EPA Public Access

Author manuscript

J Environ Eng (New York). Author manuscript; available in PMC 2019 October 01.

About author manuscripts

Submit a manuscript

Published in final edited form as:

J Environ Eng (New York). 2018 ; 144(10): . doi:10.1061/(ASCE)EE.1943-7870.0001423.

Signal Decomposition of Conductivity Sensor Measurements on the Allegheny River, Pennsylvania

Kathleen J. Brown^a, Gary A. Norris^{a,*}, Kasey D. Kovalcik^a, Ali S. Kamal^a, Kathleen Patnode^b, Matthew S. Landis^a

^aU.S. Environmental Protection Agency, Office of Research and Development, Research Triangle Park, NC 27709, United States

^bU.S. Fish and Wildlife Service, Pennsylvania Field Office, 110 Radnor Road, Suite 101, State College, Pennsylvania 16801

Abstract

Surface water conductivity measurements were used to evaluate the combined contribution of anions in western Pennsylvania from brines discharged by sources such as oil and gas wastewater treatment, coal-fired power plants, and coal mining activities. Conductivity sensor data were collected in the Allegheny River during a US Environmental Protection Agency and US Fish and Wildlife study that included seven sites covering 256 river km during the fall of 2012. Intermittent discharges, such as oil and gas wastewater, and continuous sources contributing to the conductivity were quantified using constrained and adaptive decomposition of time-series (CADETS) frequency analysis. CADETS was able to quantify the intermittent or short-term component of conductivity at sites where the intermittent fraction was 1 to 22% of the total conductivity. The demonstrated efficacy of the CADETS method for surface water quality analysis suggests it could be widely used to evaluate other water sensor data in rivers with both continuous and intermittent source impacts.

Introduction

The Allegheny River in western Pennsylvania is the source of drinking water for over half a million people, and its tributaries are impacted by many different types of water pollution sources. The major discharge sources include centralized waste-treatment facilities for oil and gas brine (CWTFs), coal-fired electric power generating stations, acid mine drainage (AMD) from historic mining wastes (Sams and Beer 2000), current mining operations, natural oil seepage, industrial manufacturing facilities, publically owned sewage treatment plants [or publically owned treatment works (POTWs)], urban stormwater runoff, roadway deicing materials, and industrial facility sewage-treatment plants (USEPA 2015; Landis et al. 2016). These sources discharge a complex mixture of inorganic and organic pollutants. A common method for evaluating general water contaminant levels in surface waters is

*Corresponding author: norris.gary@epa.gov.

Supplemental Materials

Supplemental Data

Figs. S1–S10 and Program Code S1 are available online

measuring electrical conductivity, providing an overall measure of ionic source discharges (USEPA 1982). Conductivity is a measure of electrical current through a solution expressed in units of micro-Siemens per cm ($\mu\text{S cm}^{-1}$) and is a function of the concentration of ions in solution. Conductivity measurements of all solutions change as a function of temperature; typically, specific conductivity is used to normalize measurements back to a standard temperature of 25°C.

Specific conductivity (SC) can be used to evaluate potential ecosystem impacts due to pollutant discharges into the Allegheny River. A high SC CWTF discharge on the Allegheny River was associated with impairment of juvenile northern riffleshell mussel survival (Patnode et al. 2015). Levels in the immediate vicinity of the CWTF discharge were 12,000 $\mu\text{S cm}^{-1}$ and were still elevated at 4,000 $\mu\text{S cm}^{-1}$ 400 m downstream. Native mussel abundance and diversity were lower within and downstream of the high-conductivity zone compared with upstream. Water quality criteria have been developed for chloride, an anion in CWTF discharges that is highly toxic to mussels (Canadian Council of Ministers of the Environment 2011). Increased SC has also been associated with impacts in benthic macroinvertebrates in Appalachian Mountain streams (Pond et al. 2008; USEPA 2011).

One of the primary sources of SC in the Upper Allegheny River is discharges from CWTFs treating fluids produced by oil and gas wells (Landis et al. 2016). Oil and brine treatment by CWTFs removes oil and dissolved solids (Ferrar et al. 2013) but does not effectively remove soluble monovalent ions such as Br^- , Cl^- , and Na^+ (USEPA 2015). Increased levels of Br^- in the Allegheny River have been associated with elevated levels of brominated disinfection by-products from oil and gas, and unconventional and conventional wastewater (States et al. 2013; Landis et al. 2016). The sources of Br^- , Cl^- , NO_3^- , and SO_4^{2-} ions in the Allegheny River during 2012 were quantified using source apportionment (USEPA 2015) for samples collected at 9:00 a.m. and 12:00 p.m. Eastern Standard Time (EST) during low-flow conditions (summer and autumn). The main sources of Cl^- were CWTF discharges, POTW sewage discharges, and coal-fired generating stations equipped with sulfur dioxide air pollution control flue gas desulfurization (FGD) scrubbers that discharge wastewater to the river. Nitrate sources were POTWs and FGD, and the sulfate sources were AMD and FGD.

The Br^- contribution from CWTFs to the Allegheny River was also evaluated by Weaver et al. (2015) with both steady-state and intermittent transport simulations. The simulations showed that the intermittent peak concentrations from shorter-term discharges can be four times higher than steady-state or continuous discharges, and that CWTF Br^- concentrations decreased downstream due to dilution with the addition of water from tributaries. The discharge of conventional wastewater from CWTFs continues to be a significant source of anions contaminants such as Br^- and Cl^- ; for example, an estimated 8.5×10^8 L of wastewater was discharged in 2013 (Getzinger et al. 2015; Landis et al. 2016). An evaluation of the 2012 well completion records for customers of a CWTF discharging to the Allegheny River found that 76% of the wastewater was from hydraulically fractured conventional wells (USEPA 2015).

The main anion sources discharging into the Allegheny are CWTFs and FGD scrubber discharge (USEPA 2015; Landis et al. 2016). Samples were collected from the CWTFs

impacting US Fish and Wildlife (FWS) and USEPA sites, and the conductivity was 131,000 and 133,555 $\mu\text{S cm}^{-1}$, respectively. FGD wastewater and acid mine drainage also had measured conductivity of 31,597 and 1,219 $\mu\text{S cm}^{-1}$, respectively. The internal outfalls for the wastewater discharges at a coal-fired electrical generating station were 2,945, 850, and 1,330 $\mu\text{S cm}^{-1}$ for the cooling tower, coal pile runoff, and POTW, respectively. The CWTs are intermittent discharges because they discharge typically between 6:00 a.m. and 1:00 p.m. EST during weekdays (Landis et al. 2016). FGD discharges can either be continuous or intermittent discharges depending on the wastewater-treatment process.

Conductivity sensors generate a large amount of data, creating a challenge to interpret their variability, especially across multiple monitoring sites. Frequency analysis has been used previously to evaluate irregularly spaced water sample data to identify street sweeping as a source of metals (Dilmaghani et al. 2007). In addition, the variability in river discharge signals has been used to separate the base flow from rainfall contribution events (Arnold and Allen 1999; Welderufael and Woyessa 2010). Traditional approaches such as evaluating peaks in time series have been used to identify and quantify source impacts (Ort and Siegrist 2009; Halliday et al. 2014; Landis et al. 2016); however, dilution of short-term intermittent discharges into rivers attenuates these peaks, and separation of these less pronounced signals requires frequency analysis. Wavelet frequency analysis has been used to evaluate temporal patterns in hydrology data (Whitefield and Dohan 1997; Labat 2005; Parmar and Bhardwaj 2013), and Kang and Lin (2007) demonstrated wavelet analysis of hydrology data and seasonal trends in Cl^- , Na^+ , and NO_3^- .

The primary goal of this study is to evaluate the efficacy of the Constrained and Adaptive Decomposition of Time Series (CADETS) algorithm developed by Vedantham et al. (2015) for analysis of surface water SC data and quantification of the impact of intermittent CWT discharges on downstream sampling sites. Specifically, the objectives were to (1) evaluate the anions contributing to conductivity in the upper reaches of the Allegheny River; (2) quantify the continuous and intermittent components contributing to measured conductivity on the Allegheny River using a frequency analysis approach (CADETS); (3) evaluate the sensitivity of the CADETS algorithm; and (4) compare the frequency analysis results using three different discharge scenarios and published Allegheny River results. High-frequency conductivity data contain useful information that will be lost when reducing data to daily averages. Frequency analysis uses the power of high-frequency data by using all of the data to make inferences about the system.

Materials and Methods

Field Study

Onset (Cape Cod, Massachusetts) HOBO Model U24–01 in-river sondes were deployed to measure temperature and electrical conductivity at all of the sampling sites. The measurements from two studies were combined in this analysis. The Allegheny River at Warren, Pennsylvania data collection (FWS) was described by Patnode et al. (2015), and the Allegheny River at Franklin, Pennsylvania to Freeport, Pennsylvania (EPA) was described by Landis et al. (2016) and USEPA (2015). The Allegheny River at Warren FWS sondes

collected conductivity measurements every 15 min and the EPA sondes from Franklin, Pennsylvania to Freeport, Pennsylvania collected measurements every 5 min. The FWS sondes were calibrated with an YSI model 6820V2 multiparameter instrument (Yellow Springs, Ohio) or MS-5 multiprobe (OTT Hydromet, Kempten, Germany) instrument, and EPA in-river sondes were calibrated with a YSI Model 556 multiparameter instrument. Conductance standards were used to calibrate the YSI and Hydrolab conductivity probes on a daily basis (Landis et al. 2016).

This analysis focused on the SC data collected at seven sampling sites as shown in Fig. 1. Two of the sampling sites were near intakes for public drinking water (PDW) systems. The data included in this analysis were collected from September 18, 2012, through October 4, 2012, because it allowed for the integration of SC data from EPA (Sites 1, 2, 3, 4, PDW1, and PDW2) and FWS (M1 and M6) field studies (Patnode et al. 2015). The river discharge volume increased 41% between the Warren USGS Gauge Station 03015310 and USGS Station 03025500 just upstream of Public Drinking Water System 1 (PDW1) (Franklin, Pennsylvania) with a mean of 55.3 and 78.0 m³/s, respectively (Figs. S1–S3 show hydrographs for each gauge station). The river discharge volume for PDW2 increased to 183 m³/s at the USGS Station 03049500 (Natrona, Pennsylvania) as a result of the upstream contribution of the Kiskiminetas River into the Allegheny.

Constrained and Adaptive Decomposition of Time Series

The CADETS algorithm decomposes time series data into a slow-varying time series and a fast-varying time series using techniques from signal processing. Generally, the slow-varying components can be thought of as a continuous trend, and the fast-varying component as intermittent or short-term fluctuations. Typically, signal processing decomposes time-series data into signal and noise components; this can result in negative values, which precludes quantitative analysis. CADETS constrains both components to be nonnegative, and this allows for the two resulting conductivity components to be added together to get back to the original data. CADETS separates these two components based on the fundamental frequency observed in the data set. Three main steps are used to separate the components as shown in Fig. 2: (1) identifying the low-frequency component using a filter or curve with a smoothing based on fast Fourier transform (FFT); (2) adjustment of the curve to be below the data; and (3) subtraction of the frequency curve from the measured data to quantify the high-frequency component. The first step shown in Fig. 2(a) is to subtract from the baseline (gray), apply Butterworth filter (red) based on PSD peak frequency, and divide the curve into low-frequency segments using valley points (blue symbols). The second step shown in Fig. 2(b) is adjust low-frequency segments (defined by the blue segments) so they are below the input data and connect segments with cubic spline (red). The third step shown in Fig. 2(c) is subtract the slow-varying component (red line in the middle panel) from the input data (gray) to obtain the high-frequency/intermittent component (blue line).

Fast Fourier transform takes time-series data (e.g., conductivity versus time) and transforms those data into frequencies and amplitudes. The time-series data can be thought of as the sum of many sine waves with various frequencies and amplitudes. In Fourier space, the

frequencies and amplitudes inherent to the data can be seen more clearly. Before any filtering of the data takes place, the Fourier transform can be used to plot the power spectral density to determine the fundamental frequency. The power spectral density (PSD) is a plot of frequencies present in the time series versus their amplitude, or strength of those frequencies (energy). It is useful to plot the period (1/frequency), rather than the frequency, on the x -axis to observe hourly or daily cycles present in the data. The fundamental frequency is frequency corresponding to the first prominent peak in the PSD going from fast frequencies to slow.

The fundamental frequency is given to the CADETS algorithm as the cutoff value that will serve as the bifurcation point that divides the two frequencies, which is the basis for the decomposition. The CADETS algorithm uses the fundamental frequency as a cutoff value in a low-pass filter to smooth the data by filtering off all frequencies greater than the cutoff value. This smoothed version of the data is the basis for the slowly varying component. The CADETS algorithm adaptively decomposes the two components so that the constraints (nonnegativity and additivity) are satisfied. This is accomplished by adaptively modifying the smoothed curve to create a lower envelope of the data.

Each run of the CADETS algorithm is limited to one site or location. However, CADETS can be run in multiple iterations for multiple sites. When using CADETS to analyze data from multiple sites (e.g., along a stretch of a river), the same cutoff frequency must be used for all sites to allow for direct comparison. This allows the user to understand the time-series dynamics in their data across sites. Additional information on the CADETS algorithm are presented in the Supplemental Data.

Results and Discussion

Average Conductivity and Anion Concentration

A summary of anion concentrations for the time period covering the EPA and FWS studies is shown in Table 1. The average EPA site conductivity for the date range used in this analysis were within 7.5% of the low-discharge results reported by Landis et al. (2016), which included samples collected in July, September, and the start of October 2012. The largest Cl^- difference among sites was found between the sites in Warren, Pennsylvania (M1 and M6), which were impacted by a CWTF discharge; data for additional Warren sites have been given by Patnode et al. (2015). The PDW2 intake had an elevated SO_4^{2-} concentration (89.40 mg L^{-1}) compared with the PDW1 intake (10.10 mg L^{-1}) due to the acid mine drainage in the Kiskiminetas River (USEPA 2015). Bromide concentrations also increased from PDW1 (0.092 mg L^{-1}) to PDW2 (0.149 mg L^{-1}) due to a CWTF on the Blacklick Creek and FGD scrubber wastewater discharges.

Milliequivalents per liter were calculated for each of the anion species, and each species' contribution to the total is shown in Table 2. Cl^- explained the greatest percentage at sites M1 to PDW1 (73.2–82.1%). The anion Meq L^{-1} contribution shifted from Cl^- to SO_4^{2-} between PDW1 and PDW2 due to sulfates from the Kiskiminetas River, which is impacted

by CWTF, acid mine drainage, and coal-fired power plant discharges (USEPA 2015). The smallest anion contribution was from Br⁻, which contributed less than 1% at all of the sites.

Parameters Used for Application of CADETS Algorithm

The CADETS algorithm was applied to the specific conductivity data for the seven sampling locations. A continuous conductivity component was found at all of the sites. The conductivity minus a baseline level was analyzed using CADETS, where the baseline was the minimum value of the most upstream site (M1), 133 $\mu\text{S cm}^{-1}$. Several parameters within the CADETS algorithm were chosen and fixed in the analysis of all the sites. A common cutoff frequency was used based on the PSD plots of each site. The default parameters selected for the analysis of all sites were a fifth-order Butterworth filter (Butterworth 1930) with a normalized cutoff frequency based on the evaluation of each sites data. The PSDs were examined for common peaks, and Fig. 3 contains the PSDs for each of the sites, along with the location of the selected cutoff value displayed by the red star. Site M1 displayed it first peak with a slower frequency compared with the other sites, at 1.185 days or approximately 28 h. However, the power of the peak was much lower $\{6.04 \times 10^3[(\text{mScm}^{-1})^2 \times \text{days}^2)]\}$ compared to the other sites (9.38×10^6 to 3.96×10^8). The PSD for M6 and PDW2 showed their first peak with a slightly faster frequency compared with the other sites at 0.889 days or approximately 21 h. A final cutoff value corresponding to 0.9481 days, or approximately 23 h, was chosen because it had most similarity for the first peaks of the PSD for the various sites. The cutoff period of 0.9481 days translates to a normalized cutoff frequency of 0.004. However, because the data from the FWS sites had an increment of 15 min (rather than 5 min), the normalized cutoff frequency for the sites M1 and M6 was 0.011 due to the calculation given in CADETS Step 2.

CADETS Applied to the Seven Sites

The conductivity data from this study are regularly spaced (every 5 min for some sites and every 15 min for other sites) with no missing values. The CADETS algorithm is designed to take this type of time-series data and translate it into an intermittent (high-frequency) and a continuous (low-frequency) signal. The CADETS algorithm was used to analyze the conductivity data from each of the seven sites. The same parameters were used for each site: baseline conductivity of 133 $\mu\text{S cm}^{-1}$, cutoff frequency corresponding to 0.9481 days (normalized cutoff frequency was 0.004 for 5-min increment data, and 0.011 for 15-min increment data), and fifth-order Butterworth filter.

The results of the CADETS algorithm are given in Fig. 4 for Site 3 and PDW1 and in Fig. 5 for Site M6. Results for all other sites are given in Figs. S4–S10. These plots are presenting conductivity minus baseline, where baseline is the constant minimum conductivity level of the most upstream site (M1), 133 $\mu\text{S cm}^{-1}$. This baseline represents the conductivity from the natural waters and discharges such as a POTW between M1 and the Allegheny Reservoir. CADETS separates the data into continuous, intermittent, and baseline components to allow for a comparison of the differences among the sites. The original conductivity time series are shown in Fig. 4 in gray as a reference. The initial smoothing of the raw data using the low-pass Butterworth filter is the basis for the final CADETS slowly varying component [red curve in Figs. 4(a and b)]. The fast-varying component [blue curve

in Figs. 4(c and d)] is the result of subtracting the slowly varying curve from the original time-series data.

Evaluation of the CADETS results is primarily based on the PSD plots (Fig. 3) and time-series plots, which show a comparison of the intermittent, continuous, and measured data (Fig. 4). Comparing sites with different discharge impacts provides support that CADETS is correctly separating out the intermittent component. The continuous component could only be evaluated by comparing the concentrations among sites because the time period of this analysis was only 16 days. The M6 site was 3.7 km downstream of both CWTF and POTW discharges (Patnode et al. 2015). The M6 site had a high intermittent signal ($33 \mu\text{S cm}^{-1}$) that was 18% of the total conductivity and had both a sharp bimodal on/off discharge pattern and an increasing continuous component (Fig. 5). The CADETS algorithm captured both the high-conductivity variability and increase in the continuous component from Days 12 to 14. Another site with a similar conductivity signal 45 m downstream of a CWTF discharge (CWTF-P45) was evaluated by Landis et al. (2016), and the PSD identified a shorter fundamental period of 0.33 days, which corresponds to the operating schedule for the CWTF, which was approximately 6:00 a.m. to 1:00 p.m. EST Monday through Friday. The continuous component of conductivity was not consistently fit at this site, indicating that CADETS may not be appropriate for evaluating data collected in close proximity to discharge locations.

A summary of the CADETS intermittent and continuous components results from all seven sites is given in Table 3, and the slow-varying and fast-varying signals for each site are shown in Figs. S4–S10. The summary statistics (average percent and median conductivity) assume that the two CADETS conductivity components are independent of time. The intermittent and continuous conductivity components are presented in Table 3 as a fraction of the time-varying components (continuous plus intermittent) and of the intermittent as a fraction of the total conductivity (baseline plus continuous plus intermittent).

In addition, the intermittent fraction of the continuous plus baseline is given in Table 3, and the intermittent fraction of the conductivity represents from 1 to 22% of the total conductivity. It is useful to look at the CADETS results comparing Site 3 and PDW1, as shown in Fig. 4. Site 3 has higher-amplitude intermittent pulses compared with PDW1. The slow-moving or continuous component peaks for Site 3 and PDW1 have similar peak patterns and are offset due to transport time. However, the PDW1 continuous conductivity contribution is higher due to sources such as historic AMD, which increased the SO_4^{2-} concentrations.

CADETS results were compared with low-flow CWTF-impacted ($n=10$) conductivity results presented by Landis et al. (2016), which were composite grab samples collected at 9:00 a.m. and 12:00 p.m. EST. The Landis et al. (2016) analysis was able to classify samples as impacted by a CWTF based on information from the facility; an advantage of the CADETS model is that it does not require any a priori discharge information. Although the data evaluated in this analysis were a subset of the all of the low river discharge data of Landis et al. (2016), a comparison of the CWTF intermittent signal is similar for the Site 3 contribution (CADETS intermittent contribution of $26 \mu\text{S cm}^{-1}$ compared with $28 \mu\text{S cm}^{-1}$

based on composite grab sample results). Differences in the conductivity decompositions are expected for the two approaches (CADETS and Landis et al. 2016) due the differences in the data sets: continuous time-series data are not reliant on the classification of being CWTF discharge impacted or nonimpacted, and CADETS is based on all available data. The results are also consistent in showing that the Site 3 CWTF or intermittent total conductivity contribution using two different approaches was low (15%).

Sensitivity Analysis of CADETS Algorithm

The sensitivity of the CADETS results was evaluated by changing the order of the Butterworth filter and the cutoff frequency for Sites 3 and PDW1. The order of the Butterworth filter was tested for $n=1$ to 7, keeping all other parameters the same as described previously. Similarly, the normalized cutoff frequency was tested for several other plausible first peaks in the power spectral density plots (period = {0.8889, 0.9481, 1.185, 1.422, 1.778} days). For the two sites, the changes of these parameters were examined for their effect on the partitioning between the intermittent and continuous components (percentage).

The order of the Butterworth filter and cutoff frequency were varied to quantify the sensitivity of the solution to these user-specified parameters (Tables 4 and 5). The sensitivity of the CADETS results were evaluated using the conductivity data from Site 3 and PDW1 because these sites represent different source impact scenarios, with Site 3 being 12.3 km downstream of a CWTF discharge and PDW1 being 51.8 km downstream. The order of the Butterworth filter, a key parameter, was varied from 1 to 7 to determine the sensitivity of the solution (Table 4). Compared with the results for the selected fifth-order Butterworth filter, decreasing the order increased the continuous component by 14% (or 5%, omitting first order) for Site 3 and by 9% for PDW1. Increasing the order to 7 resulted in a decrease of 0.8% for Site 3 and 5% for PDW1. Changing the order ± 1 did not greatly affect the results. However, for $n>7$, the data are smoothed to a monotonic curve, and the CADETS algorithm breaks down. In addition, several cutoff frequencies (Table 5) were tested. Changing the cutoff frequency from the selected value of 0.9481 days decreased the continuous component by as much as 4% and increased it by as much as 2% for Site 3. Modifying the cutoff frequency for PDW1 had a larger effect for PDW1, with the continuous component increasing as much as 6%.

Discharge Scenarios

Three discharge scenarios were present on the Allegheny River, allowing for the evaluation of the CADETS results: (1) M6 to Site 1, which had a stable background and no significant additional intermittent sources; (2) Site 3 to PDW1, which had a combination of a significant CWTF intermittent source (Site 3) that was diluted due to flow accretion and continuous sources (AMD and POTW), leading to an increase in the continuous component; and (3) PDW1 to PDW2, which had increases in both components, which in turn indicate that intermittent and continuous sources are contributing to the elevated conductivity.

For the first case, with no significant additional intermittent sources in the river section from M6 to Site 1, the intermittent contribution was attenuated from 61 to 15% moving downstream, and it was a small fraction of the total conductivity at Site 1 (3%). This is likely

due to the dilution of the discharged anions contributing to conductivity with water from direct runoff and tributary inflow. Landis et al. (2016) shows an example of CWTF signal attenuation from a site near the Franklin CWTF discharge and a site located 45 m downstream.

The second case of the river section between Site 3 and PDW1 shows an increase in the intermittent contribution due to the CWTF discharge just upstream of Site 3, followed by the increase in the continuous component due to dilution and other continuous sources on the Allegheny, including acid mine drainage. An 85-MW bituminous waste coal-fired power plant was located between Sites 3 and 4 and discharges were both continuous (cooling tower blow down) and intermittent (stormwater runoff). The comparison of the CADETS results in Fig. 4 shows that the intermittent pulse decreases similar to the monocacy results presented by Weaver et al. (2015).

The last case from PDW1 to PDW2 demonstrated an increase in conductivity from 221 to 391 $\mu\text{S cm}^{-1}$, respectively (Table 1), in conductivity. The intermittent component increased 60 $\mu\text{S cm}^{-1}$ from PDW1 to PDW2, showing that additional intermittent sources are contributing to the conductivity at PDW2. In addition, the continuous component increased 72 $\mu\text{S cm}^{-1}$ from PDW1 to PDW2, which may be due to acid mine drainage. The intermittent fraction of the total conductivity was 22% at PDW2, indicating that even for a site impacted by multiple intermittent sources, the background and continuous components explained most of the conductivity. The time series of the continuous and intermittent conductivity do not have a consistent pattern between PDW1 and PDW2 due to the multiple intermittent and continuous sources of conductivity, and the main Cl^- and SO_4^{2-} sources at PDW2 were from a CWTF and coal-fired generating stations discharging from Blacklick Creek to Kiskiminetas River (USEPA 2015).

The CADETS time-series analysis of the intermittent and continuous contributions among sites provides novel information on the sources of intermittent and continuous SC increases (Table 3). Sites downstream of known sources had the greatest increase in intermittent contributions, with sites M1–M6 increasing 31 $\mu\text{S cm}^{-1}$, Sites 1–3 increasing 20 $\mu\text{S cm}^{-1}$, and PDW1 to PDW2 increasing 60 $\mu\text{S cm}^{-1}$. It is more difficult to pinpoint the sources that resulted in the monotonic increase along the entire study segment for the conductivity associated with continuous sources; however, it is possible to identify potential sources. For example, the Site 1 continuous discharge increase may be due to other sources, such as a POTW treating high-halide waste (USEPA 2015). The intermittent fraction of the total conductivity was 18 and 13% of the total conductivity at sites downstream of known short-term discharges (M6 and Site 3). This shows that although conductivity is a common water sensor, its signal is dominated by baseline natural water conductivity and other sources with slower-moving contributions.

Conclusions

The CADETS algorithm was applied to river conductivity time-series data and produced decomposed time series that could be interpreted as baseline, intermittent, and continuous components. CADETS results were compared with CWTF impacted versus nonimpacted

peak analysis reported by Landis et al. (2016), and similar increases in conductivity were found by CADETS without classifying the data based on discharges. The continuous component of the conductivity increased at the downstream sites, which was consistent with the SO_4^{2-} data indicating increased acid mine drainage and other sources such a generating station. Sensitivity analysis demonstrated that parameter decisions for the CADETS algorithm may influence the interpretation of the results up to 14%, although overall the sensitivity analysis showed fairly consistent results.

The CADETS algorithm quantifies the intermittent and continuous signals and provides a method for identifying sources' contribution to conductivity with commercially available conductivity sensors. Multiple components contribute to the conductivity, including baseline, continuous, and intermittent; CADETS provides a way to separate out these components using an entire data set and data from multiple sites even if pronounced peaks are not present. This methodology can be used for evaluation of other water-sensor data to help quantify both known and unknown contaminant sources. In addition, distributions of continuous and intermittent signals may provide exposure metrics for ecological studies, as well as identify reaches within large rivers where cumulative discharges pose unacceptable ecological risk.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The EPA through its Office of Research and Development funded and conducted this research. The views expressed in this paper are those of the authors and do not necessarily reflect the views or policies of EPA or the FWS. It has been subjected to Agency review and approved for publication. Mention of trade names or commercial products do not constitute an endorsement or recommendation for use.

References

- Arnold JG, and Allen PM. 1999 "Automated methods for estimating baseflow and ground water recharge from streamflow." *J. Am. Water Resour. Assoc* 35 (2): 411–424. 10.1111/j.1752-1688.1999.tb03599.x.
- Butterworth S1930 "On the theory of filter amplifiers." *Wirel. Eng* 7 (10): 536–541.
- Canadian Council of Ministers of the Environment. 2011 Canadian water quality guidelines: Chloride ion scientific criteria document. Winnipeg, Canada: Canadian Council of Ministers of the Environment.
- Dilmaghani S, Henry I, Soonthornnonda P, Christensen ER, and Henry RC. 2007 "Harmonic analysis of environmental time series with missing data or irregular sample spacing." *Environ. Sci. Technol* 41 (20): 7030–7038. 10.1021/es0700247. [PubMed: 17993144]
- Ferrar KJ, Michanowicz DR, Christen CL, Mulcahy N, Malone SL, and Sharma RK. 2013 "Assessment of effluent contaminants from three facilities discharging Marcellus Shale wastewater to surface waters in Pennsylvania." *Environ. Sci. Technol* 47 (7): 3472–3481. 10.1021/es301411q. [PubMed: 23458378]
- Getzinger GJ, O'Connor MP, Hoelzer K, Drollette BD, Karatum O, Deshusses MA, Ferguson PL, Elsner M, and Plata DL. 2015 "Natural gas residual fluids: Sources, endpoints, and organic chemical composition after centralized waste treatment in Pennsylvania." *Environ. Sci. Technol* 49 (14): 8347–8355. [PubMed: 26147419]

- Halliday SJ, Skeffington RA, Bowes MJ, Gozzard E, Newman JR, Loewenthal M, Palmer-Felgate EJ, Jarvie HP, and Wade AJ. 2014 “The water quality of the River Enborne, UK: Observations from high-frequency monitoring in a rural, lowland river system.” *Water*6 (1): 150–180. 10.3390/w6010150.
- Kang S, and Lin H. 2007 “Wavelet analysis of hydrological and water quality signals in an agricultural watershed.” *J. Hydrol*338 (1–2): 1–14. 10.1016/j.jhydrol.2007.01.047.
- Labat D. 2005 “Recent advances in wavelet analyses. 1: A review of concepts.” *J. Hydrol*314 (1–4): 275–288. 10.1016/j.jhydrol.2005.04.003.
- Landis MS, Kamal AS, Kovalcik KD, Croghan C, Norris GA, and Bergdale A. 2016 “The impact of commercially treated oil and gas produced water discharges on bromide concentrations and modeled brominated trihalomethane disinfection byproducts at two downstream municipal drinking water plants in the upper Allegheny River, Pennsylvania, USA.” *Sci. Total Environ*542: 505–520. 10.1016/j.scitotenv.2015.10.074. [PubMed: 26520274]
- Ort C, and Siegrist H. 2009 “Assessing wastewater dilution in small rivers with high resolution conductivity probes.” *Water Sci. Technol*59 (8): 1593–1601. 10.2166/wst.2009.174. [PubMed: 19403973]
- Parmar KS, and Bhardwaj R. 2013 “Wavelet and statistical analysis of river water quality parameters.” *Appl. Math. Comput*219 (20): 10172–10182. 10.1016/j.amc.2013.03.109.
- Patnode KA, Hittle E, Anderson RM, Zimmerman L, and Fulton JW. 2015 “Effects of high salinity wastewater discharges on unionid mussels in the Allegheny River, Pennsylvania.” *J. Fish Wildlife Manage*6 (1): 55–70. 10.3996/052013-JFWM-033.
- Pond GJ, Passmore ME, Borsuk FA, Reynolds L, and Rose CJ. 2008 “Downstream effects of mountaintop coal mining: Comparing biological conditions using family- and genus-level macro invertebrate bioassessment tools.” *J. North Am. Benthol. Soc*27 (3): 717–737. 10.1899/08-015.1.
- Sams JI, and Beer KM. 2000 “Effects of coal-mine drainage on stream water quality in the Allegheny and Monongahela River Basin: Sulfate transport and trends” *Water-Resources Investigations Rep.* 99–4208. Reston, VA: US Geological Survey.
- States S, Cyprych G, Stoner M, Wydra F, Kuchta J, Monnell J, and Casson L. 2013 “Marcellus Shale drilling and brominated THMs in Pittsburgh, PA, drinking water.” *J. Am. Water Works Assoc*105 (8): E432–E448. 10.5942/jawwa.2013.105.0093.
- USEPA. 1982 Method 120.1: Conductance (specific conductance, μmhos at 25°C) by conductivity meter. Washington, DC: US Environmental Protection Agency.
- USEPA. 2011 A field-based aquatic life benchmark for conductivity in central Appalachian streams. EPA/600/R-10-023F. Washington, DC: Office of Research and Development.
- USEPA. 2015 Sources contributing inorganic species to drinking water intakes during low flow conditions on the Allegheny River in western Pennsylvania. EPA/600/R-14/430. Washington, DC: Office of Research and Development.
- Vedantham R, Hagler GS, Holm K, Kimbrough S, and Snow R. 2015 “Adaptive decomposition of highly resolved time series into local and non-local components.” *Aerosol Air Quality Res.*15 (4): 1270–1280. 10.4209/aaqr.2015.02.0103.
- Weaver JW, Xu J, and Mravik SC. 2015 “Scenario analysis of the impact on drinking water intakes from bromide in the discharge of treated oil and gas wastewater.” *J. Environ. Eng*142 (1): 04015050 10.1061/(ASCE)EE.1943-7870.0000968.
- Welderufael WA, and Woyessa YE. 2010 “Stream flow analysis and comparison of base flow separation methods: Case study of the Modder River Basin in Central South Africa.” *Eur. Water*31: 3–12.
- Whitefield PH, and Dohan K. 1997 “Identification and characterization of water quality transients using wavelet analysis. II: Application to electronic water quality data.” *Wat. Sci. Tech*36 (5): 337–348.

Fig. 1.
Allegheny River sampling sites.

Fig. 2.
CADETS algorithm steps.

Fig. 3. Power spectral density of conductivity time-series data from sites along the Allegheny River. The red star indicates the period that was chosen as the cutoff frequency in the CADETS algorithm, which is the closest to the first peak in the PSD common to all sites.

Fig. 4. CADETS results for (a and c) Site 3; and (b and d) PDW1 with a baseline of $133 \mu\text{S cm}^{-1}$.

Fig. 5. CADETS results for site M6 showing the strong step function intermittent (blue curve) discharge pattern (baseline= $133 \mu\text{S cm}^{-1}$).

Table 1.

Mean values and ranges (minimum – maximum) for conductivity ($\mu\text{S cm}^{-1}$), and anion concentrations (mg L^{-1}) for n number of daily samples, grouped by site

Site	N	Conductivity	Br ⁻	Cl ⁻	NO ₃ ⁻	SO ₄ ²⁻
M1	1	145	0.0285	14.6	0.867	6.57
M6	1	215	0.136	29.5	1.30	7.63
Site 1	16	180 (159–194)	0.073 (0.033–0.117)	21.0 (16.2–25.8)	0.847 (0.355–1.26)	7.43 (6.78–8.60)
Site 3	16 ^a	202 (178–262)	0.117 (0.046–0.277)	25.3 (17.9–41.3)	0.696 (0.282–1.09)	8.42 (7.75–9.36)
Site 4	16	212 (191–243)	0.089 (0.045–0.154)	24.0 (18.7–31.3)	0.381 (0.012–0.807)	9.19 (7.98–11.7)
PDW1	11	221 (199–250)	0.092 (0.043–0.166)	24.4 (18.6–32.0)	0.499 (0.261–1.20)	10.10 (8.50–10.8)
PDW2	14	391 (302–459)	0.149 (0.107–0.202)	31.7 (26.1–37.3)	2.33 (1.21–3.79)	89.40 (55.8–129)

^aBr⁻Br⁻ was quantified in 15 samples.

Table 2.Average percent MeqL⁻¹ for each anion

Site	Br ⁻	Cl ⁻	NO ₃ ⁻	SO ₄ ²⁻
M1	0.1	73.2	2.5	24.3
M6	0.2	82.1	2.1	15.7
Site 1	0.1	77.8	1.8	20.3
Site 3	0.2	79.1	1.2	19.5
Site 4	0.1	77.4	0.7	21.8
PDW1	0.1	75.8	0.9	23.2
PDW2	0.1	32.0	1.3	66.6

Table 3.

Summary of CADETS results for each of the seven sites

Site	Intermittent			Continuous		
	Median ($\mu\text{S cm}^{-1}$)	Time varying (%) ^a	Total conductivity (%) ^b	Median ($\mu\text{S cm}^{-1}$)	Time varying (%) ^a	Total conductivity (%) ^b
M1	2	50	1	2	50	1
M6	33	61	18	21	39	11
Site 1	6	15	3	34	85	20
Site 3	26	39	13	40	61	20
Site 4	8	11	4	62	89	31
PDW1	16	18	7	71	82	32
PDW2	76	35	22	143	65	41

Note: Median conductivity levels and average percent of each component (intermittent and continuous) are reported.

^aFraction of time-varying components (intermittent plus continuous) excluding baseline of $133 \mu\text{S cm}^{-1}$.

^bFraction of total conductivity (baseline plus intermittent plus continuous).

Table 4.

Effect of changing order of the Butterworth filter for Site 3 and PDW1 on intermittent/continuous partition

Order	Site 3		PDW1	
	Intermittent (%)	Continuous (%)	Intermittent (%)	Continuous (%)
1	26	74	10	90
2	36	64	13	87
3	38	62	13	87
4	37	64	16	84
5	41	60	19	81
6	40	61	19	81
7	41	59	23	77

Table 5.

Cutoff values chosen based on PSD first peaks of all sites, with effects shown for Site 3 and PDW1

Period (days)	Normalized cutoff frequency	Site 3		PDW1	
		Intermittent (%)	Continuous (%)	Intermittent (%)	Continuous (%)
0.8889	0.004	41	59	16	84
0.9481	0.0037	41	60	19	81
1.185	0.003	39	61	18	82
1.422	0.0025	44	56	13	87
1.778	0.002	39	61	18	82