Environmental Earth Sciences (2018) 77:310
https://doi.org/10.1007/s12665-018-7469-4

THEMATIC ISSUE

Analysis of hydrological data with correlation matrices: technical implementation and possible applications

Johannes Christoph Haas[1,2] · Matthew Switanek[3] · Steffen Birk[1,2]

Received: 13 October 2017 / Accepted: 31 March 2018 / Published online: 19 April 2018

Abstract
Changing political frameworks in addition to novel and more cost-effective means to investigate the subsurface have led to an increase in the availability of hydrological data. This wealth of data, however, poses new challenges in effectively making use of it. Traditional tools such as spreadsheets or proprietary datalogger software often do not scale easily with a larger amount of available datasets, requiring considerable user interaction. Also, comparing different locations and types of data can be difficult and tedious. Thus, a python script is presented that enables the user to quickly visualize and compare different types of data such as for example groundwater levels or precipitation amounts. This is done by first standardizing the data using different drought indices and, subsequently, visualization of correlation matrices or plots of data on maps. This approach can be used for data quality control (identifying erroneous data, classifying data into different types), data comparison (comparing different types of data, such as groundwater and precipitation; comparing different locations) and to visualize and analyze the development of hydrological data and their correlation patterns over time. Prospects and limitations of the approach are illustrated and discussed using various example applications.

Keywords Time series · Correlation matrices · Visualization · Standardization · Data analysis

Introduction

With the advancement of hydrological monitoring tools and information technologies, more comprehensive hydrological datasets are becoming increasingly available. For precipitation and temperature, which predominantly govern surface hydrological fluxes, datasets are improving in quality and spatial coverage. Subsurface hydrology has historically relied on classic techniques like auger drilling or various other rotary methods (see for example Langguth and Voigt 2004, chapter 10; DVGW 2008; Todd and Mays 2005, chapter 5 or Delleur 2007, chapter 11 for a general overview). However, more novel methods like direct push or sonic drilling are also becoming more common and can provide a time and cost advantage, compared with the techniques mentioned above (see for example Maxwell and Hildebrand 1994 or Delleur 2007, chapter 35).

The increase in available methods to obtain, visualize and compare data as well as reduction in costs will assist research projects, site investigations and site remediation. Also, a push toward open data (whether the source is government, university or otherwise), as for example demanded by the European Union by the directive 2003/98/EC (EU 2003) or stemming from the Freedom of Information Act (FOIA 1967) in the USA will facilitate greater access and use. Platforms like ehyd.gv.at (BML-FUW 2016) used for most of the examples herein, the ZAMG HISTALP dataset (http://zamg.ac.at/histalp/, Auer et al. 2007), gkd.bayern.de (LfU 2017), waterdata.usgs.gov/nwis/nwis (USGS 2017) or bafg.de/GRDC (GRDC

This article is part of a Topical Collection in Environmental Earth Sciences on "NovCare - Novel Methods for Subsurface Characterization and Monitoring: From Theory to Practice", guest edited by Uta Sauer and Peter Dietrich.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12665-018-7469-4) contains supplementary material, which is available to authorized users.

✉ Johannes Christoph Haas
johannes.haas@uni-graz.at

[1] Institute of Earth Sciences, NAWI Graz Geocenter, University of Graz, Graz, Austria

[2] FWF-DK Climate Change, University of Graz, Graz, Austria

[3] Wegener Center for Climate and Global Change, University of Graz, Graz, Austria

2017) or various other such platforms allow one to easily access or request a wealth of country- or state-wide data.

One important challenge related to the interpretation of this wealth of data is the need to make different types of data as well as data from different locations comparable. For many researchers, appropriate tools to more easily and quickly understand and visualize the data may not be readily available.

Many of the generally used software packages, such as the proprietary tools of the datalogger manufacturers or spreadsheet software, often do not scale very well to larger datasets and require considerable user interaction. Also, these tools often make it more difficult to include relevant spatial (e.g., latitude, longitude) information and elevation heights. GIS software is an obvious fix for this issue, but the use of it requires considerable training and can incur considerable costs. More importantly, comparing data from different locations or different types of data remains a challenging issue with any of the aforementioned tools. For example, groundwater levels measured at different elevations and/or in different hydrogeological settings are only seldom easily comparable in one plot, as are time series of precipitation and groundwater levels with their different scales.

In this paper, we provide an approach that allows a user to standardize and visualize large amounts of groundwater monitoring time series in conjunction with river stages, precipitation or other hydrometeorological time series data. This approach allows one to gain a quick overview of the data at hand and to identify correlations, different types of data and outliers. Additionally, it can also preserve some spatial information of the data at hand and requires minimal user interaction.

Various methods for standardizing the data such as the Standardized Precipitation Index—SPI (McKee et al. 1993), Standardized Groundwater Index—SGI (Bloomfield and Marchant 2013), Standardized River Stages Index—SRSI (Haas and Birk 2017) or the Standardized Precipitation Evapotranspiration Index—SPEI (Vicente-Serrano et al. 2010) allow the user to compare different time series at different locations. Haas and Birk (2017) demonstrate that this approach of standardizing data and visualizing it using correlation matrices can be useful for analyzing and comparing different types of hydrological data by conducting an analysis of the Austrian Mur valley, showing the effects of drought and flood conditions and the differences between different areas of the catchment.

This paper illustrates the methodological details of this approach using python code and discusses potential applications and extensions as well as limitations using examples from further use cases and regions.

Methods

Python is a popular coding language that is relatively easy to learn (see, e.g., Millman and Aivazis 2011; Pérez and Granger 2007; Oliphant 2007). It is increasingly used in scientific fields (see, e.g., Bakker (2014), Bakker et al. (2016) for hydrology) and is available for all major operating systems. The core language can be extended with the many available packages, ranging from use cases such as web frameworks to graphical user interfaces [see Python Software Foundation (2018) for a complete overview]. A multitude of those packages is suitable for scientific research, of which the popular numpy (van der Walt et al. 2011), pandas (McKinney 2010) and matplotlib (Hunter 2007) packages are used for the work presented herein. Python also has packages available that allow it to use or interact with other programs or languages, such as for example MATLAB (The MathWorks, Inc. 2017) or R (Gautier 2017), the latter being used herein to interface with the SPEI R-package (https://CRAN.R-project.org/package=SPEI, Beguería et al. 2014).

In the supplementary material we provide python code that can either be run as a standalone executable or can serve as the base for exploratory work in ipython (Pérez and Granger 2007). This python code is described in detail in the following subsections and an overview is given in Fig. 1.

Data access and preprocessing

For the examples shown herein, we obtained time series from the ehyd.gv.at platform, which provides the data of Austrian groundwater, surface water and precipitation measurement stations (BMLFUW 2016). A detailed description of this platform can be found in Haas and Birk (2017). As mentioned in the "Introduction" section, there are other possible sources for data and the tools provided herein can be adapted to other data sources or CSV files structured differently.

For the ehyd data, a module that reads in the CSV files is provided in the supplementary material. While the ehyd data are quality controlled (see, e.g., Godina 2000; Müller 2006; BMLFUW 2017), it is important to note that some of the time series do have gaps. Thus, we have build a simple error handling into the module.

Time series that are missing time information get discarded, as do time series that contain more than 10% of data flagged as erroneous in the original files and time series that contain more than three consecutive errors. In cases less severe than those, the module simply pads missing data with the previous water level or precipitation amount. We deem this an acceptable workaround for the

low number of affected time series [less than 10% of the data used in Haas and Birk (2017) for example], but for larger issues it is upon the user to assess a more appropriate way to deal with missing data.

As can be seen in the files provided in the supplementary material, there is considerable preprocessing, to get the files into a format usable within python due to some intricacies of the German language, such as the replacement of umlauts. Thus, we also provide a simple CSV reader that can serve as a base for the user to adapt the code to their data at hand. After this preprocessing, the header can be used to obtain information about the time series and to store it in a way that can later be queried and used.

Dataframe

The pandas package (McKinney 2010) is used to import and export the data. For storage and handling of the data, we use a pandas dataframe, which is "a 2-dimensional labeled data structure with columns of potentially different types" (pydata.org 2017) and thus similar to a spreadsheet. This similarity also carries over to their ability to hold multiple index levels and the ability to select from these indices.

Thanks to its similarity with spreadsheets, dataframes can be easily exported to and imported from spreadsheets, enabling easy exchange with other programs, such as the ones mentioned in the "Introduction" section. Thus, we export the standardized dataframe to a CSV file, to have a human- and machine- readable intermediate result that can serve as a basis for further work. Also, we export a Hierarchical Data Format—HDF/H5 file, which is a much smaller, binary file, very well documented and standardized (see hdfgroup.org 2017) and with bindings to multiple tools and languages, ranging from Fortran to MATLAB. Compared to the above mentioned CSV file, however, reading a HDF file requires more involvement for users unfamiliar with it.

This intermediate step of exporting a dataframe could also be forgone by combining the provided python modules into one, and thus keeping the dataframe in memory, but we prefer this twofold approach, since this allows for an easier change or substitution of the means used to build or analyze said dataframe.

Standardization

The comparison and interpretation of monitoring data from hydrometeorological gauging stations and observation wells situated in different regions need to account for the differences in climatological, hydrological or hydrogeological conditions. For the example of a weather station, the average precipitation for a particular month will be different from one location to another and from one month to another. It is thus of interest to observe if a particular month (or season) in one location is drier or wetter than normal. By standardizing the data, we can compare on the same scale two stations that may experience a monthly average of 100 and 200 mm of precipitation, respectively. For the example of a groundwater measurement well, the standardized data would allow us to compare on the same scale two wells that show an average groundwater level of 500 and 200 m asl, respectively. Therefore, comparisons can readily be made between both stations of the same type but at different locations. The standardization also allows comparisons between different types of stations, e.g., groundwater levels measured in meters and precipitation measured in millimeters on the same scale. However, it should be noted that this lack of real measurements can also be misleading. At first sight, one might assume that for example two stations with an SPI value of 0.5 do experience the same amount of precipitation, while the actual precipitation amounts observed could vary greatly.

To standardize the data, we use different variations on the same method corresponding to the different components of the water cycle. The precipitation data are standardized using the SPI (McKee et al. 1993), the groundwater levels are standardized using the SGI (Bloomfield and Marchant 2013), the river stages are standardized using the SRSI (Haas and Birk 2017, SGI applied to river stages) and the climatic water balance is standardized using the SPEI (Vicente-Serrano et al. 2010), in each case using the standardization procedure proposed in the original literature.

It has to be noted that there is some disagreement on the feasibility of the gamma distribution used for the SPI (e.g., Guttman 1999; Blain and Meschiatti 2015). In general, however, it is deemed as an acceptable distribution, especially regarding its use for European data (Stagge et al. 2015b). The authors are not aware of any discussions regarding the adequacy of the nonparametric normal scores transform used for the SGI and SRSI, which we assume to be due to the relative recentness of the indices and the "self fitting" quality of the used method. Regarding the SPEI, the used method defaults to the log-logistic distribution but according to Stagge et al. (2015b) the generalized extreme value distribution might be slightly preferable when assessing European data.

Other indices or standardizations such as the Palmer Drought Index—PDI (Palmer 1965), the Standardized Streamflow Index—SSI (Vicente-Serrano et al. 2012) or many of the indices listed in Svoboda and Fuchs (2016) can also be implemented.

In the supplementary material, we provide python modules for the SPI, the SGI and the SRSI and a python module to use the published R-implementation of the SPEI (https://CRAN.R-project.org/package=SPEI, Beguería et al. 2014) as well as a very simple and general

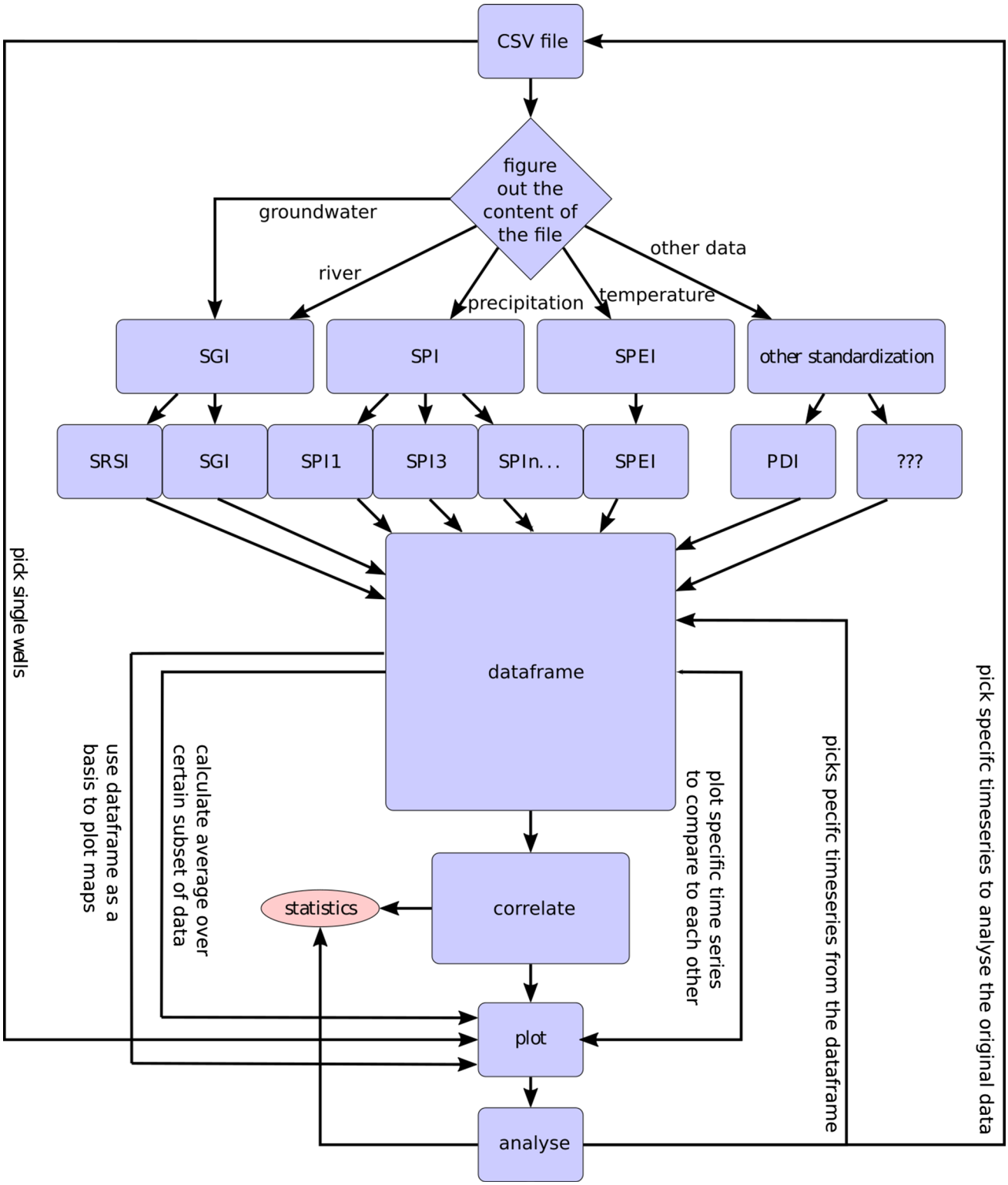

Fig. 1 Overview of the workflow described herein

standardization module using the z-scores approach, to quickly visualize all kinds of additional time series.

Matrix visualization

As previously mentioned in the "Standardization" section the standardization of data makes time series from different locations and/or different types comparable on the same scale. Additionally, it enables temporal averaging of surpluses (positive standardized values) or deficits (negative standardized values).

A main feature of the dataframe is the ability to easily construct a correlation matrix from it. This way, a Pearson correlation coefficient is obtained for every possible pairing of time series. However, since a dataset of n wells results in n^2 entries (half of which duplicates) this can be hard to readily interpret. Hence, we color code the correlation coefficients to enable a quick overview over what would otherwise be a difficult to read, huge text file or spreadsheet.

In contrast to single correlation coefficients and their map representations, as for example shown with state-wide SEDI9 versus SPI9 correlations in Kim and Rhee (2016), the correlation matrix visualization allows an easy, clutterless visualization of multiple groundwater wells, with the additional possibility to correlate and compare them with area or region wide indices such as SPI, SPEI and similar indices. Thus, we provide a python module in the supplementary material that plots these correlation matrices.

Map visualization

If a map visualization of data is needed, pythons matplotlib package provides the "basemap" module (see Hunter 2007 and https://matplotlib.org/basemap/) which enables the straightforward plotting of data on a variety of generic maps. Additionally, it is able to tie into ESRI's ArcGIS REST API, allowing for the use of ESRI maps. In the supplementary material, we provide a python module that can plot a dataset on a map.

Results and discussion

Quality control and data classification

One aspect only demonstrated tangentially by Haas and Birk (2017) is the ability to use the correlation matrix as a tool for data quality control and for the identification of one or several time series showing a behavior deviating from the others. Figure 2 shows the correlation matrix for the Aichfeld region in the Styrian Mur catchment adapted from Haas and Birk (2017), which is used to illustrate this aspect in more detail. As can be seen, most of the correlation matrix shows similar colors, indicating high correlations between the SGIs of the various time series, and thus a similar behavior of the underlying groundwater wells. However, there is also a distinct set of 5 wells that show low to negative correlations with everything else, but very high correlations with each other.

As discussed in Haas and Birk (2017), the average depth of the wells in the Aichfeld dataset is 13.5 m below ground level with a very high standard deviation of 8.5 m. Closer inspection of the underlying dataset reveals that this high standard deviation is caused by the five wells shown in Fig. 2, which have an average depth of ~ 25 m bgl, compared with an average of ~ 9.5 m bgl for the remaining 15 wells in the dataset. This difference in depth results in a different "behavior" of the groundwater and thus in low correlations of the different depth levels, suggesting that the two groups actually represent two distinct aquifers.

This different behavior is shown in Fig. 2b, where the average SGIs for the two sets of wells and the average for the complete dataset is shown. As can be seen the subset of deep wells does not affect the average too much, but this is only due to the limited number of deep wells in this example. For larger datasets, being able to identify differing datasets quickly could provide a valuable addition to more efficient quality control. Additionally, this option to calculate a representative average over a certain subset of time series enables other analysis options, such as trend analysis or a further way to compare different regions.

Haas and Birk (2017) further use the example of the construction of hydraulic structures in the Mur valley to demonstrate that the matrix visualization supports the identification of human impacts affecting the groundwater levels of some wells.

As also shown in Fig. 2, where artificial outliers (time series full of random data) have been included into the real data, those are clearly identifiable by their low, but not strongly negative, correlation with all of the real time series and with each other. "Semi outliers" such as the Danube gauging stations added into Fig. 2 seem to behave different again. While they do differ from most of the dataset, they still show moderate correlations with the river gauges in the area and some of the wells as well as noticeably lower correlations with the deep wells, making it hard to draw any distinct conclusions about those time series.

This demonstrates that correlation matrices allow one to identify time series—and thus mostly groundwater monitoring wells in the highlighted case—that deviate strongly from the average behavior in a region. However, this does not necessarily give an information about the reason for this deviation. Judging from our examples, estimates about the nature of the time series and thus the reason for the deviation in correlation can be made, however. Random time series appear to simply show a correlation coefficient

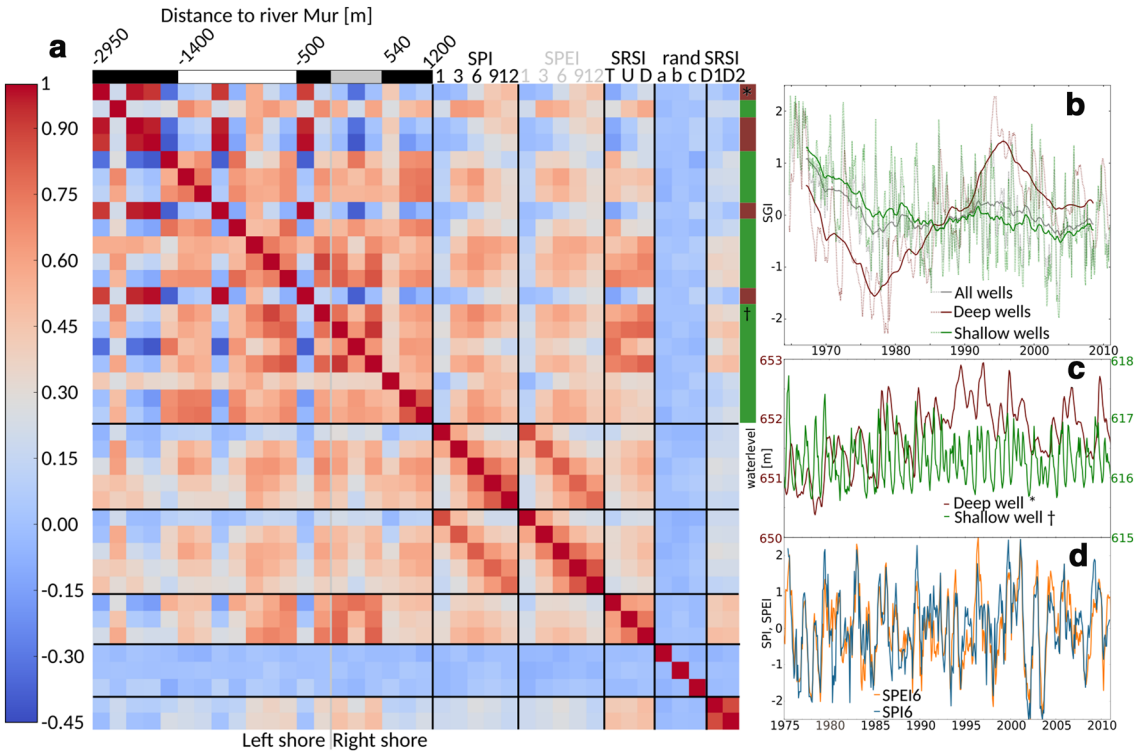

Fig. 2 Correlation matrix for the Aichfeld region, adapted from Haas and Birk (2017). Shown are the high correlations between most of the wells sorted by their distance to the river Mur, with a distinct set of deep wells lowly correlated with most other wells, but highly with each other (marked with a brown box at the right of the matrix; see also brown highlights in Fig. 5). **b** shows the average SGI for this subset of wells (deep wells, same brown color), the remaining wells (shallow wells, green color) as well as the total average for the region, with **c** showing the nonstandardized data for a representative well from the shallow and the deep wells. **d** shows the SPI6 and the SPEI6, which are highest correlated with the shallow wells. These two different indices show a high correlation and thus high similarity between the plots. Additionally, three time series (rand a–c) with random data uniformly distributed between − 3 and + 3 have been added to **a**, as well as two SRSI series from the river Danube (SRSI D1 and D2). As can be seen, the random data show correlations around 0 with everything, whereas the Danube still shows minor correlations with the rivers in the Aichfeld and some of the wells

close to zero, as expected due to their random nature. Time series of similar nature and location as the majority (e.g., the two Danube gauges inserted into Fig. 2) may still show minor correlation with many of the time series.

Comparison of different types of data

Standardizing data and plotting them in a correlation matrix allow for a comparison of different types of data, e.g., groundwater levels (SGI) and precipitation (SPI). As Haas and Birk (2017) have shown, generally shallow groundwater is highest correlated with the 6 month SPI and the most so in a foreland basin, whereas highest correlations of groundwater with surface water are seen in a narrow valley.

This comparison can be extended to other types of indices and other types of data as mentioned in the "Standardization" section. In Fig. 2, we show the addition of the Standardized Precipitation Evapotranspiration Index—SPEI (Vicente-Serrano et al. 2010). While the SPI is a purely precipitation-based drought index, the SPEI additionally includes a temperature-based estimate of potential evapotranspiration. One may expect that this allows a more appropriate representation of drought, particular under climatic conditions where a substantial part of precipitation is lost by evapotranspiration. Indeed, Kingston et al. (2015) found that a higher number of European droughts was identified by the SPEI than by the SPI. Likewise, Bachmair et al. (2015) found higher correlations for SPEI than for SPI. The findings by Bachmair et al. (2015), Stagge et al. (2015a) and Blauhut et al. (2016) similarly suggest that SPEI tends to be a better predictor of drought impacts than SPI, but a combination of different aggregation periods and indices generally is considered most favorable. In the given case, however, the behavior of the SPEI only minimally deviates from the SPI, as can be seen from the high correlation between the two indices (see Fig. 2d). As a consequence, the SPI-SGI and the SPEI-SGI correlations for the Aichfeld are also very similar, and the SPI even exhibits a tendency toward slightly higher correlations with SGI, which suggests that the impact of evapotranspiration is minor in this humid Alpine setting. This finding

agrees well with the results from a more comprehensive study for the Netherlands and Germany, where correlations between SGI and either SPI or SPEI were very similar (Van Loon et al. 2017). While the relationship between SPEI, SPI, and SGI thus deserves further investigation, particularly in more arid settings and on a regional level, as opposed to the country-wide assessment done in Stagge et al. (2015a), these examples demonstrate how the visualization of correlation matrices implemented in our Python code supports such comparisons of drought indices.

Spatial information

Correlation matrices can be used to visualize different regions in a catchment, which can then serve as a basis for comparison of the regions. Besides this one matrix per region approach, the sorting of the data in the correlation matrix can be used to convey some spatial information. The Haas and Birk (2017) paper mentioned throughout this work sorts the groundwater wells in their dataset by their distance to the main river in their regions, therefore showing the influence of the river on the groundwater. Other possible options would be for example sorting groundwater wells by their location along a stream (i.e., upstream to downstream), to sort surface water gauging stations by their location along the stream, or to sort groundwater wells by their depth or by the elevation of their locations. In the example shown in Fig. 3, the gauging stations are sorted by the size of their catchments, since this is provided as metadata in the ehyd dataset, and translates to their position along the stream. The other possible options mentioned above depend on the metadata provided with the dataset one uses, or have to be added by the user by hand. For this, two options are feasible: Adding a field with the desired data to the input data, or later adding an entry to the index of the large pandas dataframe described in the "Dataframe" section.

Also, contrary to the splitting between regions as shown by Haas and Birk (2017), regions can also be added together into single matrices (see Fig. 4), to show possible long distance correlations/teleconnections, or wells can be clustered together, for example into wells close to a know groundwater abstraction and wells deemed unaffected by human activities.

Time and event comparison

The dataframe can be split up into distinctive periods, so that a development over time can be visualized and the effects of extreme events such as the 2003 drought and the 2009 floods can be shown (see Fig. 4 and Haas and Birk 2017).

Building on this approach where time periods get selected due to outside information, we provide a short script in the

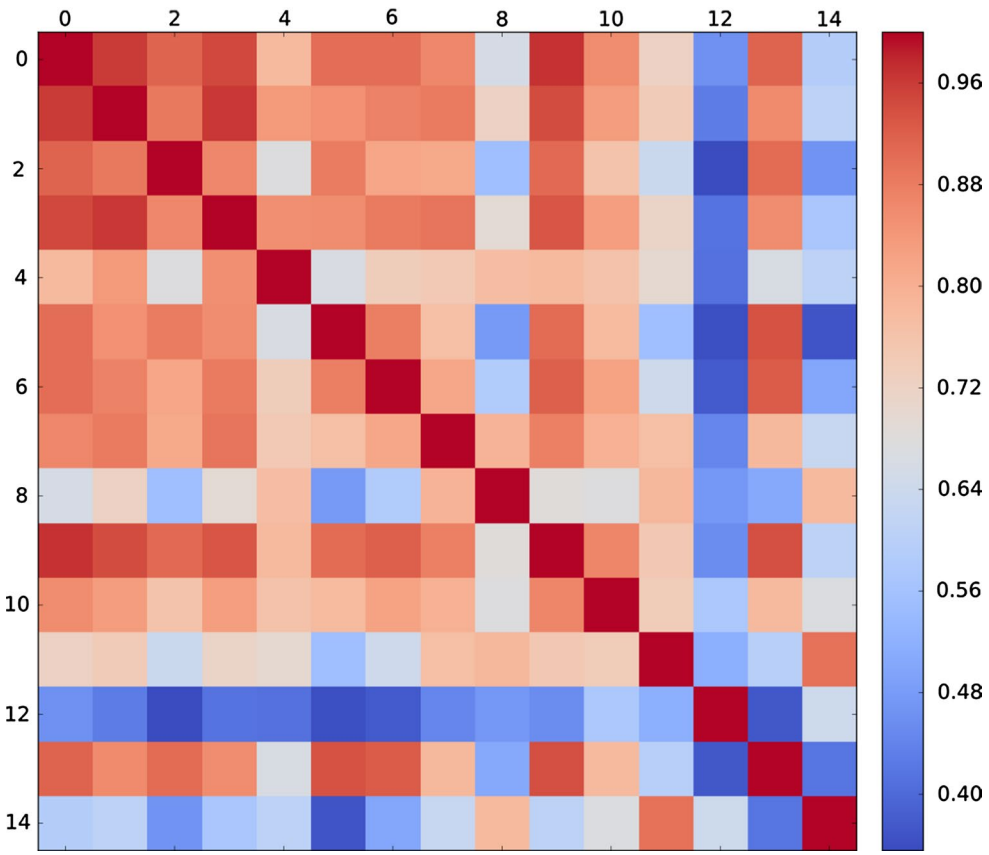

Fig. 3 Correlation matrix for of all the gauging stations at the Danube river in Austria with long-term data available. The stations are sorted from the most upstream one on the left (Achleiten) to the most downstream one on the right (Wolfsthal). Note, however, that we use a different color scale, compared with all the other plots in this paper

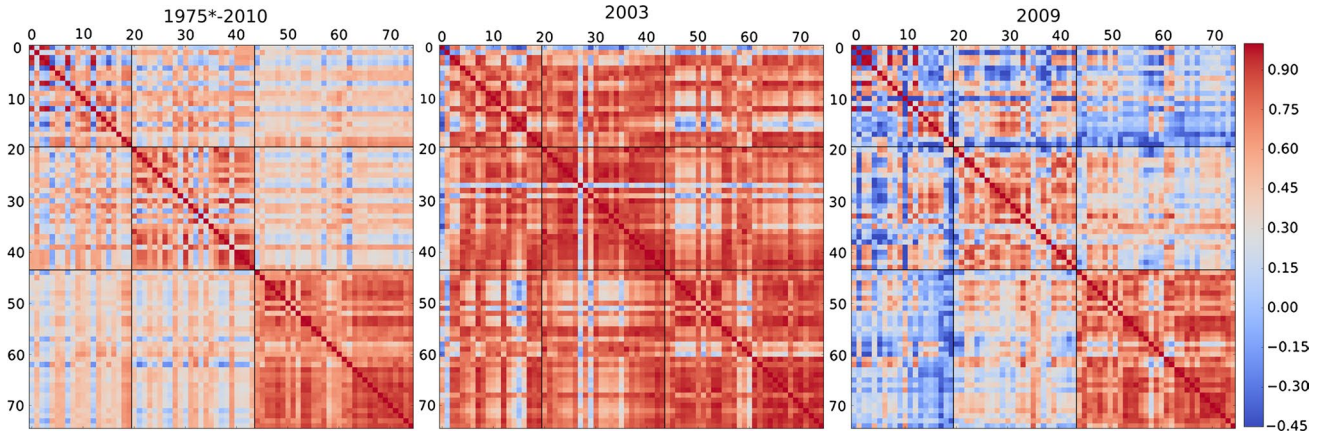

Fig. 4 Correlation matrix showing the groundwater measuring wells from all the subregions from Haas and Birk (2017) in one figure, demonstrating how the single subregions are clearly visible, but some areas are highly correlated, even though they are far away from each other. See Fig. 2 for a detailed discussion of one of the subregions. Also shown are all of the subregions for the drought year 2003 and the flood year 2009, showing the high correlations within the subregions under drought conditions as described in Haas and Birk (2017), but also high correlations between the subregions and the lower correlations within the subregions under flood conditions, as discussed in Haas and Birk (2017), but also lower correlations between the subregions, especially compared with the drought year. Note that one of the subregions only has data available from 1980 on

supplementary material that uses this ability to split a dataset into single years that are then visualized separately for every year in the data. This turns a long-term dataset into a series of yearly snapshots that can be easily browsed through with an image viewer or turned into a movie that allows for a quick way to check for changes in correlations or patterns. In this way, one can observe whether the correlation patterns are more or less stable or change through time. Identifying the causes of changes in correlation patterns, however, is not straightforward. This is illustrated by the example of the Mur catchment, where a snow-rich year shows a correlation pattern similar to a drought year and a snow-poor year behaves similar to a flood year (Haas and Birk 2017).

In addition, linking drought indices such as SPI, SPEI, or SGI to the actual occurrence of drought impacts is a challenging task. Bachmair et al. (2016) for instance, found that the accumulation periods of SPI and SPEI best linked to drought impacts vary from one region to the other and also depend on the kind of impact considered. In the Aichfeld example, the SPI6 exhibits the highest correlation with the SGI, suggesting that it is a reasonable indicator for drought impacts on groundwater, in particular pointing to low groundwater levels and thus low aquifer storage and discharge, which might be associated with other adverse environmental or economic effects. Using the threshold $SPI \leq -2$, as suggested by the original McKee et al. (1993) paper, results in six extreme droughts within the time period from 1975 to 2011. The correlation matrices of these events show a tendency to higher correlations between the SGI values of the different wells, as shown for the drought year 2003 by Haas and Birk (2017). In contrast, no distinct correlation patterns (e.g., low correlations as found by Haas and Birk (2017) for the flood year 2009) are obvious from the matrices for the six events classified as "extreme floods" (SPI $\geq +2$) when adapting the original drought classification by McKee et al. (1993) to floods. Remarkably, the SPI6 of the year 2009, which is associated with extreme floods according to the local literature (see for example BMLFUW 2011; Hornich 2009; Schatzl 2009; Stromberger et al. 2009; Ruch et al. 2010), stays below + 2, thus only matching the category of "severe flood".

These examples highlight the difficulties in interpreting the values of the drought indices, their correlation patterns and how to link them to real phenomena, such as the manifold possible drought impacts as for example discussed in Stahl et al. (2016). Thus, we suggest that the visual comparison and interpretation of correlation matrices should be complemented by independent, external information, such as reported drought or flood events, which can be used to select distinct time periods (e.g., years) that deserve closer investigation.

Another issue is the fact that very snow-rich years can produce similar patterns as the 2003 drought due to the snow cover cutting of the connection between precipitation and groundwater recharge. Likewise, very snow-poor years can produce patterns similar to the 2009 flood due to a direct connection of precipitation to groundwater recharge during the winter. However, this abundance or lack of snow does not show up in the SPI, or other such indices, thus making it impossible to tell a snow-rich year apart from a snow-poor one, using the matrix visualization alone. Related to this issue, it should be discussed what classifies as a drought or

flood for a region in question, taking into account not only a certain index value, but also other factors.

In a region such as Austria, a local meteorological drought might not necessarily result in a groundwater drought. This is attributed to the fact that groundwater can still be replenished by naturally occurring infiltration of river water. Groundwater and precipitation time series are often the only data that are easily available for past times, which can make it difficult to identify past droughts/floods as solely a function of one or both of these in the absence of a hydrological model. A proper classification would best be done by either assessing droughts/floods from a calibrated hydrological model, by using databases such as the European Drought Impact Inventory—EDII (see Stahl et al. 2016) or in the case of lacking information in the EDII by going through local records, such as newspaper archives, government reports or even "crowdsourced" data from social media.

Splitting the time series into distinct periods can also support the quality control and classification of the data. In the case of the SGI values discussed in the "Quality control and data classification" section, the negative correlations of the deep wells with the shallow wells and their high correlations with themselves is found to be amplified in many years, making them much more obvious than shown in Fig. 2. Also, most of the wells closest to the river Mur, which show high correlations in Fig. 2, keep these high correlations in most years, whereas the wells further away show a more arbitrary behavior through time.

Maps

Another option easily enabled by python is the relative ease of the plotting of maps for regions with the SGI values of various wells. This "classic" approach can be a first step in going beyond correlations and investigating their causations, especially when combined with the option to split the dataset into arbitrary time periods (see "Time and event comparison" section).

As discussed in Stromberger et al. (2009) the flood year of 2009 was characterized by multiple large floods and heavy precipitation events of often only local extent and interrupted by periods of "nice" weather. One would assume that such local phenomena are easy to spot in a map, possibly with the topography already giving helpful hints as to their causes and extent. However, these phenomena are not clearly visible in the monthly data available (see Fig. 5). Here the month of July 2009 (subfigure c) which follows the main floods of June (Stromberger et al. 2009) shows mostly only moderately wet conditions in groundwater with some wells even indicating moderately dry conditions. While this does fit the idea of locally differing conditions, the SGI values shown (0–approx. + 1) would indicate a "mild flood", using the definitions from McKee et al. (1993) for flood, are not fitting the reported severity of the floods, with the highest SGIs only seen in September 2009. In contrast, July of the drought year 2003 (subfigure b) does show very negative SGI values throughout the region shown in Fig. 5.

Regarding the possible identification of outliers, such as erroneous measurements or time series of differing nature (e.g., from different aquifers), and assessing their causation, a map view can provide both helpful hints and misleading marks, as demonstrated with the deep wells (brown highlights in Fig. 5) discussed in the "Quality control and data classification" section. These wells show a more muted behavior in the drought and flood years of 2003 and 2009, but can also behave very different, as highlighted for two cases in 1977 and 1984 (subfigure a). Thus, the map view can point toward the fact that this subset of wells behaves different. As to the causation of this difference, the first conclusion from the map view could be that this simply is due to their location on the northern bank of the river Mur. However, as discussed in the "Quality control and data classification" section, their differing behavior is presumed to be connected to their depth, instead of their location.

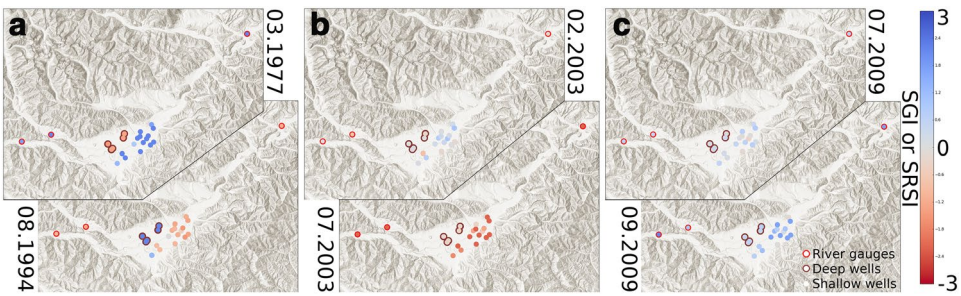

Fig. 5 Map representations for selected months of the drought year 2003 (**b**), the flood year 2009 (**c**) and two months with strongly deviating behavior of the deep wells (**a**). Note that the color scheme is the same as used in Figs. 2, 3 and 4 but unlike in those figures it does not represent correlation coefficients but rather color coded SGI and SRSI values. Deep wells are marked with a brown background. The three river gauges in the region are marked with a red background. Image width of the single maps: approx. 45 km

Conclusions

As shown in the previous sections, correlation matrices of standardized data enable one to quickly visualize complex systems in order to identify wells, stream gauges, precipitation measurements or other time series that correlate well with each other. The python implementation of this approach initially may need some involvement and adaption, but is expected to be more time efficient in the long-term than approaches using off-the-shelf software. The examples shown herein reveal opportunities and advantages for a variety of potential applications:

- The ability to quickly and easily—after the mentioned initial customization by the user—plot large amounts of data can speed up the first exploratory data analysis steps considerably.
- Following this, comparing different types of data can easily be enabled.
- Similarly, different regions can be compared.
- It can serve to classify data, i.e., by identifying time series belonging to one group that can be distinguished from another group, e.g., wells situated in a shallow aquifer as opposed to those of a deep aquifer.
- Related, the identification of outliers resulting from measurement errors or unexpected/unknown human impacts such as temporal changes in pumping rates or the construction of hydraulic structures affecting the groundwater levels of some wells can serve as a tool for data quality control.
- Following this classification and quality control, representative averages can easily be calculated, to allow for further analysis.
- Adding a temporal element, changes in correlations or correlation patterns over time can help to gain new insights into data or regions. Automated plotting of regional maps can be a first step to go beyond correlations.

While the proposed methodological approach does not resolve some fundamental issues related to data standardization and analysis, such as the suitability of a chosen index or the question whether or not an observed correlation indicates a causal relationship, it offers a valuable addition to the data analysis toolbox. The standardization procedures currently implemented in the python script can readily be changed to use other distributions or expanded by other indices. This may include existing indices such as the Palmer Drought Index or the Standardized Streamflow Index, but also newly developed standardization procedures. The possibility to integrate various indices into one data analysis and visualization tool thus may foster the future development of new, useful indices, e.g., addressing human impacts on water resources, such as water abstraction or contamination. In particular, the quality control and data classification aspect (see "Quality control and data classification" section) combined with the time aspect (see "Time and event comparison" section) offer the user a novel way to gain a greater understanding of the data at hand.

Acknowledgements Open access funding provided by Austrian Science Fund (FWF). This work was funded by the Austrian Science Fund (FWF) under research Grant W 1256-G15 (Doctoral Programme Climate Change—Uncertainties, Thresholds and Coping Strategies). Background maps for Fig. 5: ESRI World shaded relief.

References

Auer I, Böhm R, Jurkovic A, Lipa W, Orlik A, Potzmann R, Schöner W, Ungersböck M, Matulla C, Briffa K, Jones P, Efthymiadis D, Brunetti M, Nanni T, Maugeri M, Mercalli L, Mestre O, Moisselin JM, Begert M, Müller-Westermeier G, Kveton V, Bochnicek O, Stastny P, Lapin M, Szalai S, Szentimrey T, Cegnar T, Dolinar M, Gajic-Capka M, Zaninovic K, Majstorovic Z, Nieplova E (2007) HISTALP-historical instrumental climatological surface time series of the Greater Alpine Region. Int J Climatol 27(1):17–46. https://doi.org/10.1002/joc.1377

Bachmair S, Kohn I, Stahl K (2015) Exploring the link between drought indicators and impacts. Nat Hazards Earth Syst Sci 15(6):1381–1397. https://doi.org/10.5194/nhess-15-1381-2015

Bachmair S, Svensson C, Hannaford J, Barker LJ, Stahl K (2016) A quantitative analysis to objectively appraise drought indicators and model drought impacts. Hydrol Earth Syst Sci 20(7):2589–2609. https://doi.org/10.5194/hess-20-2589-2016

Bakker M (2014) Python scripting: the return to programming. Groundwater 52(6):821–822. https://doi.org/10.1111/gwat.12269

Bakker M, Post V, Langevin CD, Hughes JD, White JT, Starn JJ, Fienen MN (2016) Scripting MODFLOW model development using Python and FloPy. Groundwater 54(5):733–739. https://doi.org/10.1111/gwat.12413

Beguería S, Vicente-Serrano SM, Reig F, Latorre B (2014) Standardized precipitation evapotranspiration index (SPEI) revisited: parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. Int J Climatol 34(10):3001–3023. https://doi.org/10.1002/joc.3887

Blain GC, Meschiatti MC (2015) Inadequacy of the gamma distribution to calculate the standardized precipitation index. Revista Brasileira de Engenharia Agrícola e Ambiental 19(12):1129–1135. https://doi.org/10.1590/1807-1929/agriambi.v19n12p1129-1135

Blauhut V, Stahl K, Stagge JH, Tallaksen LM, De Stefano L, Vogt J (2016) Estimating drought risk across Europe from reported drought impacts, drought indices, and vulnerability factors. Hydrol Earth Syst Sci 20(7):2779–2800. https://doi.org/10.5194/hess-20-2779-2016

Bloomfield JP, Marchant BP (2013) Analysis of groundwater drought building on the standardised precipitation index approach. Hydrol

Earth Syst Sci 17(12):4769–4787. https://doi.org/10.5194/hess-17-4769-2013
BMLFUW, Abteilung VII 3 - Wasserhaushalt (HZB) im Bundesministerium für Land- und Forstwirtschaft, Umwelt und Wasserwirtschaft (2011) Hydrographisches Jahrbuch von Österreich 2009, vol 117. lebensministerium.at. https://www.bmlfuw.gv.at/dam/jcr:18fd2126-60a6-4a5e-813a-5f8f9c647f3e/JB2009.pdf. Last accessed 03 May 2017
BMLFUW, Bundesministerium für Land- und Forstwirtschaft, Umwelt und Wasserwirtschaft, Abteilung IV/4 - Wasserhaushalt (HZB) (2016) eHYD - Der Zugang zu hydrografischen Daten Österreichs. http://ehyd.gv.at/. Accessed 29 July 2016
BMLFUW, Bundesministerium für Land- und Forstwirtschaft, Umwelt und Wasserwirtschaft, Abteilung IV/4 - Wasserhaushalt (HZB) (2017) HyDaMS - das Hydrografische Datenmanagement System. https://www.bmlfuw.gv.at/wasser/wasser-oesterreich/wasserkreislauf/hydrographische_daten/HyDaMS.html. Accessed 18 Jan 2017
Delleur JW (2007) The handbook of groundwater engineering, 2nd edn. CRC Press, Boca Raton
DVGW, Deutsche Vereinigung des Gas- und Wasserfaches e V (2008) Arbeitsblatt W 115; Bohrungen zur Erkundung, Beobachtung und Gewinnung von Grundwasser. http://www.dvgw-regelwerk.de/plus/#technische-regel/dvgw-arbeitsblatt-w-115/70f55b. Accessed 11 Apr 2018
EU, European Parliament (2003) Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information. http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32003L0098. Accessed 11 Apr 2018
FOIA, Freedom of Information Act (1967) An Act to amend section 3 of the Administrative Procedure Act, chapter 324, of the Act of June 11, 1946 (60 Stat. 238), to clarify and protect the right of the public to information, and for other purposes
Gautier L (2017) rpy2—R in Python. https://rpy2.bitbucket.io/. Accessed 11 Apr 2018
Godina R (2000) Überblick über Daten und Datenarchive im Hydrographischen Dienst für Österreich. In: Gutknecht D, Blöschl G (eds) Niederschlag-Abfluss Modellierung—simulation und Prognose, Wiener Mitteilungen, vol 164, Gutknecht, Dieter, pp 119–128. http://www.hydro.tuwien.ac.at/forschung/publikationen/wiener-mitteilungen/wiener-mitteilungen-band-164/. Accessed 11 Apr 2018. ISBN 3-85234-055-1
GRDC—The Global Runoff Data Centre (2017). http://www.bafg.de/GRDC/. Accessed 11 Apr 2018
Guttman NB (1999) Accepting the standardized precipitation index: a calculation algorithm. JAWRA J Am Water Resour Assoc 35(2):311–322. https://doi.org/10.1111/j.1752-1688.1999.tb03592.x
Haas JC, Birk S (2017) Characterizing the spatiotemporal variability of groundwater levels of alluvial aquifers in different settings using drought indices. Hydrol Earth Syst Sci 21(5):2421–2448. https://doi.org/10.5194/hess-21-2421-2017
Hornich R (2009) Hochwasser und Hangrutschungen. Wasserland Steiermark 2:17–22. http://www.wasserwirtschaft.steiermark.at/cms/ziel/1356921/DE/. Accessed 11 Apr 2018
Hunter JD (2007) Matplotlib: a 2D graphics environment. Comput Sci Eng 9(3):90–95. https://doi.org/10.1109/MCSE.2007.55
Kim D, Rhee J (2016) A drought index based on actual evapotranspiration from the Bouchet hypothesis. Geophys Res Lett 43(19):10,277–10,285. https://doi.org/10.1002/2016GL070302
Kingston DG, Stagge JH, Tallaksen LM, Hannah DM (2015) European-scale drought: understanding connections between atmospheric circulation and meteorological drought indices. J Climate 28(2):505–516. https://doi.org/10.1175/JCLI-D-14-00001.1
Langguth HR, Voigt R (2004) Hydrogeologische Methoden. Springer, Berlin
LfU, Bayerisches Landesamt für Umwelt (2017) Gewässerkundlicher Dienst Bayern. http://www.gkd.bayern.de/grundwasser/karten/index.php?thema=gkd&rubrik=grundwasser&produkt=gwo&gknr=0. Accessed 5 Sept 2017
Maxwell RS, Hildebrand LF (1994) Improved direct push technologies: quicker and cheaper on-site methods for sampling and monitoring. Remediat J 4(4):435–442. https://doi.org/10.1002/rem.3440040406
McKee TB, Doesken NJ, Kleist J (1993) The relationship of drought frequency and duration to time scales. In: Proceedings of the 8th conference on applied climatology. American Meteorological Society, Boston, MA, 17–22, pp 179–183
McKinney W (2010) Data structures for statistical computing in python. In: van der Walt S, Millman J (eds) Proceedings of the 9th Python in science conference, pp 51–56
Millman KJ, Aivazis M (2011) Python for scientists and engineers. Comput Sci Eng 13(2):9–12. https://doi.org/10.1109/MCSE.2011.36
Müller G (2006) Datenprüfung und Verfügbarkeit beim Hydrografischen Dienst in Österreich. In: Blöschl G, Godina R, Merz R (eds) Methoden der hydrologischen Regionalisierung, Wiener Mitteilungen, vol 197, Gutknecht, Dieter, pp 55–70. http://www.hydro.tuwien.ac.at/forschung/publikationen/wiener-mitteilungen/wiener-mitteilungen-band-197/. Accessed 11 Apr 2018. ISBN 3-85234-088-8
Oliphant TE (2007) Python for scientific computing. Comput Sci Eng 9(3):10–20. https://doi.org/10.1109/MCSE.2007.58
Palmer WC (1965) Meteorological drought, vol 30. US Department of Commerce, Weather Bureau, Washington, DC
pydata.org (2017) Intro to data structures, pandas documentation. http://pandas.pydata.org/pandas-docs/stable/dsintro.html. Accessed 11 Apr 2018
Pérez F, Granger BE (2007) IPython: a system for interactive scientific computing. Comput Sci Eng 9(3):21–29. https://doi.org/10.1109/MCSE.2007.53
Python Software Foundation (2018) PyPI—the Python package index. https://pypi.python.org/pypi. Last accessed 13 Jan 2018
Ruch C, Reszler C, Schatzl R (2010) Operational flood forecasts for the Mur and Enns catchment in Austria—experiences from the June 2009 double flood event. In: EGU general assembly conference abstracts, vol 12. p 3992. http://meetingorganizer.copernicus.org/EGU2010/EGU2010-3992.pdf. Accessed 11 Apr 2018
Schatzl R (2009) Bericht des Hydrografischen Dienstes. Wasserland Steiermark 2:13–16. http://www.wasserwirtschaft.steiermark.at/cms/ziel/1356921/DE/. Accessed 11 Apr 2018
Stagge JH, Kohn I, Tallaksen LM, Stahl K (2015a) Modeling drought impact occurrence based on meteorological drought indices in europe. J Hydrol 530:37–50. https://doi.org/10.1016/j.jhydrol.2015.09.039. http://www.sciencedirect.com/science/article/pii/S0022169415007222
Stagge JH, Tallaksen LM, Gudmundsson L, Van Loon AF, Stahl K (2015b) Candidate distributions for climatological drought indices (spi and spei). Int J Climatol 35(13):4027–4040. https://doi.org/10.1002/joc.4267
Stahl K, Kohn I, Blauhut V, Urquijo J, DeStefano L, Acácio V, Dias S, Stagge JH, Tallaksen LM, Kampragou E, Van Loon AF, Barker LJ, Melsen LA, Bifulco C, Musolino D, de Carli A, Massarutto A, Assimacopoulos D, Van Lanen HAJ (2016) Impacts of European drought events: insights from an international database of text-based reports. Nat Hazards Earth Syst Sci 16(3):801–819. https://doi.org/10.5194/nhess-16-801-2016
Stromberger B, Schatzl R, Greiner D (2009) Hydrologische Übersicht für das erste Halbjahr 2009. Wasserland Steiermark 2:7–11.

http://www.wasserwirtschaft.steiermark.at/cms/ziel/1356921/DE/. Accessed 11 Apr 2018

Svoboda M, Fuchs B (2016) Handbook of drought indicators and indices. Integrated drought management tools and guidelines series 2. World Meteorological Organization (WMO) and Global Water Partnership (GWP), Geneva

The HDF Group (2017) https://www.hdfgroup.org/. Accessed 11 Apr 2018

The MathWorks, Inc (2017) MATLAB engine API for Python. https://mathworks.com/help/matlab/matlab-engine-for-python.html. Accessed 11 Apr 2018

Todd DK, Mays LW (2005) Groundwater hydrology, 3rd edn. Wiley, New York

USGS, US Geological Survey (2017) USGS water data for USA. https://waterdata.usgs.gov/nwis/nwis. Accessed 11 Apr 2018

Van Loon AF, Kumar R, Mishra V (2017) Testing the use of standardised indices and GRACE satellite data to estimate the European 2015 groundwater drought in near-real time. Hydrol Earth Syst Sci 21(4):1947–1971. https://doi.org/10.5194/hess-21-1947-2017

Vicente-Serrano SM, Beguería S, López-Moreno JI (2010) A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. J Climate 23(7):1696–1718

Vicente-Serrano SM, López-Moreno JI, Beguería S, Lorenzo-Lacruz J, Azorin-Molina C, Morán-Tejeda E (2012) Accurate computation of a streamflow drought index. J Hydrol Eng 17(2):318–332. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000433

van der Walt S, Colbert SC, Varoquaux G (2011) The NumPy array: a structure for efficient numerical computation. Comput Sci Eng 13(2):22–30. https://doi.org/10.1109/MCSE.2011.37