VAST Challenge 2018 Reviewer Guide: Mini-Challenge 2

This document provides information to support peer review of submissions to VAST Challenge 2018, Mini-Challenge 2. It covers background about the submission structure, the challenge problem, tasks and questions presented to participants, potential answers, and evidence found in the Challenge data supporting these answers. For a full description of the challenge problems and to access the data provided to the participants, please visit http://vacommunity.org/VAST+Challenge+2018.

Submissions

Participants are required to submit their entries on a standard answer form, along with a video explaining how visual analytics were used to help solve the challenges. Please consider both parts of the submission in your review. If you have difficulty reading the answer form or playing the video, please contact us at vast-challenge@ieeevis.org for assistance.

Scenario

We have oriented the VAST Challenge 2018 around an invented environmental scenario. This year's scenario extends the one we created in 2017, and if you would like additional information about that work, please see the 2017 Challenge material at http://www.vacommunity.org/VAST+Challenge+2017 . The 2018 participants did not have to participate in the 2017 Challenge, nor would they have needed to extensively review that material to submit this year. The storyline may make a little more sense if some of the 2017 narrative was reviewed, however.

Overview

Mistford is a mid-size city located to the southwest of the Boonsong Lekagul Wildlife Preserve (BLWP). The city has a small industrial area with four light-manufacturing endeavors. Mistford and the wildlife preserve are struggling with the possible endangerment of the Rose-Crested Blue Pipit, a locally loved bird. The bird's nesting pairs seem to have decreased alarmingly, prompting an investigation last year implicating Kasios Office Furniture, a Mistford manufacturing firm. Since the initial investigation, the situation has evolved: Kasios insists that they have done nothing wrong! They assert that grad student Mitch Vogel and his professors are mere media-seekers trying to draw attention away from their lackadaisical research. Kasios presents itself as an extremely eco-friendly organization. They have launched their own very public investigation into the issues raised last year and are reporting very different results! It's time to apply your visual analytics expertise to help illuminate the path to good science.

Mini-Challenge 2, "Like a Duck to Water"

Last year, the Kasios Furniture Company was implicated in environmental damage to the Boonsong Lekagul Wildlife Preserve for both dumping toxic waste and polluting the air with chemicals from its manufacturing process. But Kasios is not taking these accusations lying down, and they deny any

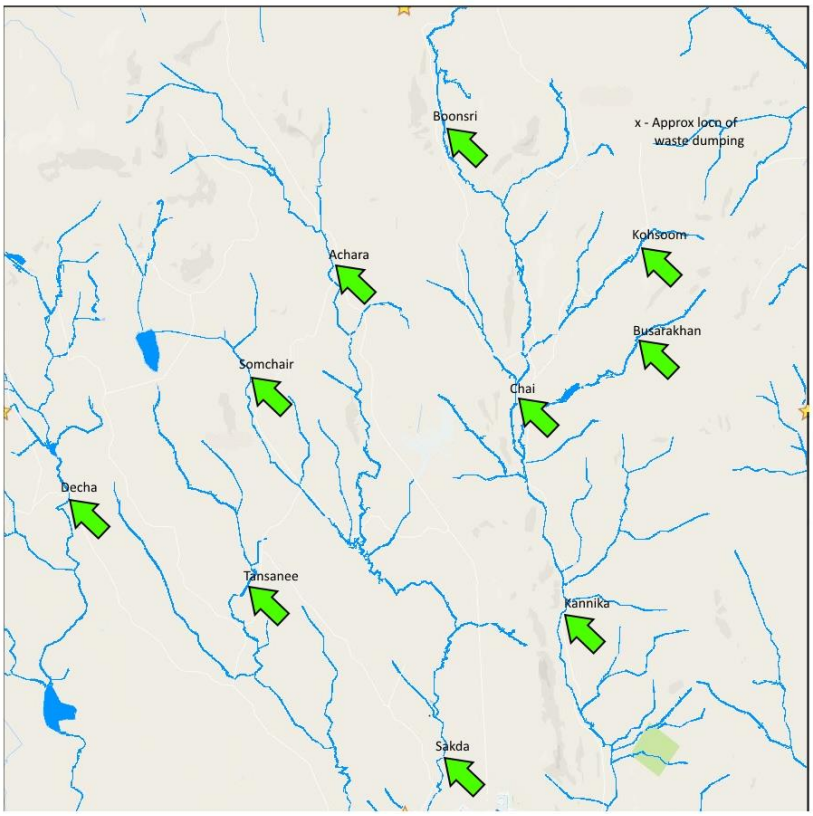
accusation of industrial waste dumping! Kasios' spokespersons state that there isn't any ground contamination near the remote ranger station that was suggested by last year's mini-challenge 1 and 3 participants, and they have inspected that area and found it as pristine as the rest of the preserve.

Outraged ornithology professors from Mistford College journeyed out to look over the dumping site themselves and perform soil analyses. They returned to report that the site looked like there had been recent excavation and building activities going on. Boonsong Preserve rangers later confirmed that a new ranger station was being built at that site! Soil samples taken from the site were inconclusive in detecting Methylosmolene (the toxic manufacturing chemical in the suspected dumping) or any other contaminant, as new top soil had been trucked in.

With a primary piece of evidence against Kasios now gone, investigators will need to take another approach. Professors in the Mistford College Hydrology Department have come forward with several years of water sensor readings from rivers and streams in the preserve. These samples were taken from different locations scattered throughout the area and contain measurements of several chemicals of possible interest, but they have never been analyzed due to lack of funding. Could visual analytics help reveal something in this data that could make up for the soil evidence that was destroyed?

The contestant's tasks were to investigate the hydrology data from across the Preserve. They were given a map of the Preserve (with the same base as last year's challenge), with named sampling sites indicated on the map (the names have local significance, but are just mnemonics for your study). They were also provided with readings from each sampling station over time for several different chemicals and water properties.

Ground Truth



A map of the Boonsong Lekagul Wildlife Preserve showing the 10 sites that have hydrology measurements

Data

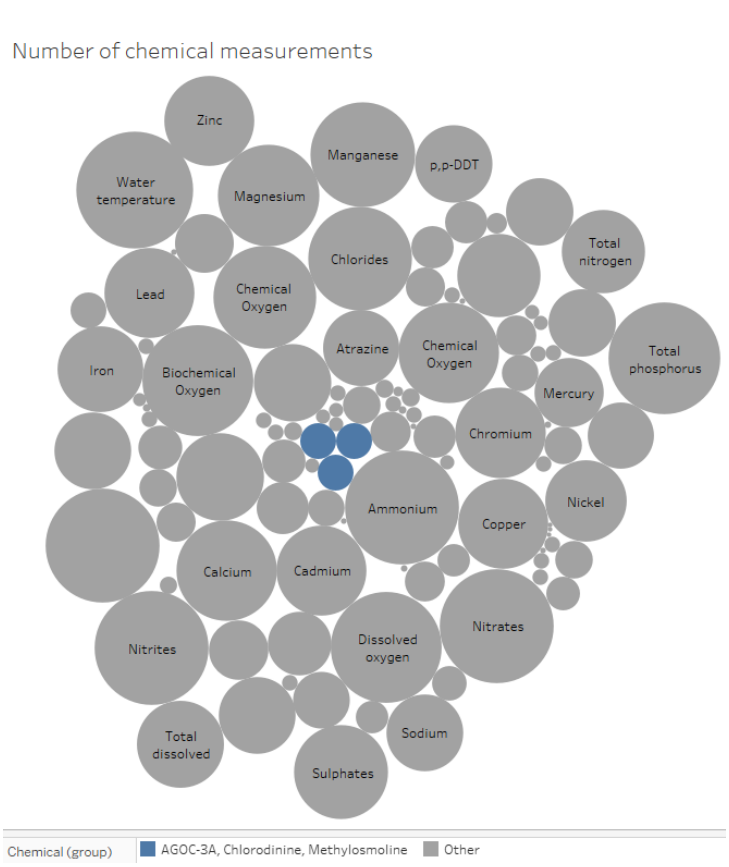
Challenge participants were provided with the above map of the BLWP showing 10 sites where hydrology measurements have been made over the years. Participants were given a csv file containing hydrology measurements, a readme file to explain the data fields, and a file giving the chemical units of measure. There are 106 different hydrology measurements in the data set. Data goes back to January of 1998 and ends in December 2016. In total, there are more than 136,000 hydrology measurements in the data set.

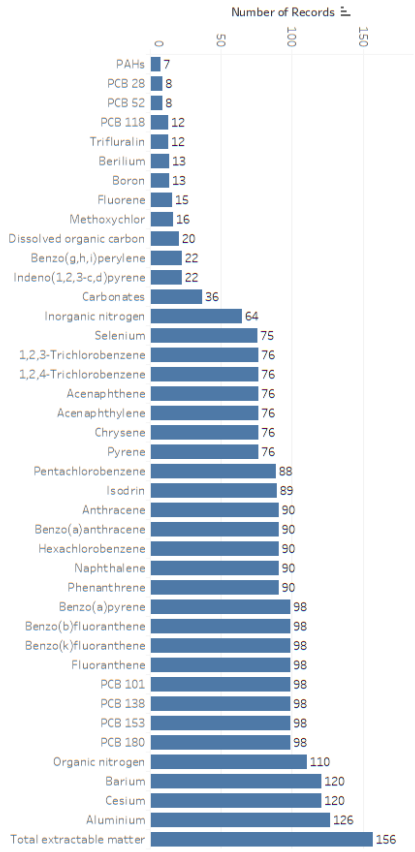
Contestant Questions

1. *Characterize the past and most recent situation with respect to chemical contamination in the Boonsong Lekagul waterways. Do you see any trends of possible interest in this investigation? Your submission for this questions should contain no more than 10 images and 1000 words.*

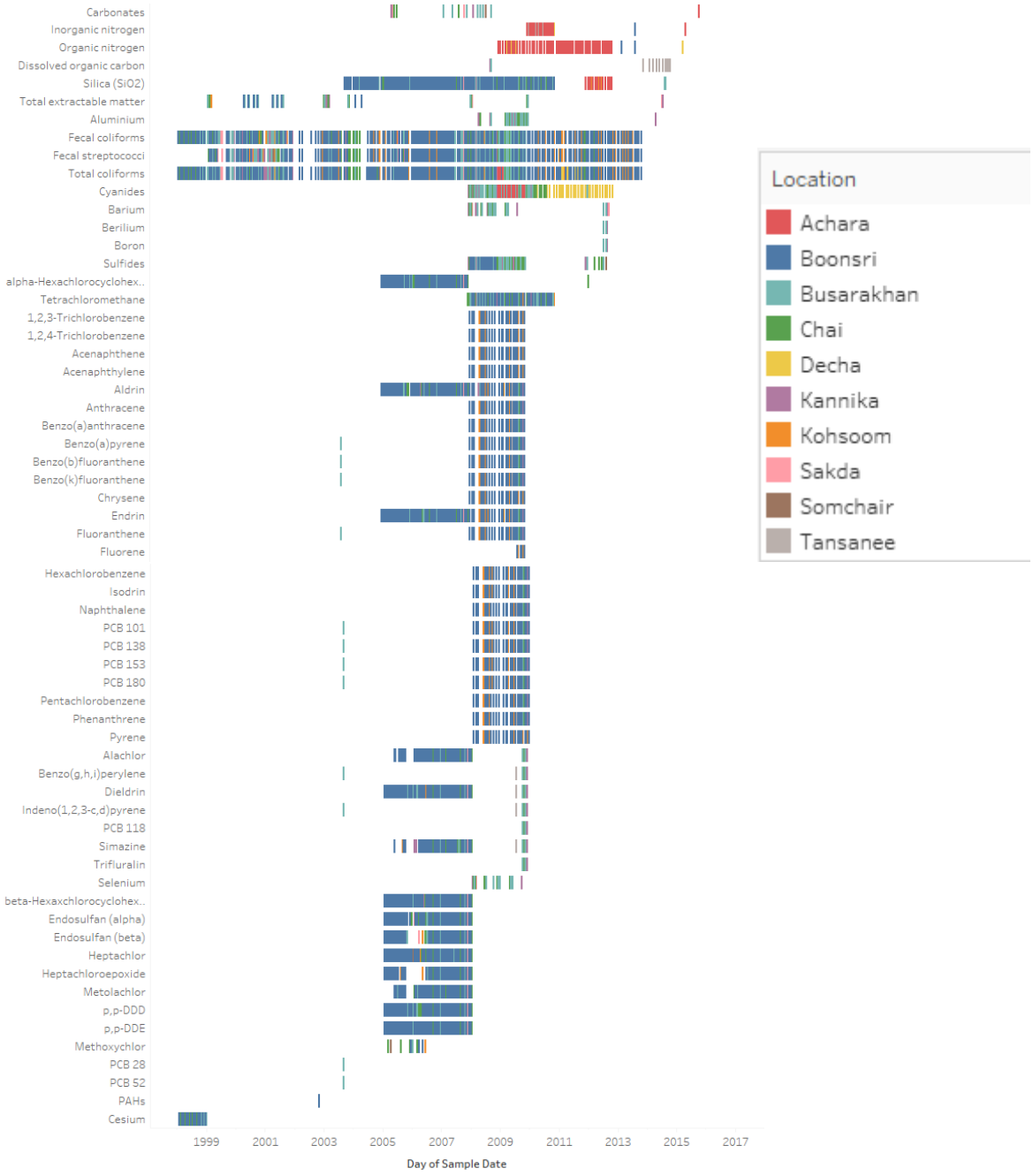
   With 10 locations, 106 chemicals, and nearly 20 years of data it is nearly impossible to succinctly and clearly present all of the chemical data. Instead, it is reasonable to characterize the past and recent situation with a representative subset of the data. Visual analytics may be used to choose that subset. For example, some chemicals are measured frequently and in many locations, while others are measured in only a few locations over a short time period. Many chemicals have too few

readings to definitively show if there have been changes over time. These are not useful for this analysis of overall trends. The visualizations below show the relative number of measurements for all of the chemicals (left) and the least frequently measured 41 chemicals (right). These 41 chemicals can be removed from this analysis





A further group of chemicals do not cover enough time to give an accurate picture over time:

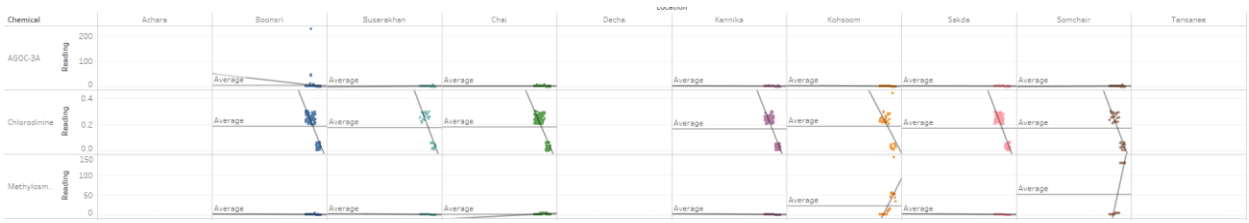


This visualization shows how time limited some chemical measurement are. Many chemicals have no recent data.

Further, there are several chemicals that show no upward or downward trend over the whole period so they are not of interest.

There are a total of 30 water measurements with interesting trends.

Visualization of trends for key chemicals of interest

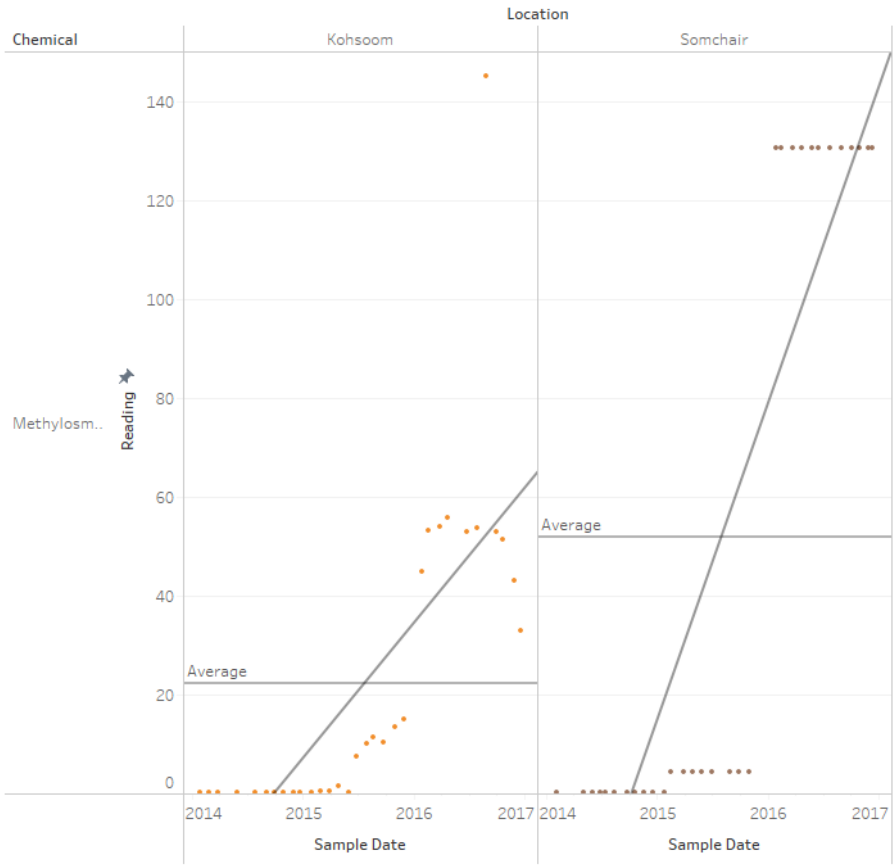


Note that there is no chlorodinine, AGOC-3A or Methylosmoline data (these were key chemicals of interest in 2017, in case any of the submissions refer to them) for Achara, Decha, and Tansanee. Chlorodinine appears to be declining in every location. AGOC-3A was not observed often – only two elevated readings at Boonsri, which is also the most measured location.

Methylosmoline is the chemical that Kasios was accused of dumping in the preserve. This video shows the full history of Methyl readings for all of the sites. It is clear that there is a significant increase in this chemical in the last year, especially in two locations: Kohsoom and Somchair.

ice_video_20180710-140945.webm (Click this link to open the video – Windows media player or VLC media player may be used to play it)

A closer look at Methyl readings reveals that there was a sharp uptick in Methylosmoline at both Kohsoom and Somchair.



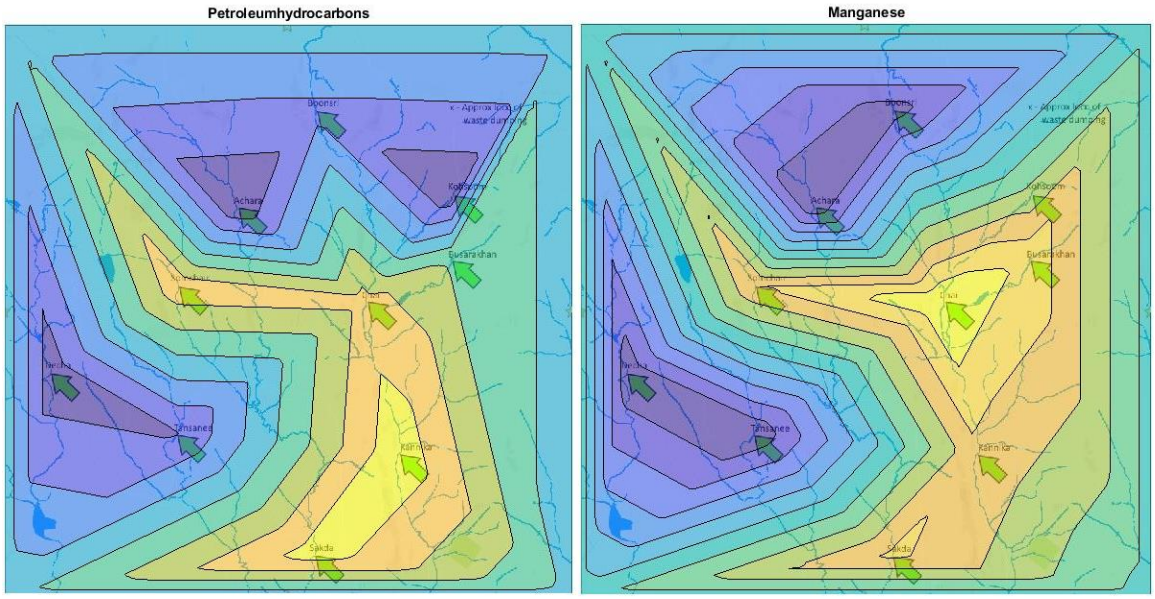
The readings at Somchair are odd because there is no variation. The measurements are flat over time then they shift up and are flat again for almost a year. Finally, there is a dramatic shift up in the last year, but no variation. This could be the result of systematic measurement error (i.e., the zero level of the measurement was not accurate and shifted, which made it appear that there was contamination when there was not. This is not the case at Kohsoom, where evidence of Methyl contamination is stronger. There is a normal variation in readings that follow a strong increasing trend since the middle of 2015. It is important to note that Kohsoom is also the closest site to the dumping area. There is little doubt that Methylosmolene was detected at this site, indicating past contamination. However, the most recent data at Kohsoom shows declining Methyl levels. This could indicate that Kasios has stopped dumping and it is simply taking time for contaminants to clear out of the watershed. Further monitoring is recommended at Kohsoom and other sites in the same watershed.

Since Kohsoom is the site where contamination was most likely detected it is worthwhile to take a closer look at other chemical readings at this site. Below are a set of relevant measurements from Kohsoom (all of the chemicals that showed a significant increasing or declining trend over the measurement period).

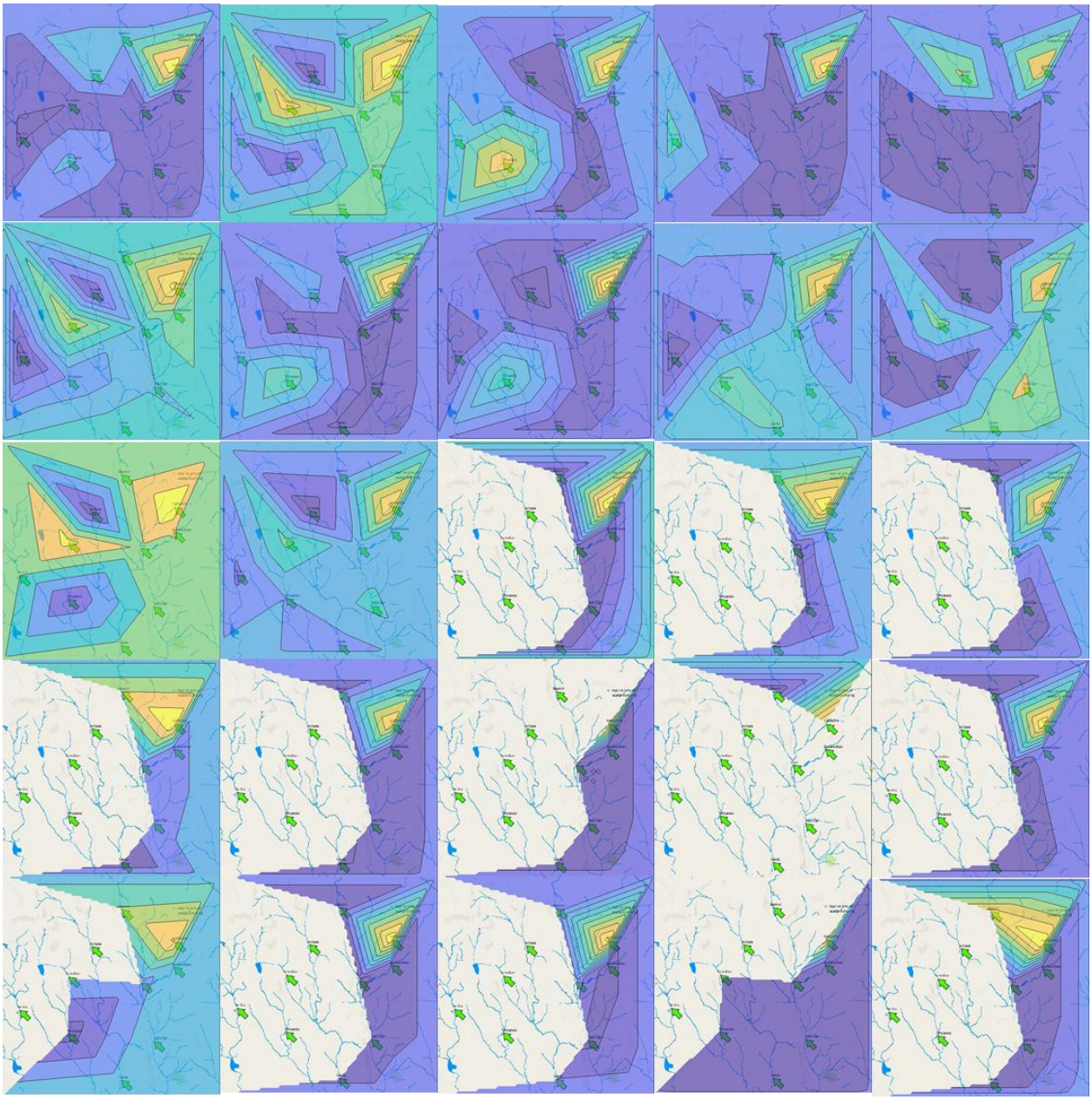


Kohsoom saw increases in Ammonium, Arsenic, dissolved silicates, Methylosmoline, Fecal Streptococci, and Fecal coliforms. There were declines in AGOC-3A, chlorodinine, Oxygen saturation, Sulphates, bicarbonates, and Zinc.

Maps of chemical averages can also reveal interesting trends. Contour plots were generated for the whole preserve using the average values of chemical readings over the entire available time period. This first group of four is representative of typical distribution of chemicals. The purple areas have lower historical averages while the yellow areas have had the highest average readings, historically. Petroleum hydrocarbons and manganese have data at every location in the preserve while the maps for aluminum and AGOC-3A show large areas of missing data (uncolored areas do not have data available).

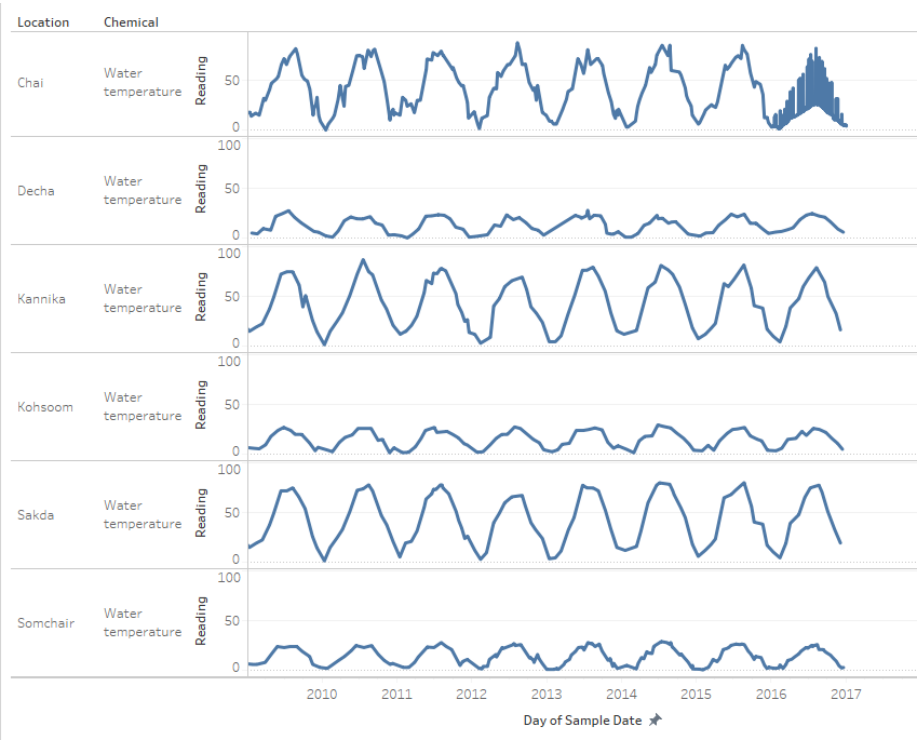


Using this mapping approach it was easy to find 25 chemicals that had high historical averages relative to the rest of the preserve. Kohsoom is often the center of elevated chemical readings at the top right of each map. The 25 elevated readings at Kohsoom include: biochemical oxygen, copper, dissolved silicates, fecal streptococci, heptachloroepoxide, hexachlorobenzene, indeno(123)cdpyrene, iron, lead, mercury, methylosmoline, nickel, orthophosphatephosphorus, PAHs, PCB-180, PCB 28, pyrene, simazine, water temperature, total dissolved phosphorus, total phosphorus, fecal coliforms, ammonium, AOX, and anionic active surfactants.

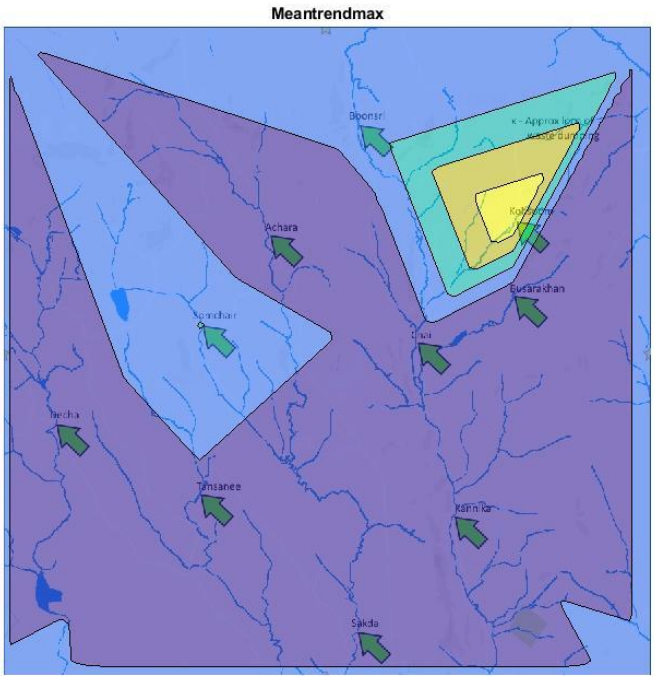


2. *What anomalies do you find in the waterway samples dataset? How do these affect your analysis of potential problems to the environment? Is the Hydrology Department collecting sufficient data to understand the comprehensive situation across the Preserve? What changes would you propose to make in the sampling approach to best understand the situation? Your submission for this question should contain no more than 6 images and 500 words.*

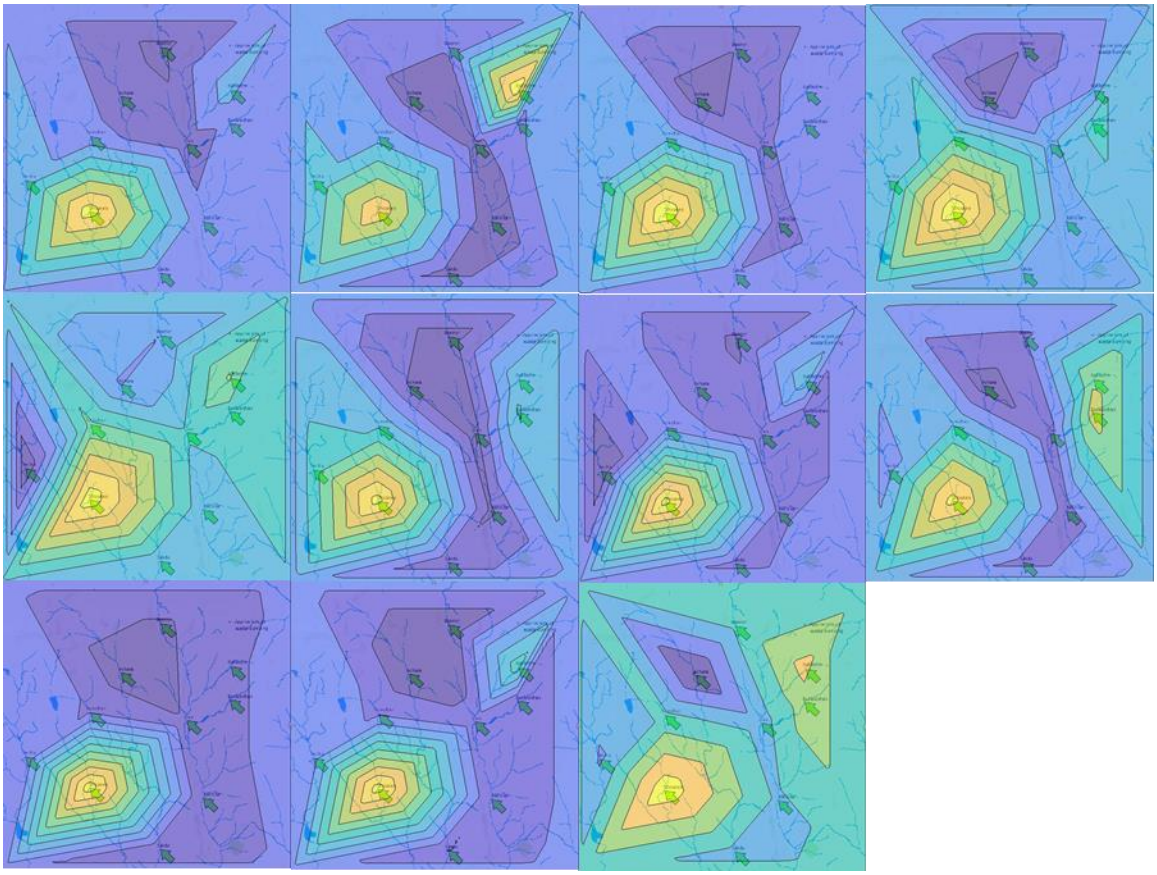
There seems to be an error in the water temperature measurements at Chai. During the last year it frequently measured much colder than historical averages. It's possible that waste that was held at lower temperature was dumped at this site. However, there are no other strong indications of contamination at Chai.



A new visualization was developed to determine which site is the most perturbed in terms of recent chemical readings compared to historical averages. In this case, Kohsoom was the clear anomaly:



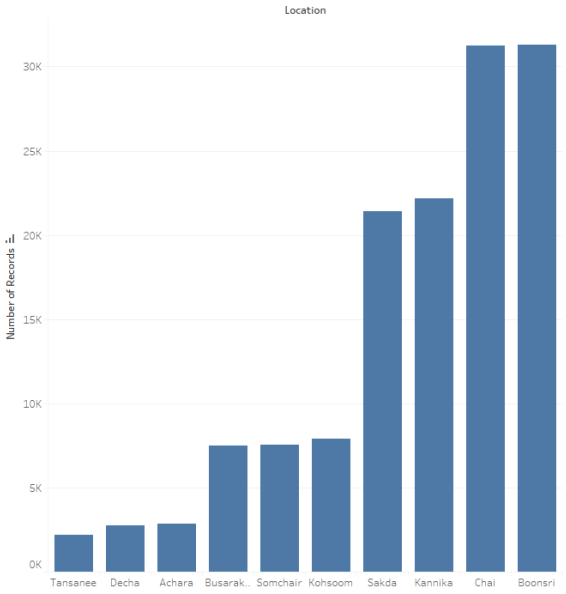
Even though the Methylosmoline data from Tansanee was not conclusive the historical average maps revealed that Tansanee has historically had high concentrations of 11 chemicals relative to the rest of the preserve. This could be an indication of past contamination. The 11 chemicals found at Tansanee are: Arsenic, bicarbonates, biochemical oxygen, chemical oxygen demand (Cr), chlorides, nitrates, nitrites, potassium, sodium, sulphates, and total nitrogen. This site should be monitored more closely in order to protect the Pipits in the preserve.



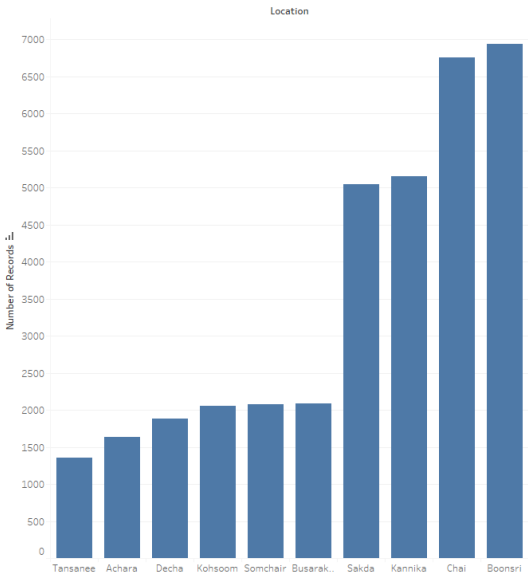
3. *After reviewing the data, do any of your findings cause particular concern for the Pipit or other wildlife? Would you suggest any changes in the sampling strategy to better understand the waterways situation in the Preserve? Your submission for this question should contain no more than 6 images and 500 words.*

Further visualization reveals that some locations have very few chemical measurements (all time shown at left and the last 5 years shown at right).

Measurements over all time

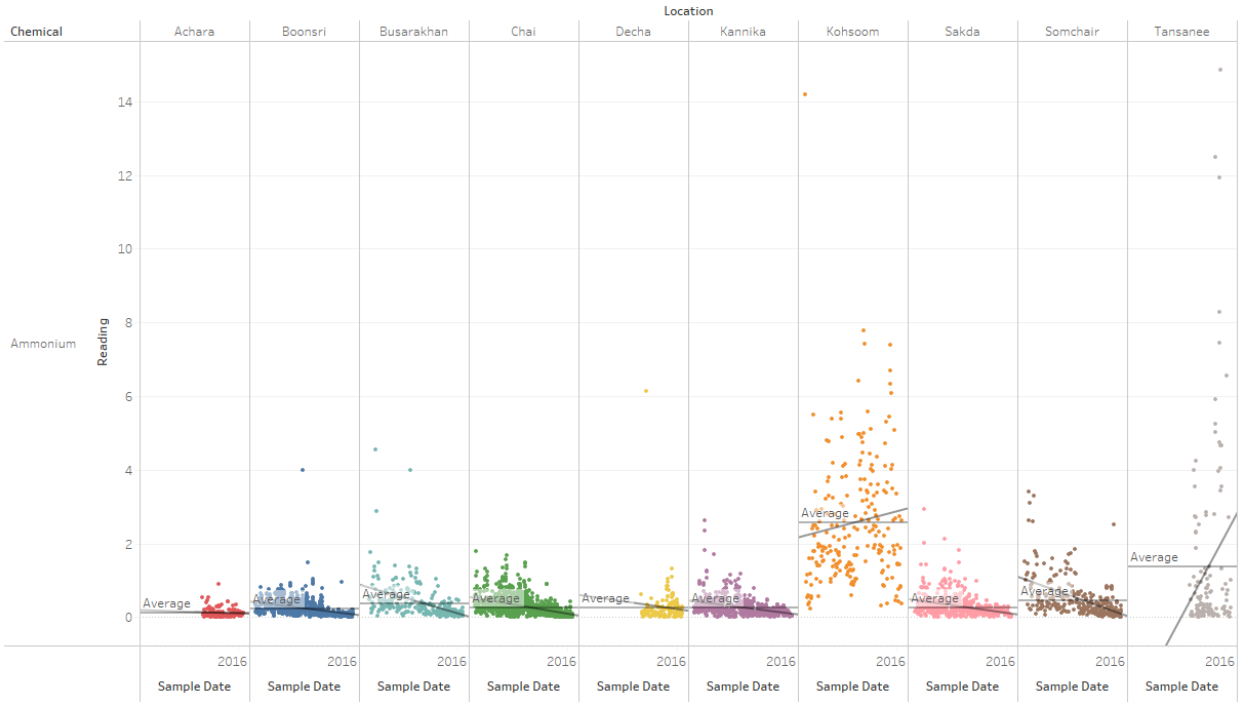


Measurements over the last 5 years



Based on this it is recommended to increase testing at Tansanee, Decha, and Achara. Luckily, these three sites are not close to the suspected dumping site. However, both Somchair and Kohsoom have relatively low levels of measurements as well. Both of these sites show evidence of contamination with Methylosmoline. Testing should be increased in these locations as well.

The ammonium measurements have historically been high at Kohsoom. This historical trend is rising whereas most other locations have declining ammonium readings. Tansanee is another exception that has elevated and increasing ammonium. This condition should be monitored more carefully.



Additional Reviewer Considerations

MC2 Questions and Approach:

- Was the submission able to sort out suspicious patterns (relevant to the scenario) from the false leads (which would not present a feasible threat to wildlife)? Were multiple variations in patterns identified?
- If there were hypotheses that remained unresolved, did the submission specify actions to be taken to resolve them?

MC2 Application of visual analytics:

- Did the team develop an innovative visual analytic tool? Alternatively, did they use an existing tool in an innovative way?
- Did visualizations enable the analysis process? Or did they merely illustrate conclusions? Did the submission rely more heavily on non-visual analytic approaches?
- Did their tool allow useful interactions?
- Did they use all the available data?
- Was the submission clear?