Environmental Modelling and Software 136 (2021) 104949

Contents lists available at ScienceDirect

Environmental Modelling and Software

journal homepage: http://www.elsevier.com/locate/envsoft

A toolbox for visualizing trends in large-scale environmental data

Claudia von Brömssen [a,*], Staffan Betnér [a], Jens Fölster [b], Karin Eklöf [b]

[a] Department of Energy and Technology, Division of Applied Statistics and Mathematics, PO Box 7032, Swedish University of Agricultural Sciences, SE-750 07 Uppsala, Sweden

[b] Department of Aquatic Sciences and Assessment, Section for Geochemistry and Hydrology, PO Box 7050, Swedish University of Agricultural Sciences, SE-750 07 Uppsala, Sweden

ARTICLE INFO

Keywords:
Generalized additive models
Visualization of trends
Surface waters
Acidification
Chemical recovery

ABSTRACT

Generalized additive models are increasingly used to identify and describe environmental trends. A major advantage of these models, as compared to simpler statistical tools such as linear regression or Mann-Kendall tests, is that they provide estimates of prevailing levels and trend magnitudes at any given point in time instead of an overall measure. For multiple time series, this versatility has to be followed by flexible visualization methods that can summarize and visualize trend analysis results for many series simultaneously. Here, we propose several types of visualizations and illustrate the methods by showing trends in variables related to the recovery from acidification in Swedish riverine data over the period 1988–2017. By this, we show that generalized additive models, together with a small number of selected plots, can comprehensively illustrate prevailing trends and summarize complex information from multiple series.

1. Introduction

Environmental questions are complex by nature. The effects of environmental pressures are often observed on a regional scale, because they are driven by large-scale factors such as airborne deposition, land-use, or climatic variables. In addition, policy measures are introduced on a national or international level and thereby influence many monitoring series simultaneously. This requires statistical evaluations that can be applied on multiple series without a substantial modeling effort and thereby provide simple summaries that can be tabulated or visualized to reveal the essential information contained in the trend analysis.

In this paper, we investigate the high potential of generalized additive models for trend analysis by developing a unified framework to visualizing long-term nonlinear trends for multiple environmental series. The overall goal is to give a comprehensive overview when, where and to what extent levels of the studied variables have changed. This is done using a small number of selected visualizations, while retaining the possibility to extract single-site model results for in-depth understanding of the present data.

To quantify trend slopes and to identify when changes over time are significant or not linear regression might be applied. However, due to the complex and non-linear nature of environmental trends in long time series, regression methods assuming a linear increase or decrease are seldom used. Mann-Kendall tests (Mann, 1945; Hirsch and Slack, 1984) and Sen-Theil slopes (Theil, 1959) can quantify trends and trend magnitudes that are not linear as long as they are monotone (i.e. either only increasing or only decreasing), and hence these approaches are often chosen for environmental trend analysis focused on single and multiple series.

Results of Mann-Kendall tests on multiple series are often presented in tables or graphs. For example, Grimvall et al. (2014) presented trends in river nutrient concentrations by color coding tables of p-values. Stoddard et al. (1999) presented trends in acid anion concentrations and alkalinity in lakes and streams by using site-specific Seasonal Kendall tests, which were visualized with bar charts of estimated slopes for different regions in North America and Europe. Similarly, Monteith et al. (2007) used Mann-Kendall tests and Sen-Theil slopes to analyze dissolved organic carbon (DOC) concentrations in lakes and streams at a number of stations in Europe and North America, and those investigators presented values for individual sites on a map and summarized trend slope distributions as boxplots for six larger regions. Futter et al. (2014) also used Sen-Theil slopes to quantify trends in water quality for acid-sensitive lakes in Sweden, and the results of that study were reported as the median and range of slope magnitudes over 35 lakes separately for 16 variables.

Similar to linear regression, in Mann-Kendall tests the results

* Corresponding author.
E-mail address: Claudia.von.Bromssen@slu.se (C. von Brömssen).

https://doi.org/10.1016/j.envsoft.2020.104949
Accepted 8 December 2020
Available online 10 December 2020

obtained are usually summarized as single numerical values (i.e. p-values or slope estimates). Unfortunately, this reduces the amount of available information about the existing trends and, even worse, can be very dependent on the time period selected for trend analysis. For example, if a time series has an initial increasing trend followed by a decreasing trend, linear regression and Mann-Kendall approaches will lead to insignificant test results for the entire series, but will indicate a decreasing trend if the analysis is performed only for the later part of the series, leading to very different conclusions.

To be able to identify the structure of present trends more correctly, individual time series plots are often used to complement trend tests. For example, Futter et al. (2014) presented individual time series of lake water quality together with annual minimum, median, and maximum values over all sites. A more common approach is to present one plot per time series, often in combination with a smooth curve that visualizes the structure of the long-term trend (Curtis and Simpson, 2014; Erlandsson et al., 2008; Monteith et al., 2014). The estimation of a smooth trend is usually achieved using a generalized additive model (GAM, Hastie and Tibshirani, 1986; Wood, 2017) or by loess smooths (Cleveland, 1979).

Several attempts based on simple tables and plots have been made to compile large amounts of information produced by such smooth trends. Orr et al. (2015) analyzed temperature measurements at 231 stations by fitting smooth trends to each series and determining the difference between the start and the end of each estimated smooth trend. For these differences, 95% confidence intervals were empirically produced by sampling series from the posterior distribution for the coefficients; thereafter, point estimates of the differences together with the 95% confidence intervals were plotted for all stations on a geographical gradient. Polansky and Robbins (2013) used generalized additive models to estimate levels of fruiting probabilities of several tree species, and the predicted levels were illustrated in a plot presenting species as rows and time points as columns. Similarly, response surfaces can be used to represent results from smoothing models. For example, Wahlin and Grimvall (2008) computed trend surfaces of nutrients that were smoothed in time on a yearly basis and over river sites sorted by the mean level of flow-normalized concentrations to illustrate trends over a larger region. Also, Xie et al. (2018) produced surface plots using empirical mode decomposition for climatic variables. In general, it is common to visualize modeled levels of the response variable rather than trend slope estimates.

Several approaches to test and visualize trends at multiple stations are based on combining such series with a single trend test or trend estimate. Mann-Kendall tests can be combined using the approaches described by van Belle and Hughes (1984) or Loftis et al. (1991), as illustrated by Stoddard et al. (1999) in an evaluation of trends in acid deposition. Similarly, Renard et al. (2008) suggested use of regional averages of Mann-Kendal test statistics to detect regional consistency in hydrological trends over hydroclimatic regions. Considering generalized additive models, Pedersen et al. (2019) explored the use of hierarchical general additive models to analyze ecological data with different smooth functions for individual groups or sites, while pooling them towards a common shape.

Many of the above-mentioned attempts to achieve a common analysis of multiple series were made in the context of acidification or climate change, that is, in research areas in which pressures are expected to have an effect on a regional level rather than on single series. Considering acidification, emission-controlling legislations such as the Convention on Long-Range Transboundary Air Pollution have led to reductions in factors such as airborne sulphur (S) deposition over the past 30 years (Vuorenmaa et al., 2018). Measurements from Sweden show a pronounced gradient with decreasing acid deposition from south to north (Garmo et al., 2014; Monteith et al., 2014; Stoddard et al., 1999; Vuorenmaa et al., 2018) and are, therefore, especially interesting to use as illustration of visualization tools that show the benefit of extensive and detailed screening of trends for follow up of regional-scale environmental pressures.

In this paper, we suggest a toolbox of specific visualizations to summarize results obtained from trend analyses done by generalized additive models. Their general nature and accessibility as easy-to-use functions in open-source software allows a fast overview of environmental trends in multiple series, while containing detailed information about single series, which can be used for further in-depth analysis.

2. Statistical and visualization methods

2.1. Statistical methods

Generalized additive mixed models (GAMMs; Hastie and Tibshirani, 1986; Wood, 2017) facilitate modeling of environmental series without prior definition of the shape of the trend curve.

A general model for trend analysis for monthly data could be

$$y_i = \mu + f_1(time_i) + f_{2,cycl}(month_i) + \boldsymbol{\beta X} + \varepsilon_i \quad i = 1, \ldots, n$$

$$\varepsilon \tilde{} N(0,\ \Lambda_\theta)$$

where the trend component $f_1(time_i)$ is a smooth function in time, e.g. using a date variable, modeled as a thin plate spline, while the seasonal variation $f_{2,cycl}(month_i)$ is described by a cyclic cubic regression spline with an annual period based on the month the observation is made. The model can easily be adjusted to more frequent data replacing month with Julian day. For less frequent data the seasonal variation could be modeled using a seasonal indicator variable instead of a smooth.

$\boldsymbol{\beta X}$ denotes the influence of one or several explanatory variables estimated by a parametric function. The explanatory variables can be either continuous (e.g., run-off or temperature) or categorical (e.g., describing analytical method changes or introduction of new policy measures, von Brömssen et al., 2018). Explanatory variables can also be included using smooth functions if the relationship to the response variable cannot be described parametrically.

The error term in a time series model is assumed to be dependent in time, and thus it is necessary to include autocorrelation in the variance-covariance matrix Λ_θ. The typical choice is a continuous autoregressive process of lag 1, assuming that correlations between two different points in time is a function of the time difference. It has been shown that inclusion of an autocorrelation term is important in generalized additive models, because otherwise there is a risk that the trend estimates will exhibit too much wiggliness, which might introduce spurious trends in a trend analysis, as has been demonstrated by Simpson (2018).

To be able to determine when interesting changes occur in time series, it is necessary to identify the uncertainty of the slope of the smooth trend curve at any time point. The computation of uncertainties for smooths has been described by Marra and Wood (2012) and Wood (2017). First derivatives of the smooth function can be computed by finite differencing, and their uncertainties can be deduced from the variance-covariance matrix of the original smooth (e.g., see documentation of the predict.gam function in the *mgcv* package; Wood, 2019). When using these uncertainty computations, a number of different confidence intervals can be constructed for smooth components, model predictions or trend smooth derivatives, and the most common types are component-wise, simultaneous, and pointwise intervals.

Simultaneous intervals (Ruppert et al., 2003) are determined by simulating representations of the whole estimated smooth function from its posterior distribution. The confidence bands determined are then required to cover (1- α)% of all simulated smooths. Component-wise intervals, which have been suggested by Wahba (1983) and Silverman (1984), cover (1- α)% of the smooth across the function. Pointwise confidence intervals represent uncertainty around the smooth at specific time points, and they are constructed by adding an error margin to the modeled smooth (Simpson, 2018; Wood, 2017). Pointwise intervals have been shown to have good component-wise coverage (Nychka, 1988), which means that their coverage is close to the nominal level

when averaged across the function, as long as the variance of the smoothed function is substantially larger than the bias. Marra and Wood (2012) have suggested an adjustment to improve the component-wise coverage in situations in which the variance is small. Here, we use confidence intervals for the derivatives of the trend smooth too identify important changes over time. The trend is significantly downward if the entire confidence band of the derivative lies below zero, but the trend is significantly upward if the entire confidence band lies above zero. This is also described by Simpson (2018).

Generalized additive models are fitted by the package mgcv (Wood, 2019). Derivative function and confidence intervals are computed using the package gratia (Simpson, 2019). In addition, we provide the function *screeningmodeling()* that fits a generalized additive model to each series in the dataset. Seasonal variation is estimated by default as a cyclic cubic regression spline. Other explanatory variables are not supported by the function at present. Autocorrelation in the error term is estimated by a continuous autoregressive process of order 1. This can be removed by the user (*autocor* = *FALSE*) and is removed automatically in case the model with autocorrelation does not converge. Both pointwise (*confidence*, default) and simultaneous (*simultaneous*) confidence intervals can be chosen. The function is described in greater detail in the supplementary.

2.2. Visualization methods

2.2.1. Visualization of single series

Single time series trends can be visualized by time series plots of observed data including the trend smooth. Using the estimated confidence intervals for the derivative of the trend smooth makes it possible to identify the periods during which the trend slope is significantly different from zero. In the plot, such periods can be illustrated with different colors (e.g., red for increasing trends and blue for decreasing), as has previously been done by Curtis and Simpson (2014) and Monteith et al. (2014) to elucidate trends. Single series plots can be produced from the output of the function *screeningmodeling()* by using the function *plot_individual_trend()*.

2.2.2. Trend screening plots

Single series plots are convenient only for a small number of different stations and variables. To be able to present a large number of series simultaneously, the information from the single series plots can be condensed series-wise, retaining the color information (i.e., the information on during which time periods the change in the estimated spline meets the requirement for significance at the 5% level). This color information for all sites is then presented in one plot, using site as row indicator and time as columns. While this is a common approach to illustrate mean levels at different sites (Wahlin and Grimvall, 2008; Xie et al., 2018) or deviations from an overall mean (Polansky and Robbins, 2013), it has, to our knowledge, not been used to visualize periods of change.

The function *plot_screeningtrends()* is applied to the output of *screeningmodeling()* to produce a trend plot with one row for each individual series that contains the color information of when significant changes have taken place. The order of the rows can be adjusted using the *sorting* statement to allow for different geographical or thematical orderings and the *wrapparvar* statement if different variabels or series should be presented in separate plots.

As an alternative, the function *plot_screeningtrends_pvalues()* presents the same model fits but illustrates significant periods by pointwise p-values. These p-values are computed under the null hypothesis that the derivative is zero. To visualize pointwise p-values, they are first transformed using $-log_{10}(p-value)$ to obtain values close to zero for high p-values. This means that low p-values are transformed into high positive values. To distinguish between positive and negative trends, the transformed p-values are multiplied by the sign of the derivative and color coded on a continuous scale from blue (decreasing) to red (increasing). Inasmuch as single very low p-values easily overshadow other trends in such plots, a limit of 0.00001 is set for the p-value (i.e., all p-values below that level are changed to 0.0001).

2.2.3. Visualization by proportion plots

Additional summarization of results can be useful to present comprehensive overviews. The function *plot_proportions()* shows, at any given time point, how many of the series exhibit increasing, decreasing, or non-significant trends at the 5% significance level. In other words, this function summarizes the information provided by *plot_screeningtrends()* over all series for each time point. A dashed line indicates what percentage of the series have observations at the given time point.

2.2.4. Visualizing the magnitude of trends

When trends are analyzed, the magnitude of any trend that is present over a longer period is especially interesting. Generalized additive models allow different trend magnitudes and directions during different periods, and hence it is not straightforward how to quantify an overall trend magnitude. Here, we choose to visualize the trend magnitude by comparing levels at any given time point with a reference level set to the mean of the first three years of the series. This gives an informative plot in series in which mean levels have changed markedly during the observation period. The proposed plot is created by the function *plot_screeningtrends_reference()*. In other situations, it might be more interesting to relate current levels to a reference level that, for example, describes pristine levels or target levels that should be reached. Such reference levels would be easy to implement but are not available at present.

Another visualization that can be interesting relates the magnitude of the trend at a specific time point to the estimated level of the smooth at the same time point. This approach can make it possible to identify situations in which a drastic relative change occurs, such as short episodes of unusually high or low values that are due to either natural conditions (e.g., snow melting periods) or data quality issues (e.g., instrumental errors or unaccounted discontinuities in the time series). These plots are created by the function *plot_screeningtrends_relative()*.

2.2.5. Visualization of local trends in multiple time series

When trends for multiple series are visualized using the methods proposed above, it is possible that similar patterns will be observed for series that are geographically close or belong to the same thematic group. Therefore, a subsequent step in the analysis can be to look closer at trends for a number of stations on a smaller regional scale or in a thematic group, or for a group of related variables at the same time. If the goal is to produce an overall trend estimate within this subset, hierarchical generalized additive models can be used as suggested by Pedersen et al. (2019), who applied a global trend function combined with series-specific deviations from that global function (denoted model GS). Visualization of such models can be done using the package gratia (Simpson, 2019). Here, we modify these plots by letting the global and series-specific trend smooths also include the intercept and thus represent the average level and trend, as well as the series-specific trends. This plot is not part of the suggested toolbox, but is useful for further in-depth analysis on a smaller regional scale.

2.3. Data

In this study, application of the presented tools for screening of environmental trends is exemplified by the well-recognized recovery from acidification since 1970–1980s (e.g., Stoddard, 1999). We used time series from 37 soft-water rivers for the period 1988–2017 selected from the Swedish national monitoring program (Fölster et al., 2014), in particular focusing on surface water pH and SO_4 concentrations. Considering that our aim is to explore statistical methods to visualize trends, not to understand the trends themselves, the dataset we use does not contain any potential drivers of variation other than seasonality. No

pre-processing of the data has been done. Stations are identified by their coordinates given in the in the SWEREF-system (N/S-E/W; SWEREF99TM EPSG:3006), where the first 7 digits number represent the latitudal gradient.

The water chemistry data used here are included in the Swedish national program for monitoring of freshwater quality. The chemical analyses were performed by the Department of Aquatic Sciences and Assessment at the Swedish University of Agricultural Sciences. All the analyses were carried out in the same laboratory over the entire study period (1988–2017), which has led to a high level of continuity. Historical flaws and changes in methodology are well documented. SO_4 was analyzed using an ion chromatography, and pH was assessed using a Metrohm 855 autosampler with a built-in pH meter in a thermostated flow cell. In the visualization of trends for multiple variables (3.7), in addition to pH and SO_4, the data we use include alkalinity, filtered absorbance, acid neutralizing capacity (ANC), dissolved base cation (BC), chloride, as well as nitrite and nitrate. All underlying chemical analyses followed standard methods, and the performing laboratory is certified for all methods applicable in the present investigation (https://www.slu.se/institutioner/vatten-miljo/laboratorier/vattenkemiska-laboratoriet/). The data set is available for download at Zenodo (von Brömssen et al., 2020).

3. Results

The following sections describe the analysis of trends in SO_4 and pH in Sweden to show how different approaches in generalized additive models and visualizations can be applied to evaluate the trends in detail. This is done by presentation of trend analysis results side-by-side for multiple series and several variables. Results are further condensed into proportion plots or by computing common trend estimates on a regional scale.

A first overview of the data is given in Fig. 1 indicating the location of the series over Sweden as well as their average levels in the beginning (1988–1990) and the end (2015–2017) of the available time series. Average values are given in Table A1 in the Supplementery.

3.1. Visualization of local trends in multiple time series

Fig. 2 illustrates single series plots with a trend smooth and color coding of significant periods of change in pH and SO_4 in the river Alsterälven (6585489-420879). Red sections reflect significant upward trends according to pointwise tests. Correspondingly, blue sections show significant downward trends, whereas yellow shows the fitted trend line when the trend is not significant. The curves indicate an initial increase in pH from 1988 to the mid 1990s (Fig. 2, left) and a clear downward trend in SO_4 over most of the time period (Fig. 2, right).

3.2. Trend screening plots

3.2.1. Screening of significant trend periods

The increase in pH and decrease in SO_4 apparent in the river Alsterälven (Fig. 2) are also common in many series throughout Sweden. Using trend screening plots, this information can be summarized in a combined figure retaining the data of significance at different time points for all individual sites, as shown in Fig. 3. It can be seen that the pH increased in the late 1980s and early 1990s in all series except a few in the northernmost part of Sweden, whereas it has remained rather constant since the late 1990s. The trends in SO_4 have been especially pronounced in south and mid Sweden, with a decrease starting in the mid 1990s and continuing until the mid or end of the 2000s. The series marked end-to-end in red or blue are those for which a linear trend was suggested as best model fit.

3.2.2. Trend screening using p-values

If the goal is instead to retain additional information about the level of significance for trends during different time periods, it can be appropriate to produce a plot like the one in Fig. 3 but colorized by obtained p-values, as illustrated in Fig. 4. Obviously, this slightly altered approach provides very similar information. The increase in pH during the first years in the series is shown in different shades of red, whereas the decrease in SO_4 is shown in stronger shades of blue, indicating the more drastic decrease that occurred in the late 1990s (i.e., the signal-to-noise ratio was higher for SO_4 than for pH). Again, this agrees with what is observed in the example shown for the Alsterälven (Fig. 2).

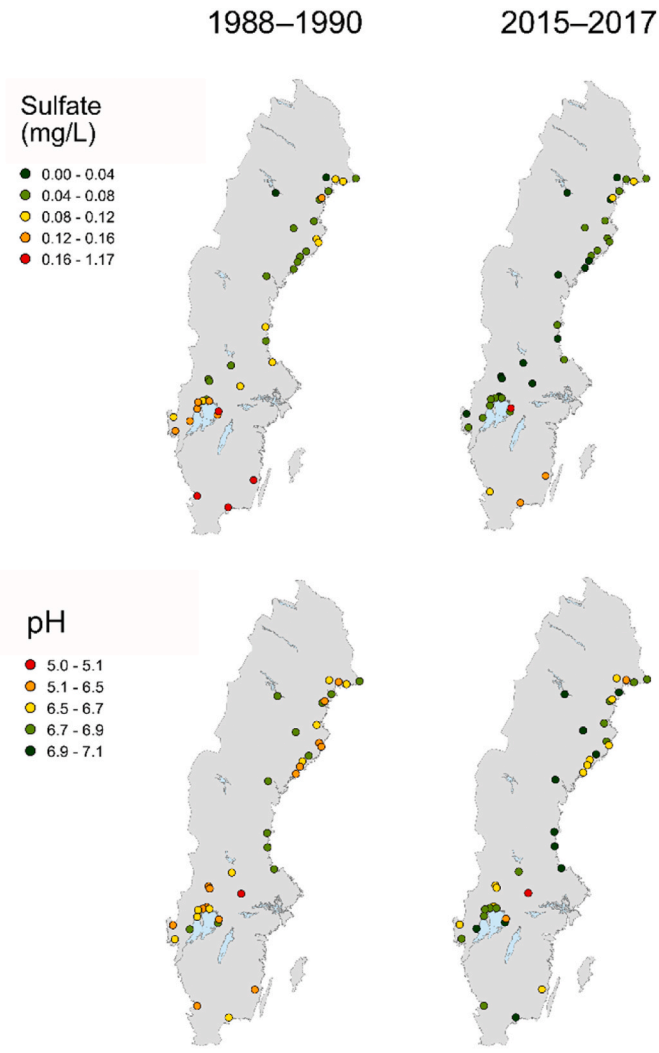

Fig. 1. Maps showing mean SO_4 and pH values in Sweden for the periods 1988–1990 and 2015–2017.

3.3. Visualization by proportion plots

By summarizing the results provided in Fig. 3, the proportion plot in Fig. 5 illustrates the percentage of stations showing upward, downward, and no significant trends at each time point. It is apparent that, in the early 1990s, about 75% of all stations in our data set showed an increase in pH and, equivalently, about 75% of the series showed a decrease in SO_4 in the late 1990s. Since 2010, significant changes have occurred in only a few series. After 2015, one and four series showed trends in pH and SO_4, respectively. Most of these series were fitted with a linear trend over the entire time range. As to whether a linear trend is indeed a good fit for these series (i.e., whether there was actually a continued increase or decrease) needs to be evaluated individually for each series.

Also in Fig. 5, a dashed line indicates what percentage of the series

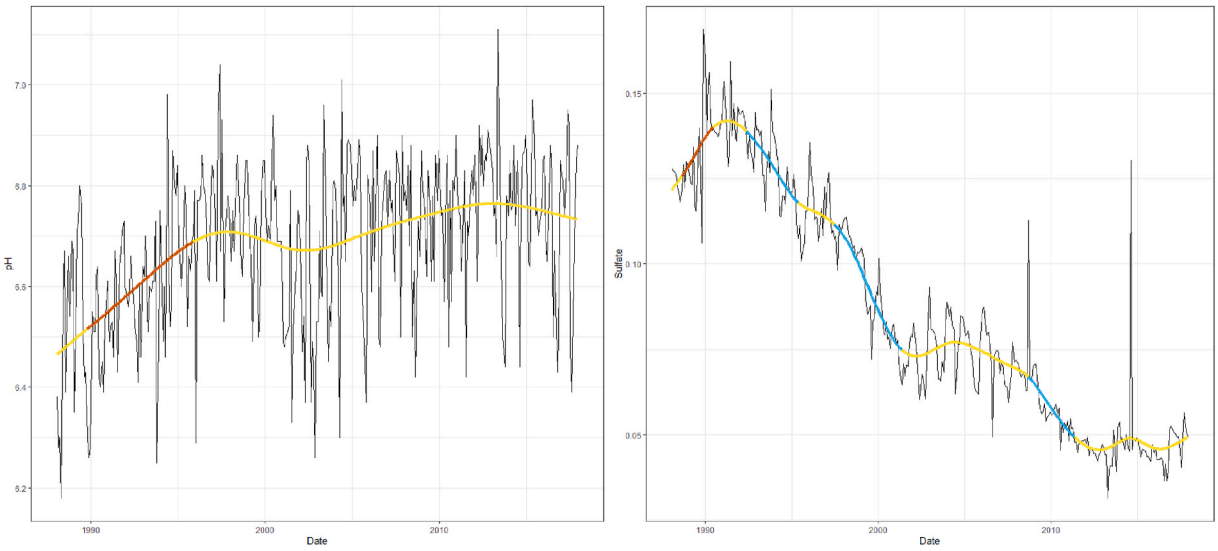

Fig. 2. Trend curves for pH and SO_4 in the river Alsterälven (6585489-420879) produced by a generalized additive model. Color indicates periods with significant trends, downward (blue) or upward (red), and no significant trend (yellow), respectively. The plots were created using the function *plot_individual_trend()*. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

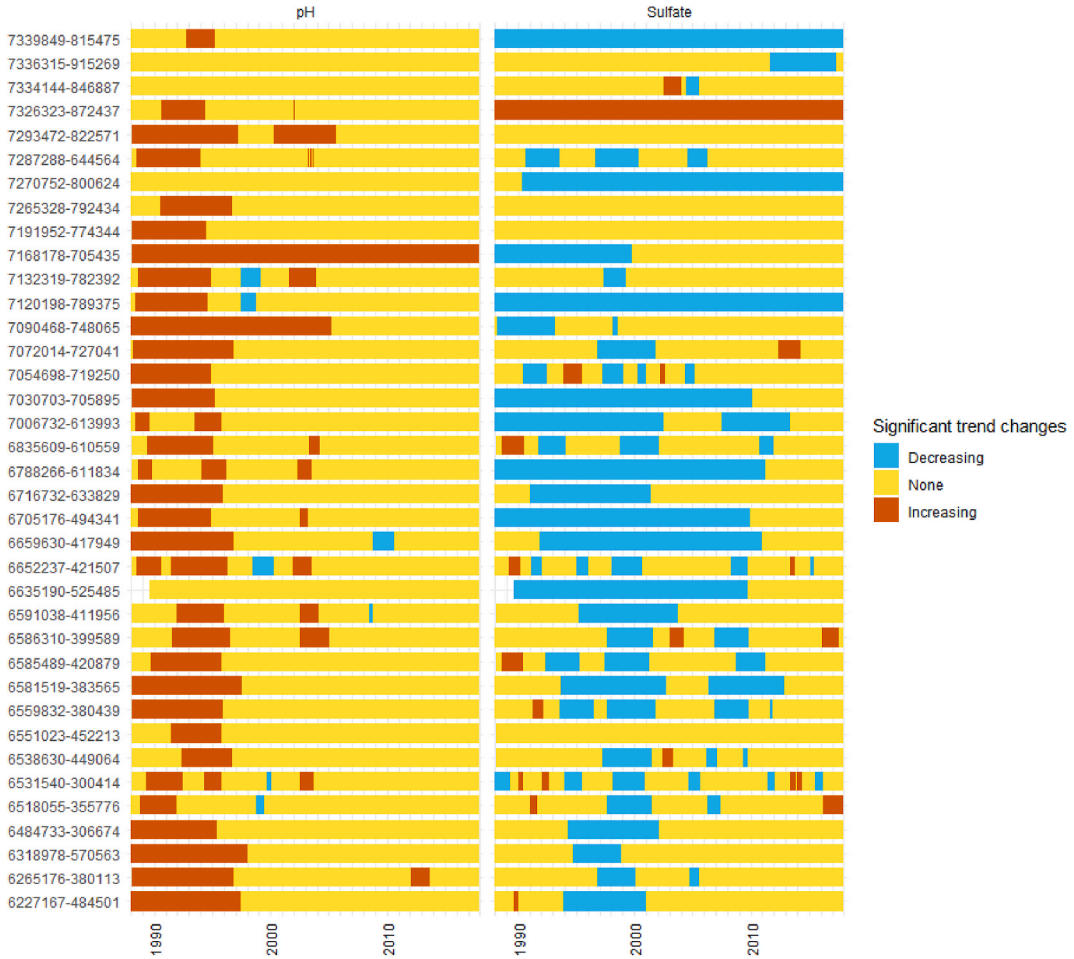

Fig. 3. Trend screening plot for pH and SO_4 in 37 Swedish rivers. Stations are denoted by their coordinates in the SWEREF-system (N/S-E/W) and are listed north to south according to the coordinates. The plot was created using the function *plot_screeningtrends()*.

have observations at the given time point. The data set used here was nearly complete, with only about 3% of the series (corresponding to only one series in our data) missing observations during the first one and a half years (Fig. 5, upper left corner).

Fig. 4. Trend screening plot for pH and SO_4 in 37 Swedish rivers visualized by p-values. Stations are denoted by their coordinates in the SWEREF-system (N/S-E/W) and are listed north to south according to the coordinates. The plot was created using the function *plot_trendscreening_pvalues()*.

3.4. Visualizing the magnitude of trends

The magnitude of the present trends is visualized as the level of the trend smooth at a specific time point divided by a reference level, which here is the mean of the first three observed years of each series (Fig. 6). The decrease in SO_4 was quite dramatic in several series in the south of Sweden, amounting to at least 50% compared to the levels occurring in 1988–1990. The increase in pH was less pronounced, with a maximum of about 7% compared to reference levels. However, this comparison is not informative about actual rises in pH, because the increase was already ongoing during this reference period (1988–1990).

Choosing the first three years of each series as reference level is reasonable in this context, as it is interesting to see how much pH and SO_4 have changed since the late 80, which was the end of a period with high deposition. However, to be able to compare changes between series it is necessary that all series are equally long. Observations at the station Laxbäcken (6635190-525485) were started in August 1989, which means that about one and a half years of observations are missing compared to the other series. Accordingly, the time period August 1989 to July 1991 is used as reference period for this station, and hence trend magnitudes computed at this site are not comparable with those at other sites. This is well illustrated for pH, because no noticeable change in pH level at any time point can be seen for Laxbäcken, whereas a clear increase is apparent for most other stations in the same area.

3.5. Visualization of different features by adjusting site order

In the illustrated plots, the sites are presented in order from north to south according to latitude. This approach enables the detection of trends on a geographical scale that can be associated with the gradient in acid deposition in Sweden. It can also be interesting to arrange sites in order according to average SO_4 or pH levels. In Fig. 7, we only present trend results for SO_4 sorted by average values for the reference years, placing sites with the lowest levels at the top of the plot. It can be seen that many series with high reference levels show strongly decreasing trends. However, there are also series with high levels that have no trends at all. For example, the station Visman Nybble (6551023-452213) has the highest SO_4 level of around 1 meq/l (Fig. 1, Table A1, Supplementary) but does not show a trend. This station is a local recipient of discharge from a paper mill and is only marginally affected by deposition, and therefore no changes due to decrease in airborne acidifying substances can be seen.

Also considering SO_4, the stations Alterälven Norrfjärden (7270752-800624) and Kalix älv Karlsborg (7326323-872437) show no trends and have relatively high levels during the reference period (0.158 and 0.103, respectively; Fig. 1, Table A1). Both of these stations are located in northern Sweden, which again is affected to only a limited degree by airborne deposition. It is plausible that the high levels of SO_4 in these rivers can be attributed to the presence of acid SO_4 soils in these areas (Becher et al., 2019).

3.6. Visualization of local trends in multiple time series

The results thus far (e.g., those shown in Fig. 3) suggest that there is considerable covariation in trends in southern and central Sweden. Therefore, it seems appropriate to take a closer look at trends on a

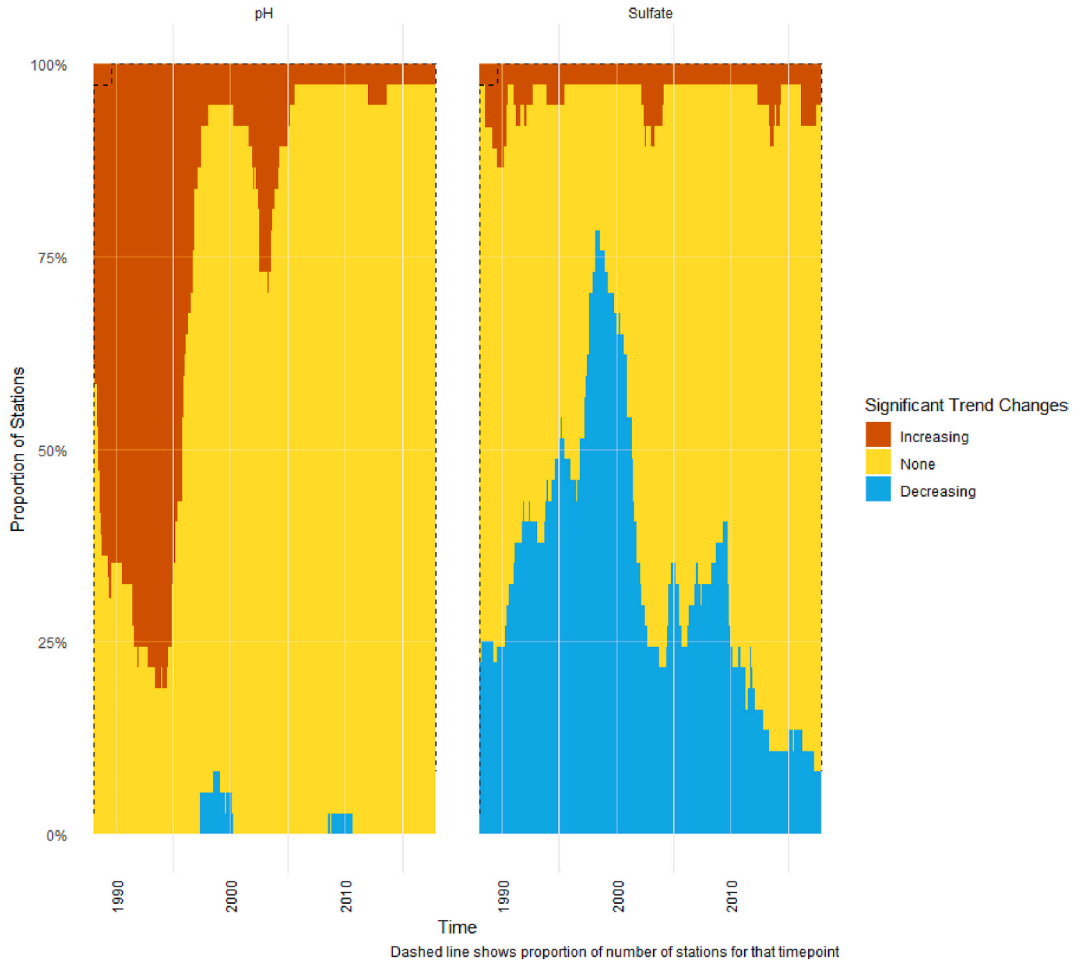

Fig. 5. Proportion plot illustrating the percentage of rivers with significantly increasing or decreasing trends in pH and SO_4 among all available stations. The percentage of stations observed for a specific time point is indicated by a dashed line. The plot was created using the function *plot_proportions()*.

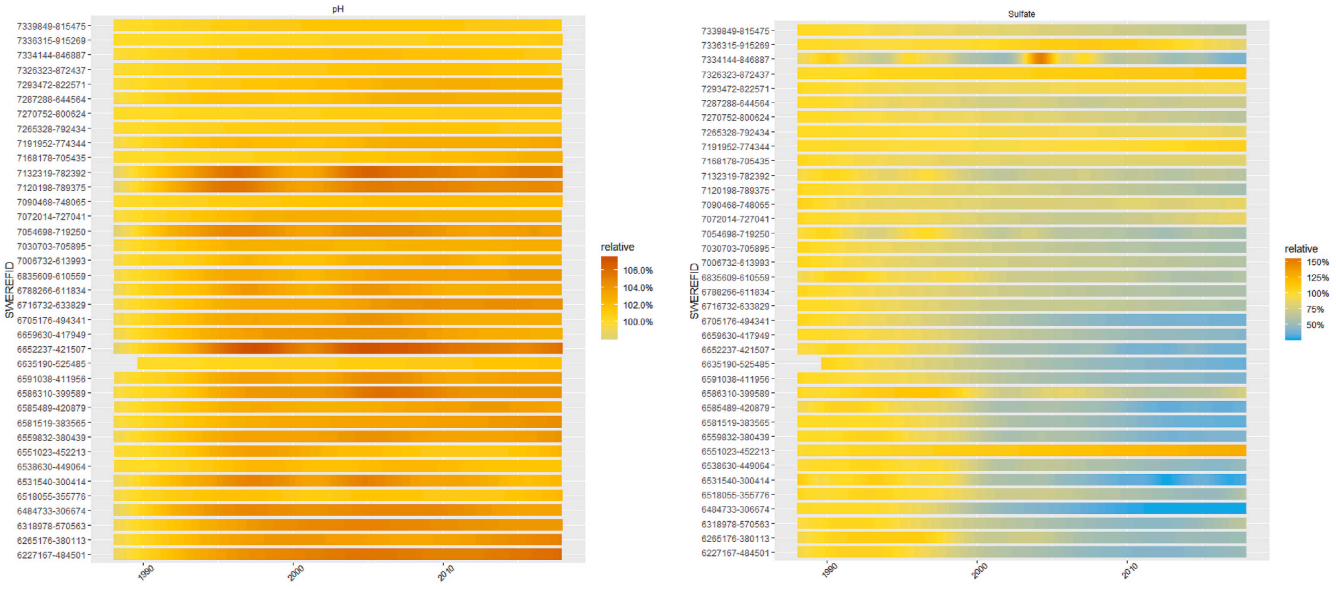

Fig. 6. Changes in pH (left) and SO_4 (right) in relation to levels observed during the reference period 1988–1990. Stations are denoted by their coordinates in the SWEREF-system (N/S-E/W) and listed north to south according to the coordinates. The plots were created using the function *plot_screeningtrends_reference()*. The behavior of river 7334144-846887 will be further discussed in section 3.8.

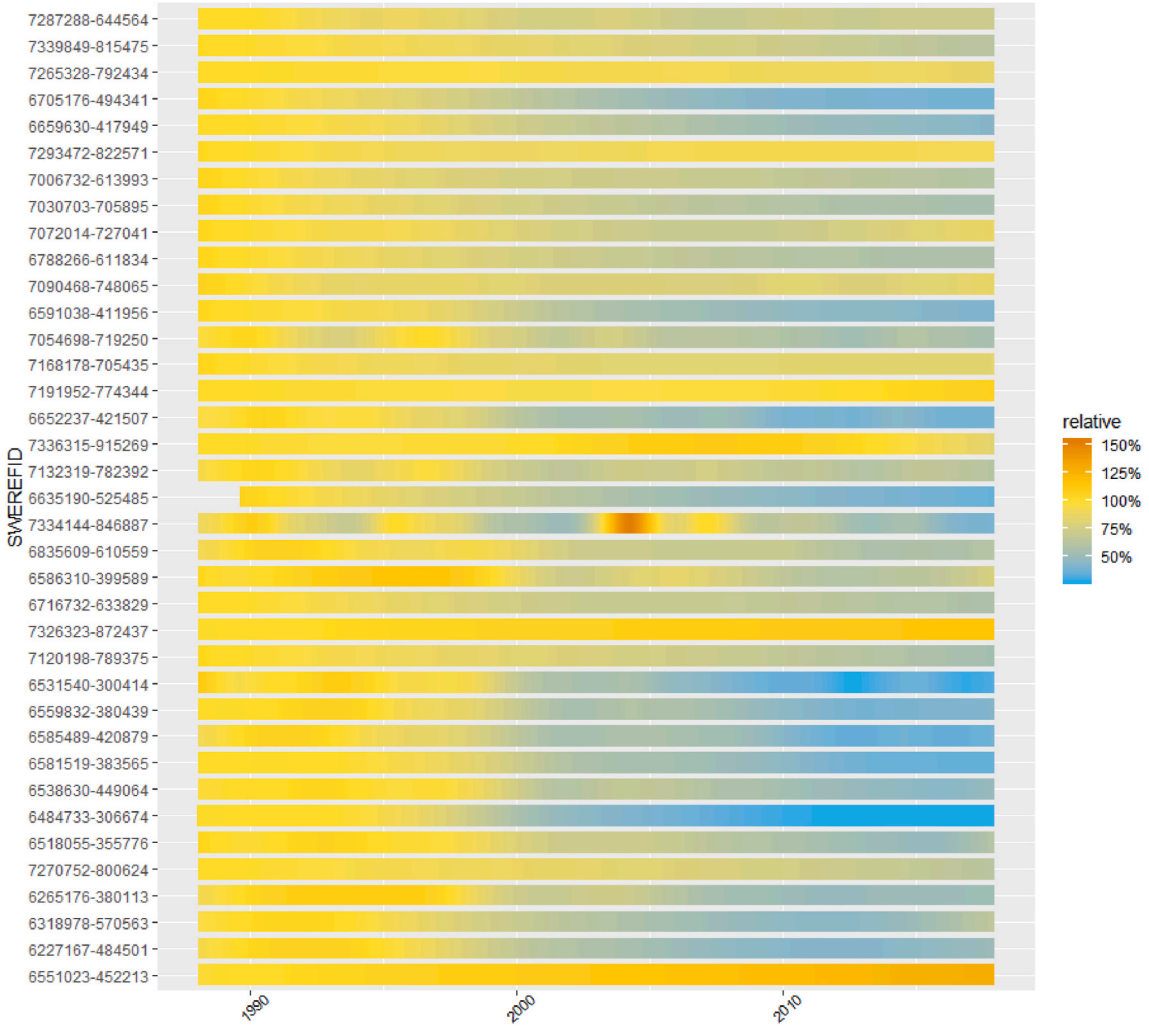

Fig. 7. Trends in SO_4 for 37 rivers arranged in order according to the reference level in 1988–1990, those with the highest levels at the bottom and those with the lowest levels at the top. This plot was created using the function *plot_screeningtrends_reference()* using the reference level per series in the sorting statement.

smaller sub-regional level. To achieve this, we select eight sites in Värmland County and fit a hierarchical generalized additive model including an overall trend function and site-specific deviations. We then plot the results including the overall intercept to represent the overall (Fig. 8, left) and the site-specific trends (Fig. 8, right). The eight series in this assessment exhibit a strong common trend. The general trend

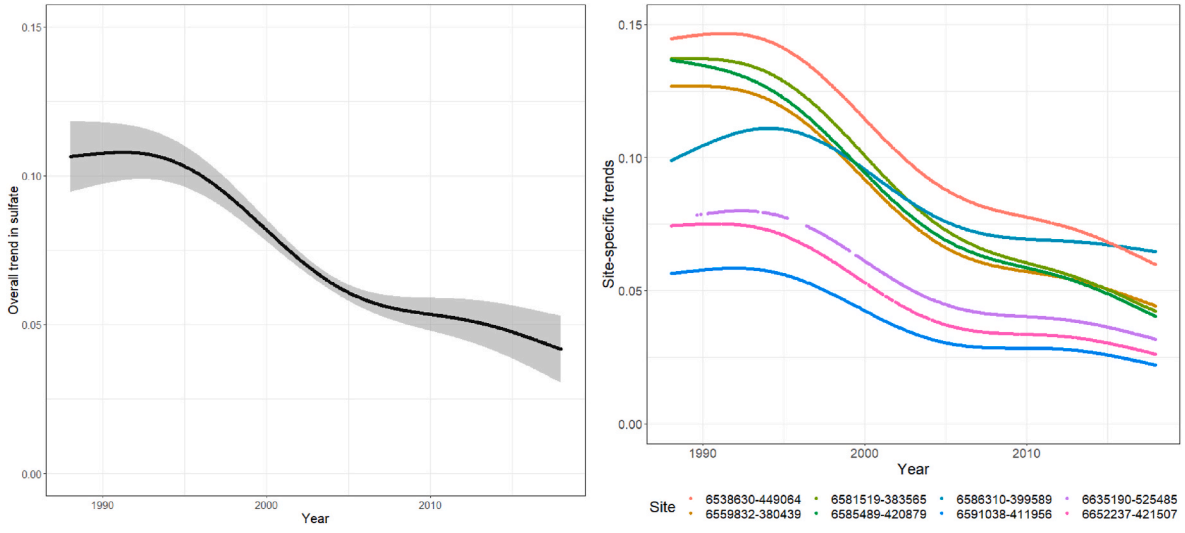

Fig. 8. The overall estimated trend for SO_4 in the Värmland County (left) and individual site-specific trends for the eight included sites (right).

estimate shows a mean decrease in SO_4 from an average of 0.107 meq/l to about 0.042 meq/l, which indicates an average magnitude of trend of about 60% in the region of Värmland over the analyzed 30 years. The absolute decrease was higher at sites with high initial levels, whereas the relative level of decrease was approximately 60% for all of the series. An exception to this is Norsälven Norsbron (6586310-399589), which started with an increase in SO_4 and thus showed a smaller decrease of only 40% over the analyzed time period.

3.7. Visualization of trends for multiple variables

Analyzing fewer series also makes it possible to take advantage of the wrapping and sorting function for the screening plots. This enables presentation of the trends for several chemical variables and several stations at the same time, as shown in Fig. 9 for the subset of series from Värmland. Even though this plot is quite "cluttered", some similarities in trends can be identified. For example, increasing trends early in the series can be seen for absorbance, alkalinity, ANC, BC, chloride, and pH at several stations, whereas mainly decreasing trends can be noted for SO_4 as well as nitrite and nitrate. A closer examination of the trend estimates' concurvity could give additional information about underlying processes.

3.8. Visualization to identify high-level episodes

In the plots presenting trend magnitudes (Figs. 6 and 7), it is apparent that one station, Töre älv (7334144-846887), deviates from general results by showing a sudden strong increase in SO_4 during a short period in the mid 2000s. For that station, several episodes of high SO_4 can be seen during the period of observation (Fig. 10, left), although only the longer episode in 2004 is seen in the trend screening plots, because other episodes, that were shorter and less pronounced, and are not and should not be picked up by the trend estimate.

To better visualize such periods of drastic change, we suggest an additional plot for which the estimated trend magnitude is divided by the level of the trend smooth at the same time point (Fig. 10). In contrast to Fig. 6, which identifies only one high SO_4 episode, the plot to the left in Fig. 10 shows several peaks for Töre älv (interchanging red and blue indicators), implying three to four episodes of higher SO_4 levels that can also be observed in the single series plot to the left in Fig. 10. None of the other stations in this analysis show similar signs of high SO_4 episodes. For those, once again, the plot to the right in Fig. 10 gives an indication of the sudden drop in SO_4 that occurred in southern and central Sweden in the late 1990s, which was dramatic compared to the prevailing levels.

3.9. Summary of trend analysis for pH and SO_4

The results of our assessments show that pH increased in the late 1980s and early 1990s in all series except for a few rivers in the north of Sweden. Since the late 1990s, the pH level has remained rather constant. Decreases in SO4 concentrations started in the mid 1990s in most cases and continued until the middle or end of the 2000s. These changes were observed in about 75% of all included series. Trends of increasing pH and decreasing SO_4 were stronger in the more acid-sensitive rivers in

Fig. 9. Trend screening plots for ten different water chemistry variables at eight stations in Värmland County. The plots were created using the function *plot_screeningtrends_pvalues()* using site ID as wrapper variable and variable name for sorting.

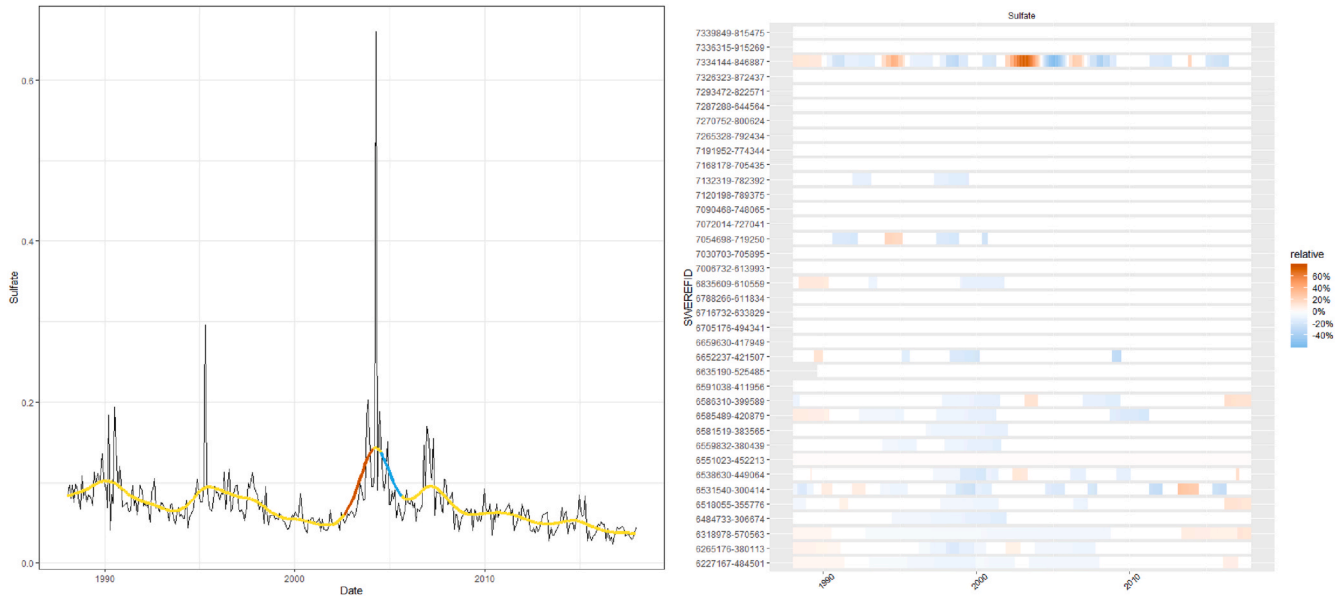

Fig. 10. SO_4 levels observed in Töre älv (left) and point-wise trends in relation to the average level at the same time (right). The plot on the left was created using the function *plot_individual_trend()*, and the plot on the right was created using the function *plot_screeningtrends_relative()*.

south and mid Sweden than in the rivers in the northern part of the country. At some stations the concentrations at the end of the observation period are at a magnitude of only 40% of the prevailing levels during the late 1980s. Episodes with high SO_4 levels were observed only in a single series in the river Töre älv, which has acid sulphur soils in the catchment. Runoff from an adjacent river indicated unusually dry conditions the year before the highest peak in Sulphur, which can lead to oxidized soils resulting in discharge of suphuric acid.

For a subset of series in the region of Värmland an additional analysis was done to illustrate individual and common trend magnitudes for SO_4, showing concordant trend curves and trend magnitudes that vary with initial SO_4 levels. A side by side presentation of trends for 10 different variables suggested association between several of these variables.

4. Discussion

Monteith et al. (2014) identified a marked decrease in SO_4 in lakes and rivers in the UK during the latter half of the 1990s by using first derivatives of trend smooths produced by generalized additive models. Such methods have a high potential for in-depth statistical analysis of complex trends for long time series. An issue that has not yet been satisfactorily addressed is the visualization of such long-term nonlinear trends for multiple series. Here, we propose a number of trend screening plots that allow a combined presentation of many trend analyses. The overall goal of the proposed functions and plots is to give a comprehensive overview of when, where, and to what extent levels of the studied environmental variables have changed. It is important that this information is provided in a small number of different plots. Using the proposed models and visualization methods addresses several of the problems with simple summaries produced by more generally used methods, such as the Mann-Kendall test and Theil-Sen slope.

First, results based on generalized additive modesl are not dependent on the time period of investigation, but are given point-wise in time. Mann-Kendall tests suffer from the subjectivity of defining specific time periods for the trend analysis, which can greatly influence the conclusions drawn. To determine if trends level out over time it is, furthermore, necessary to use distinct time periods as used by e.g. Garmo et al. (2014), who compared estimates of SO_4 trends in acidified lakes and rivers for the time period 1990–1999 with those for 1999–2008. Similarly, Vuorenmaa et al. (2018) computed mean annual changes in SO_4 in deposition and runoff in forested catchments in Europe for the periods 1990–2000 and 2001–2015. Using visualization plots for trend analysis based on generalized additive models we can determine in much greater detail when significant changes were observed and how long they persisted.

Even though the trend test results do not directly depend on the length of the series, some dependence remains, as the smoothing parameter in the generalized additive models is determined globally (i.e., the same amount of smoothing is used for the entire time series). In series with strong trends during limited time periods, adaptive smoothing can be used (Wood, 2017), although this makes additional demands on the data and can potentially lead to overfitting issues (Simpson, 2018) and has therefore not been investigated here.

A second advantage of generalized additive models and the ways we suggest to present their results is that they can be used with series of different lengths. Inasmuch as the model fit is data-driven, and the determination of significance of trends is made pointwise rather than for the entire series, the results are comparable between series, even if start and end times are not the same. For proportion plots, information about how many series contain missing values is given to allow interpretation. In our proposal, magnitude of trends is quantified compared to a reference period defined as the first three years of each series, respectively. Accordingly, this specific type of plot is vulnerable to series length and is difficult to interpret if start times differ between series. This is not a problem in our series, because they all started at approximately the same time point, and thus the plot provides a good overview of the total increase or decrease over time. For other applications, the magnitude of trends can be interpreted for single series or compared with reference values that are not dependent on the series length or start time (e.g., a target value). We also present a plot showing the magnitude of trend at a specific time point in relation to the prevailing level of the variable at the same time point, which can help identify short-term episodes of high values in the time series driven by factors like snow melting periods or data quality issues. Such plots are not dependent on the start and end time points of the analyzed series.

A third advantage of the proposed methods compared to more traditional approaches is that they allow a number of different ways to sort and group sites and variables to enhance visualization of various features. We use geographical ordering as well as sorting with respect to reference levels. Several other ways to sort series are conceivable. For

example, depending on the context, it can be possible to use mean values of drivers, such as water discharge or percentage of arable soil or forest.

The proposed plots present similar information to what could be obtained by plots of the absolute or standardised levels, but are also different in a very important aspect. They focus on the time periods of change rather than when high levels are observed, i.e. they focus on the process of ongoing pollution or recovery rather than the obtained status. This is an advantage when the level itself is not of interest, as is the case with our example data, where levels of pH and SO_4 are naturally varying over the geographical gradient of Sweden. Furthermore, we also identified one locally polluted station and stations with high natural level due to acid soils where the effect of changes in deposition could not be observed.

The formal test of a statistical hypothesis is not a priority in this context, because the focus is on exploring and finding covarying trends in larger areas or multiple chemical variables. Instead, the trend screening can be used to generate new hypotheses, especially when applied together with analysis of potential drivers. Such hypotheses might then be followed up on a local scale using sites that monitor these variables more intensively.

The suggested visualization methods are general and can be used for any types of smooth trends. It could, e.g, be interesting to illustrate trends caused by pressures or climatic changes that are more pronounced during certain seasons (e.g., driven by agricultural or forestry practices or snow cover and snow melting periods). For this, trend smooths can be computed for separate seasonal series and visualized by the proposed graphs. Similarly, in some situation it is not the main interest to model the trends in the mean of a variable, but how extremes, e.g. quantiles, change over time. A trend model for quantiles can be accomplished with generalized additive models and, again, the same visualizations can be used to summarize the results.

In this study, we have initiated a trend analysis for SO_4 and pH in 37 rivers in Sweden. As expected, our assessment indicates that most of the rivers, especially those located in the southern or central parts of the country, show strong decreases in river SO_4 and increases in pH as a consequence of reduced S deposition. The results obtained distinctly describe the steeper decrease in SO_4 in the 1990s compared to after the 2000s, as well as the increase in pH up to the mid 1990s that subsequently leveled out. The time-continuous feature of the trend tests used give a substantial advantage over previous investigations that used Mann-Kendall tests, where results were dependent on the selected time period, e.g. studies on changes in riverine SO_4 or pH in Sweden have been described by Löfgren et al. (2011, 2009), Skjelkvåle et al. (2007), and Fölster and Wilander (2002).

The use of Mann-Kendall test also makes it more complex to investigate how strong trends are during different time periods. Our proposed visualizations clearly show the differences in timing and magnitude of increases in pH and decreases in SO_4 related to geographical location, with a higher recovery rate in the more acid-sensitive rivers in the south of Sweden compared to the north. Furthermore, water chemistry variables can be affected by each other or by climatic conditions, which can be explored by using informative plots for multiple variables and, in the process, series properties such as data quality problems or high level episodes can easily be detected.

5. Conclusion

Generalized additive models provide a reliable basis for trend analysis of environmental data. These flexible models, in combination with a variety of plots to visualize their results, constitute powerful tools for analyzing many series and many variables at the same time. The present study shows that use of a small number of screening plots can comprehensively illustrate trends in SO_4 and pH in Sweden over the time period 1988–2017. A more in-depth analysis is needed to understand drivers and characteristics of the individual trends, but the approach proposed here can guide users in the right direction.

Software and data availability

The data set and function code are available at https://github.com/claudiavonbromssen/Trend-screening and the function versions used for this article can be downloaded at Zenodo (von Brömssen et al. (2020), https://doi.org/10.5281/zenodo.3935305. Functions are written in the freely available statistical computing software R (R Core Team, 2020).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envsoft.2020.104949.

References

Becher, M., Sohlenius, G., Öhrling, C., 2019. Sur Sulfatjord - Egenskaper Och Utbredning, vol. 37. SGU-Rapp.

Cleveland, W.S., 1979. Robust locally weighted regression and smoothing scatterplots. J. Am. Stat. Assoc. 74, 829–836. https://doi.org/10.1080/01621459.1979.10481038.

Curtis, C.J., Simpson, G.L., 2014. Trends in bulk deposition of acidity in the UK, 1988-2007, assessed using additive models. Ecol. Indicat. 37, 274–286. https://doi.org/10.1016/j.ecolind.2012.10.023.

Erlandsson, M., Buffam, I., Folster, J., Laudon, H., Temnerud, J., Weyhenmeyer, G.A., Bishop, K., 2008. Thirty-five years of synchrony in the organic matter concentrations of Swedish rivers explained by variation in flow and sulphate. Global Change Biol. 14, 1191–1198. https://doi.org/10.1111/j.1365-2486.2008.01551.x.

Fölster, J., Johnson, R.K., Futter, M.N., Wilander, A., 2014. The Swedish monitoring of surface waters: 50 years of adaptive monitoring. Ambio 43, 3–18. https://doi.org/10.1007/s13280-014-0558-z.

Fölster, J., Wilander, A., 2002. Recovery from acidification in Swedish forest streams. Environ. Pollut. 117, 379–389. https://doi.org/10.1016/S0269-7491(01)00201-9.

Futter, M.N., Valinia, S., Löfgren, S., Köhler, S.J., Fölster, J., 2014. Long-term trends in water chemistry of acid-sensitive Swedish lakes show slow recovery from historic acidification. Ambio 43, 77–90. https://doi.org/10.1007/s13280-014-0563-2.

Garmo, Ø.A., Skjelkvåle, B.L., de Wit, H.A., Colombo, L., Curtis, C., Fölster, J., Hoffmann, A., Hruška, J., Høgåsen, T., Jeffries, D.S., Keller, W.B., Krám, P., Majer, V., Monteith, D.T., Paterson, A.M., Rogora, M., Rzychon, D., Steingruber, S., Stoddard, J.L., Vuorenmaa, J., Worsztynowicz, A., 2014. Trends in surface water chemistry in acidified areas in Europe and North America from 1990 to 2008. Water. Air. Soil Pollut 225, 1880. https://doi.org/10.1007/s11270-014-1880-6.

Grimvall, A., von Brömssen, C., Lindstrom, G., 2014. Using process-based models to filter out natural variability in observed concentrations of nitrogen and phosphorus in river water. Environ. Monit. Assess. 186, 5135–5152. https://doi.org/10.1007/s10661-014-3765-y.

Hastie, T., Tibshirani, R., 1986. Generalized additive models. Stat. Sci. 1, 297–310. https://doi.org/10.1214/ss/1177013604.

Hirsch, R.M., Slack, J.R., 1984. A nonparametric trend test for seasonal data with serial dependence. Water Resour. Res. 20, 727–732. https://doi.org/10.1029/WR020i006p00727.

Löfgren, S., Aastrup, M., Bringmark, L., Hultberg, H., Lewin-Pihlblad, L., Lundin, L., Karlsson, G.P., Thunholm, B., 2011. Recovery of soil water, groundwater, and streamwater from acidification at the Swedish integrated monitoring catchments. Ambio 40, 836–856. https://doi.org/10.1007/s13280-011-0207-8.

Löfgren, S., Cory, N., Zetterberg, T., Larsson, P.-E., Kronnäs, V., 2009. The long-term effects of catchment liming and reduced sulphur deposition on forest soils and runoff chemistry in southwest Sweden. For. Ecol. Manag. 258, 567–578. https://doi.org/10.1016/j.foreco.2009.04.030.

Loftis, J.C., Taylor, C.H., Chapman, P.L., 1991. Multivariate tests for trend in water quality. Water Resour. Res. 27, 1419–1429. https://doi.org/10.1029/90WR01802.

Mann, H.B., 1945. Nonparametric tests against trend. Econometrica 13, 245–259. https://doi.org/10.2307/1907187.

Marra, G., Wood, S.N., 2012. Coverage properties of confidence intervals for generalized additive model components. Scand. J. Stat. 39, 53–74. https://doi.org/10.1111/j.1467-9469.2011.00760.x.

Monteith, D.T., Evans, C.D., Henrys, P.A., Simpson, G.L., Malcolm, I.A., 2014. Trends in the hydrochemistry of acid-sensitive surface waters in the UK 1988-2008. Ecol. Indicat. 37, 287–303. https://doi.org/10.1016/j.ecolind.2012.08.013.

Monteith, D.T., Stoddard, J.L., Evans, C.D., de Wit, H.A., Forsius, M., Hogasen, T., Wilander, A., Skjelkvale, B.L., Jeffries, D.S., Vuorenmaa, J., Keller, B., Kopacek, J., Vesely, J., 2007. Dissolved organic carbon trends resulting from changes in

atmospheric deposition chemistry. Nature 450, 537. https://doi.org/10.1038/nature06316. U9.
Nychka, D., 1988. Bayesian confidence intervals for smoothing splines. J. Am. Stat. Assoc. 83, 1134–1143. https://doi.org/10.2307/2290146.
Orr, H.G., Simpson, G.L., Clers, S. des, Watts, G., Hughes, M., Hannaford, J., Dunbar, M. J., Laizé, C.L.R., Wilby, R.L., Battarbee, R.W., Evans, R., 2015. Detecting changing river temperatures in England and Wales. Hydrol. Process. 29, 752–766. https://doi.org/10.1002/hyp.10181.
Pedersen, E.J., Miller, D.L., Simpson, G.L., Ross, N., 2019. Hierarchical generalized additive models in ecology: an introduction with mgcv. PeerJ 7, e6876. https://doi.org/10.7717/peerj.6876.
Polansky, L., Robbins, M.M., 2013. Generalized additive mixed models for disentangling long-term trends, local anomalies, and seasonality in fruit tree phenology. Ecol. Evol. 3, 3141–3151. https://doi.org/10.1002/ece3.707.
R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL. https://www.R-project.org/.
Renard, B., Lang, M., Bois, P., Dupeyrat, A., Mestre, O., Niel, H., Sauquet, E., Prudhomme, C., Parey, S., Paquet, E., Neppel, L., Gailhard, J., 2008. Regional methods for trend detection: assessing field significance and regional consistency. Water Resour. Res. 44, W08419. https://doi.org/10.1029/2007WR006268.
Ruppert, D., Wand, M.P., Carroll, R.J., 2003. Semiparametric Regression. In: Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.
Silverman, B.W., 1984. Spline smoothing: the equivalent variable kernel method. Ann. Stat. 12, 898–916.
Simpson, G.L., 2019. Gratia: graceful 'ggplot'-based graphics and other functions for GAMs fitted using "mgcv.
Simpson, G.L., 2018. Modelling Palaeoecological Time Series Using Generalized Additive Models. https://doi.org/10.1101/322248.
Skjelkvåle, B.L., Borg, H., Hindar, A., Wilander, A., 2007. Large scale patterns of chemical recovery in lakes in Norway and Sweden: importance of seasalt episodes and changes in dissolved organic carbon. Prague, Czech Republic Appl. Geochem., Selected Papers from the 7th International Conference on Acid Deposition 1174–1180. https://doi.org/10.1016/j.apgeochem.2007.03.040, 12–17 June, 2005 22.
Stoddard, J.L., Jeffries, D.S., Lükewille, A., Clair, T.A., Dillon, P.J., Driscoll, C.T., Forsius, M., Johannessen, M., Kahl, J.S., Kellogg, J.H., Kemp, A., Mannio, J., Monteith, D.T., Murdoch, P.S., Patrick, S., Rebsdorf, A., Skjelkvåle, B.L., Stainton, M. P., Traaen, T., van Dam, H., Webster, K.E., Wieting, J., Wilander, A., 1999. Regional trends in aquatic recovery from acidification in North America and Europe. Nature 401, 575–578. https://doi.org/10.1038/44114.
Theil, H., 1959. A rank-invariant method of linear and polynomial regression analysis. I, II, III. Nederl Akad Wetensch 53, 386–392.
van Belle, G., Hughes, J.P., 1984. Nonparametric tests for trend in water quality. Water Resour. Res. 127–136. https://doi.org/10.1029/WR020i001p00127.
von Brömssen, C., Betnér, S., Fölster, J., Eklöf, K., 2020. Visualizing trends in large-scale environmental data. Zenodo. https://doi.org/10.5281/zenodo.4051247. Dataset and code.
von Brömssen, C., Fölster, J., Futter, M., McEwan, K., 2018. Statistical models for evaluating suspected artefacts in long-term environmental monitoring data. Environ. Monit. Assess. 190, 558. https://doi.org/10.1007/s10661-018-6900-3.
Vuorenmaa, J., Augustaitis, A., Beudert, B., Bochenek, W., Clarke, N., de Wit, H.A., Dirnböck, T., Frey, J., Hakola, H., Kleemola, S., Kobler, J., Krám, P., Lindroos, A.-J., Lundin, L., Löfgren, S., Marchetto, A., Pecka, T., Schulte-Bisping, H., Skotak, K., Srybny, A., Szpikowski, J., Ukonmaanaho, L., Váňa, M., Åkerblom, S., Forsius, M., 2018. Long-term changes (1990–2015) in the atmospheric deposition and runoff water chemistry of sulphate, inorganic nitrogen and acidity for forested catchments in Europe in relation to changes in emissions and hydrometeorological conditions. Sci. Total Environ. 625, 1129–1145. https://doi.org/10.1016/j.scitotenv.2017.12.245.
Wahba, G., 1983. Bayesian "confidence intervals" for the cross-validated smoothing spline. J. R. Stat. Soc. Ser. B Methodol. 45, 133–150.
Wahlin, K., Grimvall, A., 2008. Uncertainty in water quality data and its implications for trend detection: lessons from Swedish environmental data. Environ. Sci. Policy, Monitoring and modelling diffuse pollution from agriculture for policy support: UK and European experience 11, 115–124. https://doi.org/10.1016/j.envsci.2007.12.001.
Wood, S.N., 2019. Mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation.
Wood, S.N., 2017. Generalized Additive Models: an Introduction with R, second ed. Chapman & Hall/CRC texts in statistical science. CRC Press/Taylor & Francis Group, Boca Raton.
Xie, Y., Zhang, Y., Lan, H., Mao, L., Zeng, S., Chen, Y., 2018. Investigating long-term trends of climate change and their spatial variations caused by regional and local environments through data mining. J. Geogr. Sci. 28, 802–818. https://doi.org/10.1007/s11442-018-1506-9.