likelihood estimation of the gamma distribution parameters, α and β. In simple terms, the process described above allows the rainfall distribution at the station to be effectively represented by a mathematical cumulative probability function. Therefore, based on the historic rainfall data, an analyst can then tell what is the probability of the rainfall being less than or equal to a certain amount. Thus, the probability of rainfall being less than or equal to the average rainfall for that area will be about 0.5, while the probability of rainfall being less than or equal to an amount much smaller than the average will be also be lower (0.2, 0.1, 0.01 etc, depending on the amount). Therefore if a particular rainfall event gives a low probability on the cumulative probability function, then this is indicative of a likely drought event. Alternatively, a rainfall event which gives a high probability on the cumulative probability function is an anomalously wet event. For more information on the mathematical background to the SPI, please refer to Dan Edward's masters' thesis, available at ftp://ccc.atmos.colostate.edu/pub/spi.pdf . You may also visit the SPI webpage of the Colorado Climate Center at http://ulysses.atmos.colostate.edu/SPI.html)

In the instance described above, rainfall is the variate in a gamma distribution function. The function will have a standard deviation and a mean which depends on the rainfall characteristics of that area. If a probability function for a station in a different area is calculated, it will most likely have a very different standard deviation and a different mean. Therefore it will be very difficult to compare rainfall events for two or more different areas in terms of drought, as drought is really a "below-normal" rainfall event. And what is "normal rainfall" in one area can be surplus rainfall in another area, speaking strictly in terms of rainfall amounts.

What therefore is the solution? The solution is to transform your cumulative probability gamma function into a standard normal random variable Z with mean of zero and standard deviation of one. A new variate is formed, and the transformation is done in such a way that each rainfall amount in the old (gamma) function has got a corresponding value in the new (transformed) Z function. And the probability that the rainfall is less than or equal to any rainfall amount will be the same as the probability that the new variate is less than or equal to the corresponding value of that rainfall amount. All probability functions which have been fitted for different rainfall station data can be transformed in this way, and the resultant transformed variate is always in the same units. This is where SPI comes in.

I mentioned earlier that the standard normal distribution has a mean of zero and a standard deviation of one. The SPI is a representation of the number of standard deviations from the mean at which an event occurs, often called a “z-score”. The unit of the SPI can thus be considered to be “standard deviations”. Standard deviation is often described as the value along a distribution at which the cumulative probability of an event occurring is 0.1587. In a like manner, the cumulative probability of any SPI value can be found, and this will be equal to the cumulative probability of the corresponding rainfall event. Table 1. below gives the cumulative probabilities for various SPI values.

SPI	Cumulative Probability
-3.0	0.0014
-2.5	0.0062
-2.0	0.0228
-1.5	0.0668
-1.0	0.1587
-0.5	0.3085
0.0	0.5000
0.5	0.6915
1.0	0.8413
1.5	0.9332
2.0	0.9772
2.5	0.9938
3.0	0.9986

Table 1: SPI and cumulative Probabilities

In summary therefore, the SPI can effectively represent the amount of rainfall over a given time scale, with the advantage that it provides not only information on the amount of rainfall, but that it also gives an indication of what this amount is in relation to the normal, thus leading to the definition of whether a station is experiencing drought or not. It gives output in units of standard deviation from the average based on as-long-a-rainfall-distribution-as-there-is-data-for. The longer the period used to calculate the distribution parameters, the more likely you are to get better results (e.g. 50 years better than 20 years). Therefore, you can use a very long time period (e.g. 1920-1998) to calculate the parameters of the distribution and then extract the SPI values for only a given time period (could be one year, or a number of years to give a time series).

The theory section (above) is based on Chapter 3 of Dan Edward’s masters’ thesis.

<u>Uses of the SPI</u>

Some of the advantages which can be derived from the SPI are as follows. Plotting a time series of year against SPI gives a good indication of the drought history of a paricular station. Figure 1 below shows the SPI calculated for 6 month rainfall totals (October-November-December-January-February-March) from 1980 to 1998, based on a gamma distribution function fitted for data from 1922 – 1998. Figure 2 below is a graph showing the actual 6-monthly rainfall totals from the corresponding time period. As an interpretation key, the amount shown on the scale for 1987 is the sum of rainfall from October 1986 to March 1987. These two graphs show quite well, from the similar trend they follow, the close relationship between the SPI and actual rainfall.

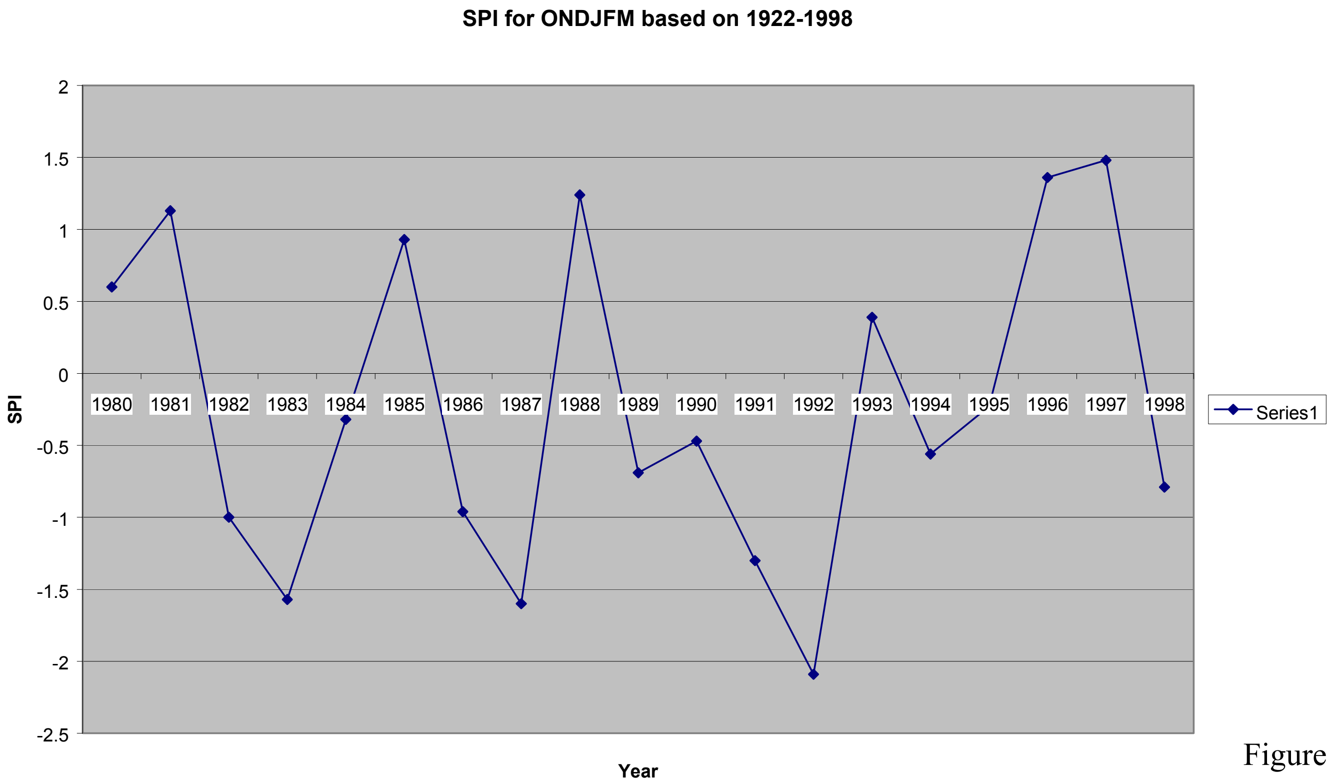


Figure 1

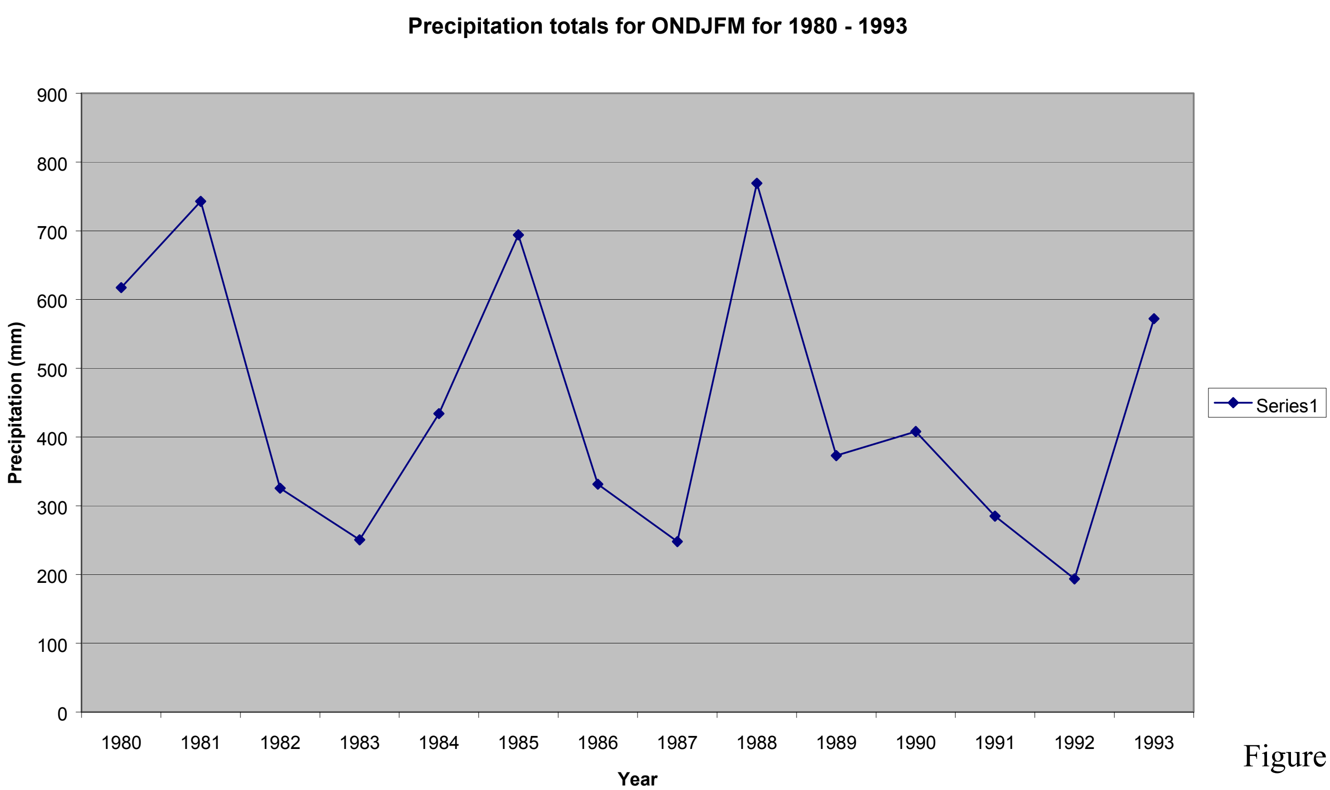


Figure 2

To summarize the interpretation of the SPI:

High SPI value (closer to 3) : heavy precipitation event over time period specified
Medium SPI value (approximately = 0) : normal precipitation event over time period specified
Low SPI value (closer to -3) : low precipitation event over time period specified (drought ?)

Remember that the heaviness or lowness of a precipitation event in the SPI is relative to the rainfall characteristics of that area.

The advantage of the SPI approach in this regard however, is that it can be used more beneficially than rainfall in spatial analysis for drought. This is because of the possibility of comparing different stations in different climatic regions (or climatic sub-regions) regardless of the fact that they may have different normal rainfalls. This being because in SPI, the rainfall is already normalized and compares the current rainfall with the average. Therefore the rainfall of two areas with different rainfall characteristics can be compared in terms of how badly they are experiencing drought conditions since the comparison is in terms of their normal rainfall. The concept is somewhat similar to a "rainfall difference-from-average" map. The advantage of the SPI over the difference-from-average map however is that while the latter will only indicate the numerical magnitude of variation (e.g. 40mm less, 10mm more etc) without stating how much less than what (is it 40mm less than 80mm, or 40mm less than 200mm), the SPI shows the statistical magnitude of deviation from the average, and therefore better portrays the seriousness of the shortage

At a more advanced level of analysis, a raster SPI map can be generated form a series o rainfall estimation images, using advanced GIS analyses. The SEDI routine could be used to generate such rainfall estimation images from as far back as the data goes (c.f. SADC Climate Information System). Another possibility in terms of use of SPI would be monitoring by assessing the change in SPI over short time periods (e.g., time series for SPI_{oct}, SPI_{nov}, SPI_{dec}, SPI_{jan}, SPI_{feb}, SPI_{mar}. One plausible question one might like to ask is where an investigation into the relationship between SPI and CCDs could lead. A lot of functions for which rainfall and related data is currently being used could equally be used with SPI for drought monitoring.

One other use to which the SPI is currently being applied is the studying of a site or region's drought history, including an analysis of frequency and duration for a magnitude ranking of sorts. This is currently being done on the UNIX version, through other programs which use the SPI.

The advantage of the SPI from a temporal view-point is related to the program itself. The program allows you to calculate SPI for rainfall totals of different time periods – namely, 3 month, 6 month, 12 month and 24 month rainfall totals. This allows the analyst to study time rainfall events at different time scales. The time scales are, unfortunately hard coded into the program so they cannot be changed to time scales which we are more familiar with, such as one dekad, or one month. This is only true for the PC version o the SPI program, as it seems that time lengths can be specified by the user on the UNIX version. It should however be possible to recompile the PC version to allow the entry of user defined time lengths. Appendix 1 shows some graphs which were generated from SPI for 3 month, 6 month and 12 month rainfall totals over different time periods.

Time periods which could prove to be interesting to the analyst are suggested below

OND
JFM
ONDJFM
JFMAMJJASOND
JASONDJFMAMJ

Each month is represented by its first letter e.g. OND = October November December (Total precipitation for those three months).

Program operation

The SPI program is relatively easy to operate.
Prepare an input file with all your data for one station in the following format

```
Header
yyyy mm pppp
yyyy mm pppp
yyyy mm pppp
yyyy mm pppp
yyyy mm pppp
yyyy mm pppp
yyyy mm pppp
yyyy mm pppp
etc
```

where Header = a string which describes the file, or something about the station, etc
yyyy = year
mm = month (in digit format 1,2,3 etc)
pppp = precipitation multiplied by 100
do not put "etc" in the input file, but more like entries.

The yyyy mm and pppp may either be separated by space or commas.
Missing values are denoted by -9900

Name this file stdin (with no extension)and put it in the same directory as the SPI program. Go to that directory at the DOS prompt and type SPI. Your output will be written to the file stdout (in the same directory), and any error messages will be written to the file stderr.

If the program runs correctly, there should be nothing written to the stderr file (and it will be zero bytes in size). However, some possible causes of error which may cause the program to not run properly are:

1. The precipitation values entered contain decimals <u>after</u> multiplying by 100. The actual input precipitation values should be integers (after multiplying by 100, that is), so your actual precipitation should not contain more than 2 decimal places
2. The input file contains too few values. There should be at least a minimum number of precipitation values for each month. I do not know how many is the minimum, but if you enter data for only one year, you will get an error, and nothing will be calculated. An example of what your input file should look like is given in appendix 2.

The output of the program is in the following format

Header
yyyy mm spi3 spi6 spi12 spi24
yyyy mm spi3 spi6 spi12 spi24
yyyy mm spi3 spi6 spi12 spi24
yyyy mm spi3 spi6 spi12 spi24
yyyy mm spi3 spi6 spi12 spi24
yyyy mm spi3 spi6 spi12 spi24
yyyy mm spi3 spi6 spi12 spi24
etc
Where yyyy and mm are as before
spi3 = SPI for a 3 month rainfall total
spi6 = SPI for a 6 month rainfall total
spi12 = SPI for a 12 month rainfall total
spi24 = SPI for a 24 month rainfall total

An sample output file is shown in appendix 3.

Put in as many years as possible within a single climatic period (I mention this because of the possibility of climate change) and only extract the SPIs you are interested in. I have written some short programs to extract periods of interest, and I will make them publicly available as soon as they are more user-friendly. In the same vein, macros can be written in Excel to automatically plot any curves of interest.

In the interpretation of the results be sure to remember that the SPI value in, say, the 3-month column, is based on the rainfall total for 3 months, including the month against which the SPI value is plotted. For example, the 3-month SPI for March, 1977 is based on the total rainfall for January, February, and March 1977. The same is true for 6 month, 12 month, and 24 month totals

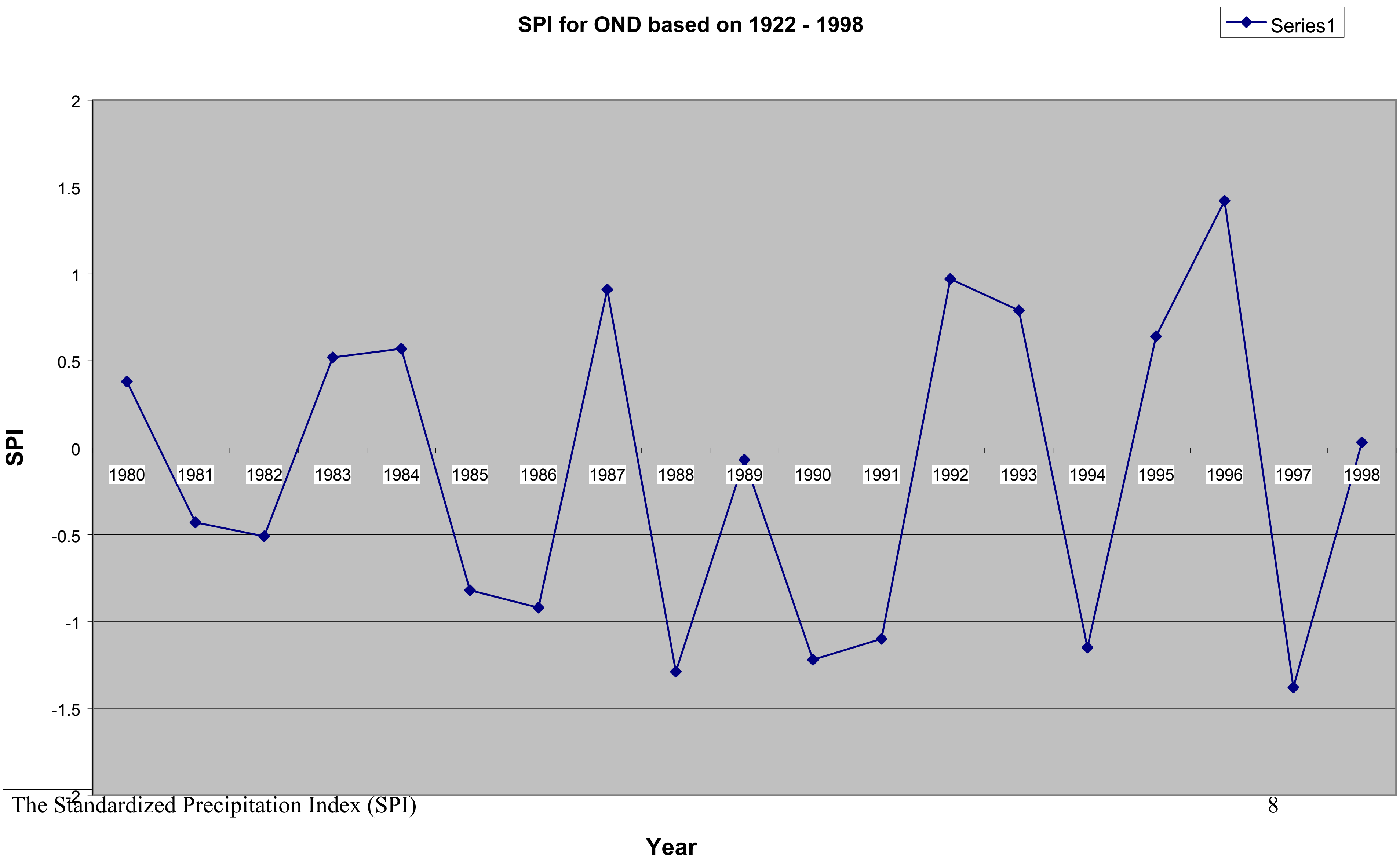
Disadvantages of the program

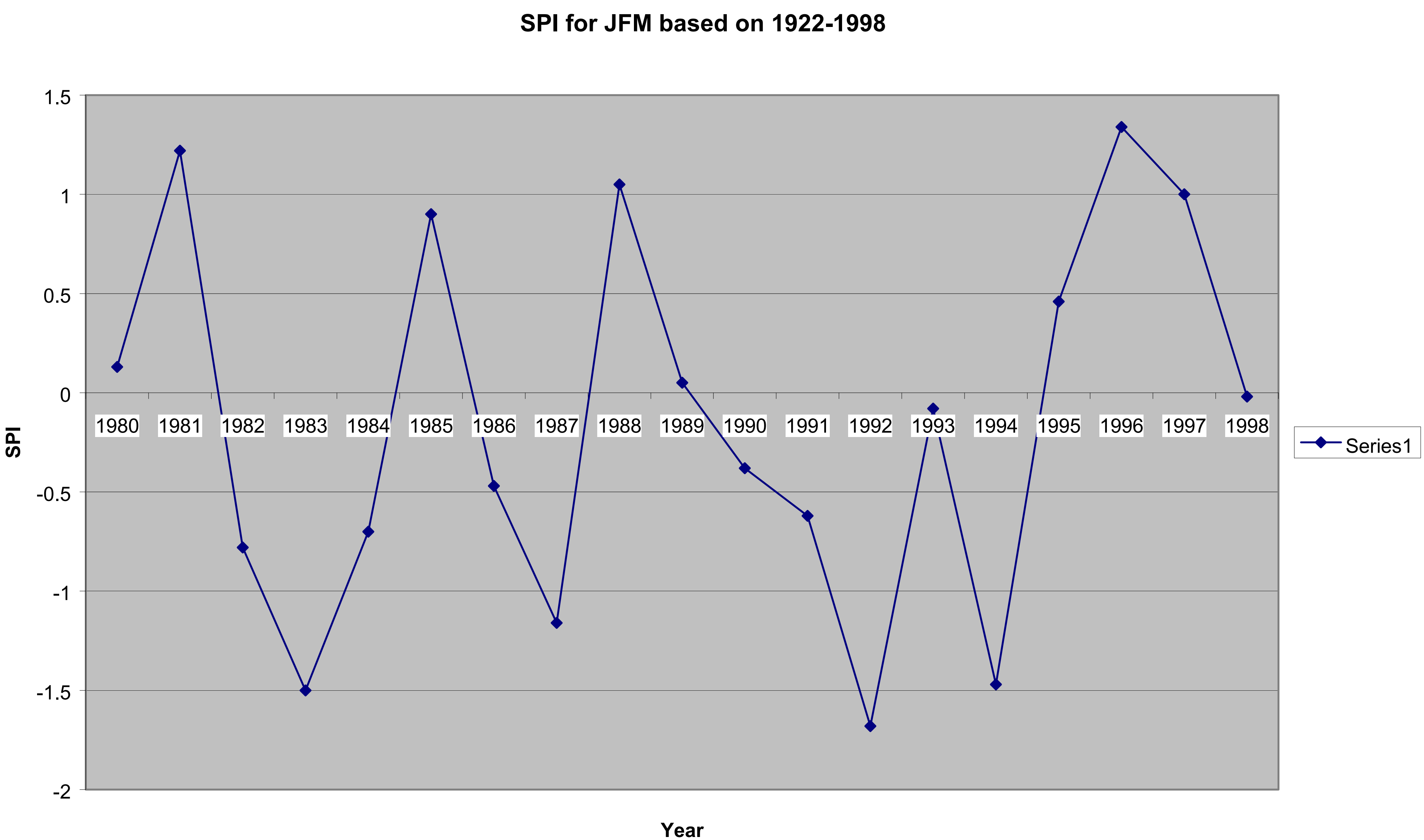
Currently there are 2 disadvantages to the SPI program. One is that time scale of analysis are hardcoded into the program and cannot be changed. The second is that, according to the source code, it would seem that the program can only be used for rainfall data between 1880 and 2000. We might say that it is not year 2001 compliant. This however can be corrected by recompiling the program.

SADC Regional Remote Sensing Unit

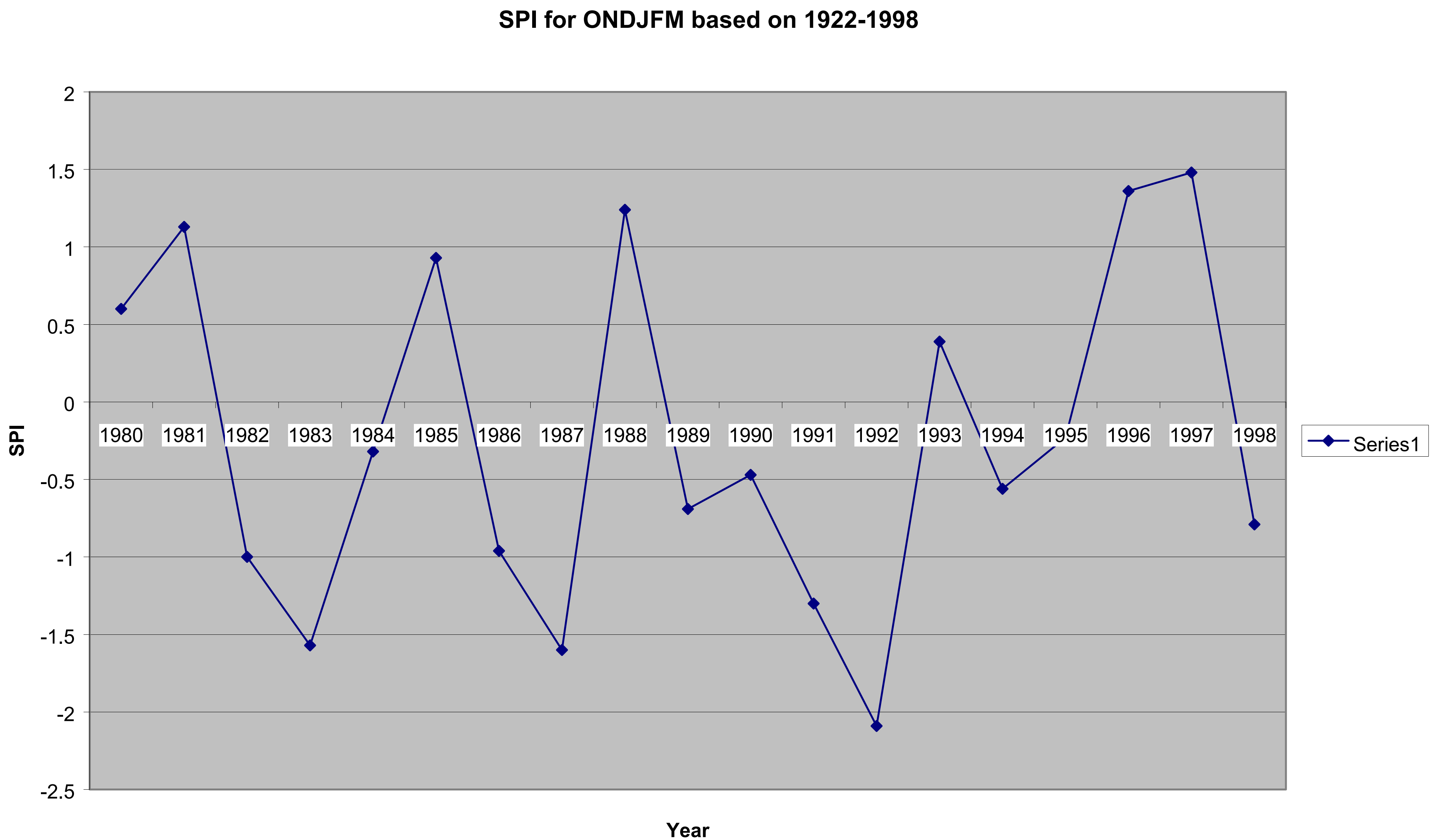
January 2000

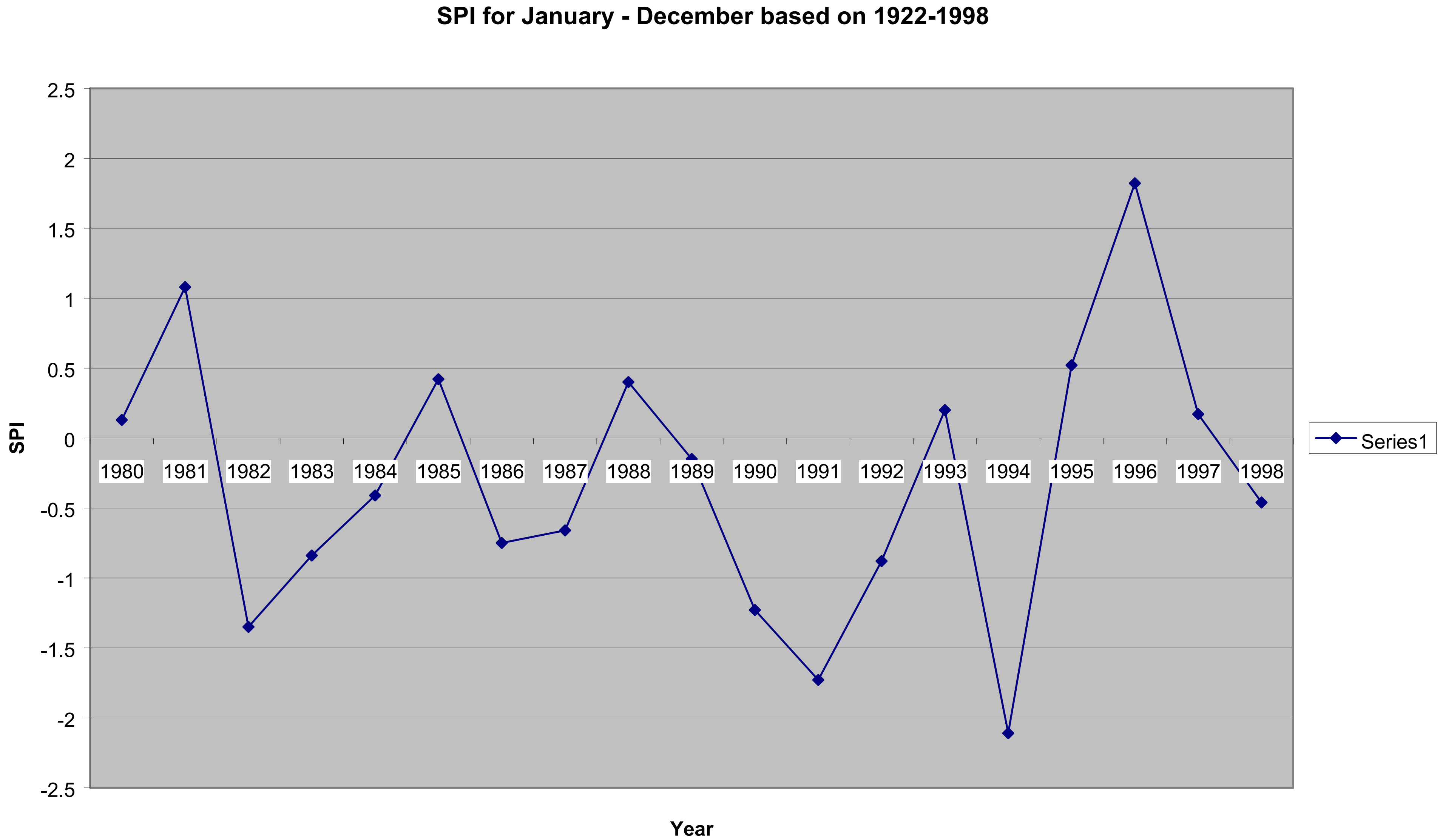
Appendix 1: Graphs generated from SPI for 3, 6, and 12 month rainfall totals





SPI for JFM based on 1922-1998
1.5
1
0.5
0
-0.5
-1
-1.5
-2
SPI
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
Series1
Year


SPI for ONDJFM based on 1922-1998
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
SPI
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
Year
Series1


SPI for January - December based on 1922-1998
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
SPI
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
Year
Series1

Appendix 2: Sample input file – The file has been truncated

The preciptation for 1970 through 1990 multiplied by 100 (for SPI program)
1970 1 610
1970 2 4290
1970 3 430
1970 4 2570
1970 5 0
1970 6 3330
1970 7 200
1970 8 0
1970 9 0
1970 10 740
1970 11 4420
1970 12 9750
1971 1 16130
1971 2 1730
1971 3 280
1971 4 5000
1971 5 970
1971 6 0
1971 7 0
1971 8 0
1971 9 770
1971 10 2760
1971 11 11480
1971 12 7690
1972 1 36020
1972 2 7970
1972 3 2430
1972 4 10150
1972 5 3050
1972 6 230
1972 7 0
1972 8 110
1972 9 510
1972 10 3960
1972 11 2410
1972 12 1720
1973 1 5420
1973 2 370
1973 3 830
1973 4 950
1973 5 380
1973 6 900
1973 7 130

1973 8 840
1973 9 5430
1973 10 3290
1973 11 8080
1973 12 36770
1974 1 15730
1974 2 14670
1974 3 4600
1974 4 3390
1974 5 540
1974 6 0
1974 7 100
1974 8 80
1974 9 1700
1974 10 290
1974 11 18290
1974 12 24270
1975 1 8050
1975 2 19920
1975 3 7980
1975 4 9930
1975 5 3020
1975 6 210
1975 7 0
1975 8 1360
1975 9 0
1975 10 3030
1975 11 2010
1975 12 10040
1976 1 3820
1976 2 12080
1976 3 15650
1976 4 2870
1976 5 3180
1976 6 380
1976 7 70
1976 8 0
1976 9 470
1976 10 5160
1976 11 9610
1976 12 8320

Appendix 3: Sample output file – the file has been truncated

The preciptation for 1970 through 1990 multiplied by 100 (for SPI program)

```
1970  3  -2.12 -99.00 -99.00 -99.00
1970  4  -1.06 -99.00 -99.00 -99.00
1970  5  -1.14 -99.00 -99.00 -99.00
1970  6    .42  -1.59 -99.00 -99.00
1970  7   1.08   -.82 -99.00 -99.00
1970  8   1.73   -.56 -99.00 -99.00
1970  9  -1.39   -.01 -99.00 -99.00
1970 10  -1.89   -.66 -99.00 -99.00
1970 11  -1.11   -.53 -99.00 -99.00
1970 12   -.57   -.67  -2.13 -99.00
1971  1    .15   -.20  -1.02 -99.00
1971  2   -.30   -.67  -1.01 -99.00
1971  3   -.58   -.83   -.97 -99.00
1971  4  -1.10   -.44   -.68 -99.00
1971  5   -.36   -.40   -.58 -99.00
1971  6    .43   -.46   -.74 -99.00
1971  7   -.45  -1.19   -.75 -99.00
1971  8  -1.67   -.62   -.76 -99.00
1971  9   -.45    .15   -.77 -99.00
1971 10   -.35   -.58   -.64 -99.00
1971 11    .63    .47   -.37 -99.00
1971 12    .19    .10   -.56  -1.78
1972  1   1.33   1.26    .48   -.35
1972  2    .88    .90    .70   -.23
1972  3   1.05    .76    .77   -.16
1972  4    .23    .94    .87    .10
1972  5    .91    .92    .90    .19
1972  6   1.60   1.30    .91    .11
1972  7    .98    .31    .92    .10
1972  8   -.56    .83    .93    .10
1972  9   -.62   1.46    .95    .11
1972 10   -.03    .28    .97    .21
1972 11   -.69   -.84    .73    .16
1972 12  -1.62  -1.59    .66   -.03
1973  1  -1.57  -1.64  -1.31   -.50
1973  2  -2.14  -2.11  -1.64   -.55
1973  3  -1.88  -2.22  -1.68   -.52
1973  4  -2.17  -2.15  -2.11   -.62
1973  5  -1.45  -2.20  -2.22   -.61
1973  6   -.62  -1.90  -2.14   -.58
1973  7   -.07  -2.04  -2.15   -.58
1973  8    .90  -1.17  -2.10   -.55
```

1973	9	2.06	.55	-1.81	-.41
1973	10	1.06	.88	-1.79	-.39
1973	11	.86	.93	-1.59	-.53
1973	12	2.10	2.20	.40	.53
1974	1	1.54	1.73	.83	-.23
1974	2	1.45	1.50	1.30	-.02
1974	3	.51	1.26	1.39	.05
1974	4	.38	1.16	1.30	-.15
1974	5	.03	1.18	1.26	-.22
1974	6	-.05	.40	1.23	-.22
1974	7	-.83	.27	1.24	-.22
1974	8	-.91	-.16	1.22	-.23
1974	9	.39	-.08	1.16	-.19
1974	10	-.95	-1.25	1.04	-.31
1974	11	1.25	1.13	1.47	.19
1974	12	1.78	1.67	1.36	1.01
1975	1	1.14	.98	.94	1.06
1975	2	.90	1.11	1.04	1.49
1975	3	.56	1.12	1.13	1.65
1975	4	1.22	1.29	1.21	1.79
1975	5	1.41	1.10	1.24	1.79
1975	6	1.56	.87	1.24	1.78
1975	7	.96	1.28	1.25	1.79
1975	8	.71	1.49	1.29	1.80
1975	9	.06	1.53	1.29	1.70
1975	10	-.08	.23	1.34	1.70
1975	11	-1.14	-.97	.92	1.65
1975	12	-.55	-.52	.51	1.08
1976	1	-.89	-1.02	.19	.66
1976	2	-.41	-.78	-.20	.50
1976	3	.32	-.12	.18	.82
1976	4	.85	-.05	-.10	.78
1976	5	1.47	.22	-.07	.82
1976	6	.52	.36	-.06	.84
1976	7	1.13	.93	-.06	.84
1976	8	-.37	1.46	-.12	.84
1976	9	-.73	.20	-.12	.82
1976	10	.25	.56	-.02	.95
1976	11	.66	.56	.29	.76
1976	12	.30	.19	.30	.38
1977	1	.19	.15	.69	.49
1977	2	.51	.57	1.01	.48
1977	3	1.17	.90	.93	.67
1977	4	1.14	.69	.75	.36
1977	5	.89	.63	.65	.29
1977	6	-1.26	.92	.64	.29

1977	7	-.44	1.09	.64	.30
1977	8	.10	.86	.67	.28
1977	9	2.19	.41	.88	.45
1977	10	1.26	1.00	.82	.48
1977	11	.57	.53	.70	.59
1977	12	.59	.94	1.35	.95
1978	1	1.02	1.25	1.81	1.56
1978	2	1.45	1.42	1.66	1.73
1978	3	1.46	1.24	1.35	1.48
1978	4	1.47	1.37	1.55	1.65
1978	5	1.41	1.53	1.50	1.55
1978	6	1.65	1.68	1.51	1.55
1978	7	.25	1.47	1.52	1.57
1978	8	-.42	1.39	1.51	1.57
1978	9	.41	1.70	1.44	1.61
1978	10	1.23	1.12	1.55	1.71
1978	11	.93	.84	1.77	1.74
1978	12	.61	.59	1.98	2.01
1979	1	.09	.33	1.35	1.96
1979	2	-.06	.18	.61	1.47
1979	3	.17	.27	.62	1.27
1979	4	.05	-.01	.13	1.23
1979	5	.04	-.09	.08	1.18
1979	6	-1.33	-.09	.10	1.18
1979	7	-.65	-.06	.10	1.20
1979	8	.34	.01	.12	1.19
1979	9	-.19	-1.46	.05	1.06
1979	10	.90	.60	.02	1.17
1979	11	1.04	1.02	.14	1.38
1979	12	1.05	.93	.38	1.44
1980	1	.30	.43	.14	.93
1980	2	.40	.61	.53	.65
1980	3	.21	.51	.29	.51
1980	4	.29	.25	.29	.15
1980	5	-.91	.10	.30	.12
1980	6	-1.17	-.03	.30	.12
1980	7	-.84	.18	.29	.12
1980	8	.10	-1.00	.30	.13
1980	9	.84	-.64	.37	.14
1980	10	-.68	-1.00	.06	-.12
1980	11	.63	.58	.16	.06
1980	12	.33	.41	.05	.13