The Standardised Groundwater drought Index (SGI)

Worked example 5.7, Example of how to estimate SGI using data from Stonor Park, UK

Table of Contents

In order to compare features of groundwater droughts using groundwater level data from different boreholes, Bloomfield and Marchant (2013) introduced the Standardised Groundwater level Index (SGI). The SGI uses the normal scores transform (Everitt, 2002), a nonparametric normalisation method which assigns a value to ranked observation of groundwater levels for a given month from a given hydrograph. A non-parametric approach to standardisation was favoured by Bloomfield and Marchant (2013) as they showed that no consistent parametric models could be fitted to a wide range of groundwater hydrographs, and that even when a hydrograph for a single site is considered no consistent parametric model could be fitted for all months of the year. Unlike SPI, SGI is based on a continuous variable and requires no accumulation period, however, Bloomfield and Marchant (2013) defined an SPI accumulation period (, in months) that gave a maximum correlation between SPI and SGI for a given site.

There is no commonly agreed definition of groundwater drought status based on SGI. However, recently Bloomfield et al. (2019) defined any month with an SGI of −1 or less as being a groundwater drought month and periods of continuously negative SGI reach a monthly intensity of −1 or less was defined as an episode of groundwater by analogy with the World Meteorological Organisation definition of an SPI drought (WMO, 2012).

# Load the data

Here we illustrate how to estimate SGI from a groundwater level time series using data from a well at Stonor Park, UK, previously described in Chapter 3. It is recommended that the standardisation is applied to data from a period of at least 30 years and that when comparing SGI from more than one site that standardisation is undertaken over a common time period. In this case, groundwater level data for Stonor Park is available for a 40 year period.

library(tidyverse)  
library(lubridate)  
library(hydroDrought)  
stonor

## # A tibble: 2,024 x 2  
## time level  
## <date> <dbl>  
## 1 1969-12-28 75.0  
## 2 1970-01-04 74.6  
## 3 1970-01-11 74.2  
## 4 1970-01-18 73.9  
## 5 1970-01-25 73.6  
## 6 1970-02-01 73.5  
## 7 1970-02-08 73.7  
## 8 1970-02-15 74.0  
## 9 1970-02-22 74.2  
## 10 1970-03-01 74.5  
## # … with 2,014 more rows

# Create a regular times series of monthly data

* Step 0: The estimation of SGI requires data to be on a regular time step, in this case we will be using monthly data. The level data (recorded as metres above sea level, m aSL) from Stonor Park is already approximately on a monthly basis so we have linearly interpolated the levels to the first day of each month. Use your interpolation method of choice, or if you have more frequent observations, such as those produced by data logging systems, to sub-set the data onto a monthly time step.

times <- seq(as.Date("1970-01-01"), as.Date("2009-12-01"), by = "1 month")  
  
stonor.monthly <- approx(x = stonor$time, y = stonor$level, xout = times) %>%  
 as\_tibble() %>%  
 rename(time = x, level = y) %>%   
 mutate(  
 month = month(time, label = TRUE, abbr = FALSE)  
 )  
  
stonor.monthly

## # A tibble: 480 x 3  
## time level month   
## <date> <dbl> <ord>   
## 1 1970-01-01 74.8 January   
## 2 1970-02-01 73.5 February   
## 3 1970-03-01 74.5 March   
## 4 1970-04-01 76.1 April   
## 5 1970-05-01 77.1 May   
## 6 1970-06-01 77.6 June   
## 7 1970-07-01 77.3 July   
## 8 1970-08-01 76.1 August   
## 9 1970-09-01 74.4 September  
## 10 1970-10-01 72.8 October   
## # … with 470 more rows

# Calculate an SGI values

* Step 1: Extract the level data for an individual month from the full groundwater level time series. For example, in the spreadsheet example we have extracted the groundwater levels for each January in the Stonor Park record.
* Step 2: Order the level data for a given month from lowest to highest and estimate the standardised rank for each level, i.e. rank/number of observations in a given month + 1.
* Step 3: Estimate the inverse standardised normal cumulative value (mean 1, s.d. 0) from the standardise rank for each level. This value is the SGI value. In Microsoft Excel this value is returned by the =NORM.S.INV(cell) function, in R it is returned by the qnorm() function.

x <- stonor.monthly %>%  
 group\_by(month) %>%  
 mutate(  
 rank = rank(level),   
 standardised.rank = rank / (n() + 1),  
 sgi = qnorm(standardised.rank)  
 )  
  
x

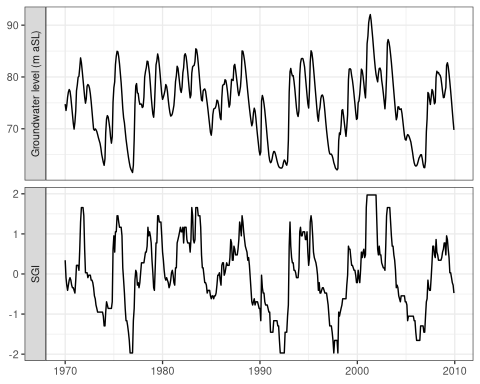
## # A tibble: 480 x 6  
## # Groups: month [12]  
## time level month rank standardised.rank sgi  
## <date> <dbl> <ord> <dbl> <dbl> <dbl>  
## 1 1970-01-01 74.8 January 26 0.634 0.343   
## 2 1970-02-01 73.5 February 19 0.463 -0.0918  
## 3 1970-03-01 74.5 March 16 0.390 -0.279   
## 4 1970-04-01 76.1 April 14 0.341 -0.408   
## 5 1970-05-01 77.1 May 16 0.390 -0.279   
## 6 1970-06-01 77.6 June 18 0.439 -0.153   
## 7 1970-07-01 77.3 July 19 0.463 -0.0918  
## 8 1970-08-01 76.1 August 18 0.439 -0.153   
## 9 1970-09-01 74.4 September 16 0.390 -0.279   
## 10 1970-10-01 72.8 October 15 0.366 -0.343   
## # … with 470 more rows

* Step 4: Repeat steps 1 to 3 for data for each calendar month separately. You will end up with 12 sets of monthly level data with associated inverse standardised normal cumulative values, or SGI values.

x %>%  
 nest()

## # A tibble: 12 x 2  
## # Groups: month [12]  
## month data   
## <ord> <list>   
## 1 January <tibble [40 × 5]>  
## 2 February <tibble [40 × 5]>  
## 3 March <tibble [40 × 5]>  
## 4 April <tibble [40 × 5]>  
## 5 May <tibble [40 × 5]>  
## 6 June <tibble [40 × 5]>  
## 7 July <tibble [40 × 5]>  
## 8 August <tibble [40 × 5]>  
## 9 September <tibble [40 × 5]>  
## 10 October <tibble [40 × 5]>  
## 11 November <tibble [40 × 5]>  
## 12 December <tibble [40 × 5]>

* Step 5: Combine SGI values with associated dates estimated in steps 3 and 4 and re-order oldest to most recent.



# Fast-Track

stonor.monthly %>%   
 group\_by(month) %>%  
 mutate(sgi = sgi(level))

## # A tibble: 480 x 4  
## # Groups: month [12]  
## time level month sgi  
## <date> <dbl> <ord> <dbl>  
## 1 1970-01-01 74.8 January 0.343   
## 2 1970-02-01 73.5 February -0.0918  
## 3 1970-03-01 74.5 March -0.279   
## 4 1970-04-01 76.1 April -0.408   
## 5 1970-05-01 77.1 May -0.279   
## 6 1970-06-01 77.6 June -0.153   
## 7 1970-07-01 77.3 July -0.0918  
## 8 1970-08-01 76.1 August -0.153   
## 9 1970-09-01 74.4 September -0.279   
## 10 1970-10-01 72.8 October -0.343   
## # … with 470 more rows