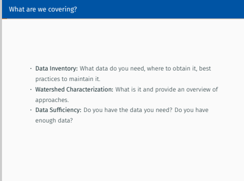
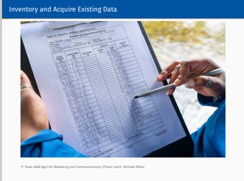
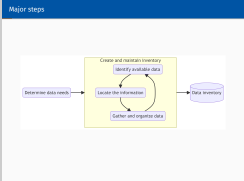


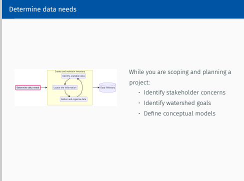
Before you get to this portion of any water quality project, do your homework, talk to stakeholders, know what water quality problem or research question you need to answer. This presentation assumes you have done some of that. This presentation focuses on the initial steps of finding and maintaining data, some methods for assessing water quality in watersheds, and evaluating if you have enough data to answer water quality related questions.



Here, we will primarily discuss how you acquire, document, and store your data.



A data inventory is how you store, maintain, and update your data in a structured way. Here we will break it down into four major steps and provide general advice for storing your data.



Your project scoping period helps determine data needs. A lot of this is established by watershed planning processes. What questions do you need to answer? For example, you may not have a good grasp on what the general water quality is or maybe you are trying to track the source(s) of the impairment. Or you know the problem but don't know what is needed to improve water quality. Finally, maybe you just need to evaluate if progress is being made. This will help you define the required models or approaches, and define the data oyu need to gather or collect.

Identify available data

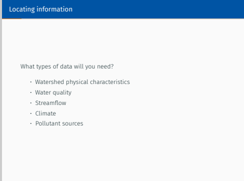
- Use existing data if/when available!
- Assess the quality/reliability of that data.
- Talk to organizations that have conducted projects in the area.

Your second step is to identify sources of available data. Resources are limited, use existing data when available assuming the quality matches your needs. Chances are, there are organizations, agencies, researchers in the area that have projects in the area. Contact them!

Locating information

What types of data will you need?

Can anyone tell me some general types of data you might want for a water quality assessment or watershed characeterization project



Besides water quality data, you will want information on waterbodies and the physical contributing watershed. For most projects you will want to have at least a basic grasp of potential point and non-point sources of pollution. If you are doing any modeling work, you will probably need more detailed climate, streamflow, and other data such as soil types, land use. land cover, crop types, and other model specific data.

Locating information

Watershed physical characteristics

- Watershed boundaries
- Hydrographic/Topographic
- Soil
- Vegetation, wildlife, etc.

When I talk about watershed physical characteristics, we can think about things like watershed boundaries and what is within those boundaries such as land cover, elevation, vegetation, flowlines, waterbodies and so on.

Locating information

Watershed physical characteristics	• USGS NHD, NHD HR: www.usgs.gov/national-hydrography/access-national-hydrography-products • NLCD, Imperviousness, Tree Canopy, etc.: www.mrlc.gov/ • NRCS Soil Survey:websoilsurvey.nrcs.usda.gov/app/

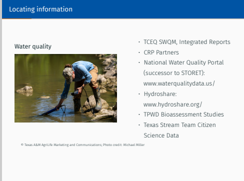
Your primary sources of information will be in the USGS national hydrographic dataset (NHD) or NHD high res. This includes raster and vector data defining natural and man-made flow lines, waterbodies, and can be used to delineate watershed boundaries. There is a pre-defined watershed boundary dataset as well that may or may not be suitable for your project area. The multi-resolution land characteristics consortium or MRLC is a group of federal agencies that provide data for land cover, tree canopy, imperviousness over the US. The USDA NRCS also provides data on soil types with numerous associated characteristics. NRCS is also a good source for a cropland data layer that provides info on the types of crops grown on agricultural lands nationwide.

Locating information

Water quality

- Routine water quality
- Stormwater/runoff monitoring
- Storm/flow biased water quality
- Benthic and other stream surveys

You will almost certainly need water quality data for whatever parameters you are interested in. This might be ambient or routine water quality data, stormwater data, biased water quality data that targets storm or high flow events, or data on other indicators of water quality health like benthic surveys, biological surveys, HAB data, and so on.



Besides TCEQ SWQM dataset, there are a number of sources of water quality data. Your local Clean Rivers Program Partner or municipal stormwater/watershed utility may have localized study data such as runoff monitoring. The federal government provides a portal for water quality data which combines state program collected data with various federal agency collected data. Hydroshare is a primarily academic data repository that includes monitoring as well as modelling data. Parks and Wildlife has plenty of data used in bioassessment and fisheries survey studies. And citizen science data is available and particularlly well suited for identifiy potentially problemenatic areas that might not be captured in the SWQM dataset.

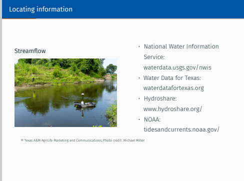
Locating information

Streamflow

© Texas A&M AgriLife Marketing and Communications; Photo credit: Michael Miller

- Daily/instantaneous streamflow
- Inundation/tidal data

I can't emphasize enough how important streamflow data is in a water quality assessment. We are often missing that data, but please think about collecting that data when designing projects. I will discuss that in my other presentation. Luckily USGS and their partners have an extensive network of streamflow gages. But depending on your study site, you might want to incorporate things like resevoir level, flood innudnation, tidal flows, or other hydrologic data.



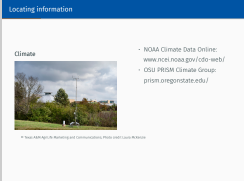
The USGS NWIS site is your primary source of data for streamflows. Texas Water Development Board provides data through Water Data For Texas on resevoir levels, groundwater levels, coastal freshwater flows, and some very useful continuous estuary water data. Hydroshare provides both water quality and hydrology data from academic research projects. NOAA is a valuable resources for tidal, wind, and salinity data used in coastal water quality projects.

Locating information

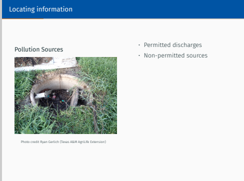
Climate

- Precipitation
- Temperature/solar radiation
- Wind direction/speed

Climatalogical data is frequently used in modeling projects but provides valuable reference for stream runoff conditions. Here we are primarily concerned with Precipitation, temperature and/or solar radiation (important components of evapotranspiration) which influences how much water goes from the soil through plants and ultimately to the atmosphere) and wind direction and speed.

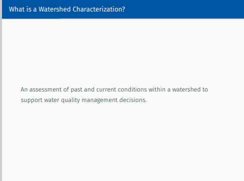


NOAA's Climate Data Online provides climatalogical records from airports and other official sites. Data type and length will vary substantial depending where you are. Oregon State Prism provides modelled climate data across the U.S. This isn't very useful for daily data, but great for monthly or annual averages and deviations from normal.



You will want to have at least a general idea of possible pollutant sources in your watershed. Permitted sources are relatively easy to discover since they are, well, permitted. Non-permitted can be a little bit more difficult. Here is a typically rural rource of pollution, what is it and how do you think you can find these?

This is a good time to get up and strech if you need. Before I jump into the next section of this presentation, are there any questions?



We can consider a watershed characterization several different things. Here I am going to discuss how we assess past and current water quality condtions to support water quality management descisions. But we can also consider characterizations things like spatial or temporal analysis to summarize and quantify all the data we just discussed to provide a snapshot of the current conditions in a watershed, and provide insight to problematic areas, periods of time, or sources of pollution.

What is a Watershed Characterization?

Can have different objectives:

- Is there a (or where is the) water quality problem?
- Compare and define current and desired conditions.
- Estimate allowable loadings for point source permits or TMDLs
- Evaluate and recommend best management practices
- Assess trends or changes in water quality

(not an exhaustive list...)

There are many reasons we might want to conduct monitoring and data assessments. First of course, we might be interested in if there is a water quality problem. This is basically what the Clean Rivers Program and TCEQ Assessments is tasked with on a large spatial scale. If we are doing watershed planning, we want to compare current and desired conditions, evaluate or recommend BMPs to acheive desired conditions or acheive reductions outlined in TMDLs. Finally, we may want to report or assess trends in water quality as we implement practices in watersheds. All of this can be done with field data, experiments, statistical analysis, and/or models.

DATA INVENTORY, WATERSHED CHARACTERIZATION & SUFFICIENT DATA

└─How do we do it?

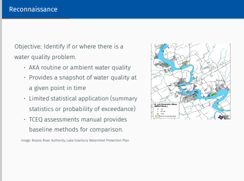
How do we do it?

Depends on our objectives!

Objectives determine the approach and data requirements (monitoring requirements).

There are many ways to accomplish this, and they are all determined by your primary objective that we just saw in the previous slide.

Here are a couple of approaches we can take. Reconnaissance study, plot studies, watershed studies. One thing I won't discuss today is using watershed models like SWAT or HSPF to conduct a watersehd assessment or scenario planning. Lucas puts togetehr a workshop with folks from AgriLife Blacklands Center that provide that training.



A reconnaissance study provides a snapshot of current water quality across a watershed using one or more sites. You can think of the 303d and clean rivers monitoring programs as state-wide or river basin reconnnaissance studies. They help identify places of concern but don't lend themselves to extensive analysis beyond summary statstics or exceedance probabilities. Here we have a map from the Lake Granbury Watershed Protection plan, this map uses ambient data to highlight one spot that is problematic, and helps focus potential implementation efforts.

Reconnaissance

Advantages:

- You might already have enough data (CRP monitoring)
- Large spatial coverage
- Useful to pinpoint problematic areas

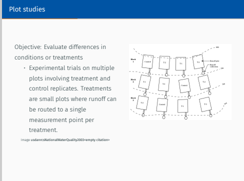
The big advantage of reconnaissance type studies are that it typically uses readily available data and we can quickly pick out potentail problem areas.

Reconnaissance

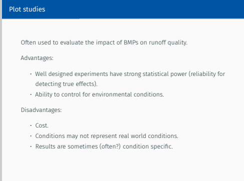
Disadvantages:

- Biased towards baseflows
- Possibly low sampling frequency in existing data
- Sampling locations are biased

Data that goes into recommaissance is typically routine monitoring data collected under a range of conditions but is usually biased towards baseflow so is inadequate for assessing loadings or if concerns are primarily non-point source. Routine data is typically collected quarterly to monthly. We miss high flow events or might have difficulty capturing seasonal or short term water quality patterns. Sampling locations are often biased towards bridge crossings or public access sites. What are some reasons biased sampling locations are problematic? Road crossings are heavily influenced by runoff, targets for illegal dumping, can be hard to assess spatial variability if sites are far apart.



Plot studies are experimental trials involving field plots, mesocosms, or some type of semi-controlled environment. They are setup to include control and treatment conditions, typically with replicates to provide stronger statistical power. Here is basic example showing an experiment with a contol and three different treatments. This is replicated across three blocks.



We often use this kind of experiment to evaluate how practices work in a given condition, watershed, or region. When designed correctly, analysis is straightforward and can account for environemental variability. However, it can potentially be a ton of work, even large plots are the same scale as large ag practices when implemented, and the results are not always transferrable.

Plot studies

Photo credit: Jason Gerlich, TWRI

Here is Lucas in his native habitat, in the process of installing automatted runoff samplers on agricultural BMP plots. Don't underestimate the work to get these experiment up and running. And even though these are huge plots, they are miniscule compare to what an ag producer might actually implement. The scaling of BMP results is an ongoing water quality research challenge.

Watershed studies

Objective: Evaluate Best Management Practices (BMPs), assess trends.

Multiple potential study designs:

Design	Causal factors
Single watershed - before/after	BMPs, Climate
Single watershed - upstream/downstream	BMP, within watershed factors
Paired watershed - BACI	BMP

Watershed studies are also called natural experiements. We use observational data with different types of study designs. These are generally classified as single or paired watershed studies. Setting up studies this way allows us to evaluate variation over time, but the experimental design is important if you need to seperate climate or other watershed factors from BMP implementation when looking for what is effecting water quality change.

Single watershed

Objective: Evaluate conditions before and after implementation, or trend analysis.

- Single monitoring site at the outlet of a watershed.
- Relatively inexpensive.
- Difficult to separate effects due to changes climate or other watershed changes.

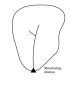
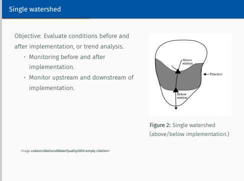


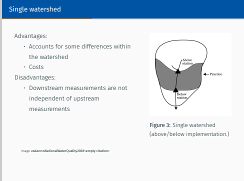
Figure 1: Single watershed (before/after implementation.)

Image usdanrcsNationalWaterQuality2003<empty citation>

Single watershed designs are simple, cheap, and the least logistically challenging. If you observe pre- and post- implementation you can evaluate potential water quality changes, but you generally cannot attribute if the cause was from implementation of BMPs, or changes in climate/runoff.



If you are evaluating the impact of a singular project, monitor before, after and upstream and downstream of the project. This allows you to mroe easily seperate the effect of within watershed changes.



This approach is more costly, but may reduce stakeholder uncertaintity depending on your objective. One disadvantge is that the upstream and downstream measurements are not independent, you will still expect the value of the downstream measurement to be somehow related to the upstream measurement.

Paired watershed

Objective: Evaluate conditions before and after implementation, or trend analysis.

- Monitor before and after in calibration and treatment watersheds
- Two or more watersheds
- One watershed always serves as a control

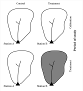


Figure 4: Paired watersheds; before-after control-impact (BACI)

Image usdanrcsNationalWaterQuality2003<empty citation>

For natural experiments, the paired watershed before-after and control-impact study is the ideal approach. Here you have one reference or control watershed and one or more treatment watersheds. They are monitored before and after treatments.

Paired watershed

Advantages:

- Can be statistically powerful.
- Can evaluate multiple treatments.
- Helps control for environmentally induced variation not associated with treatment.
- Statistical independence between treatment/pre-treatment measurements.
- Regionally relevant results.

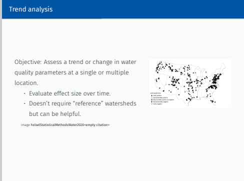
These experiments are statistcally similar to plot experiments, but less controlled. We can evaluate multiple treatments if we have enough watersheds. All measurements are independent. We are able to seperately attribute changes in water quality to BMP implementation or natural/climatic variation.

Paired watershed

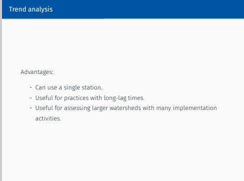
Disadvantage:

- Requires long study period.
- Lag effects can be gradual masking difference with control (**mealsLagTimeWater2010**).
- Can be hard to find two watersheds that won't have any other disturbances for the entire study period.
- Logistics.

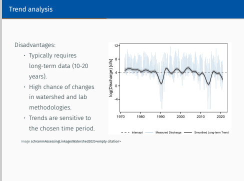
The big disadvantage is you may need a relatively long study period, 3-4 years at least. If there are lag effects (such as expected with sediment bound pollutants) you may not capture the variation. Of course finding two or more watersheds nearby, that won't have major changes for an entire study period is difficult. Expect and document interuptions. Finally these studies can be very logistically challenging.



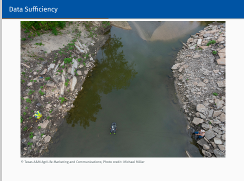
Trend analysis is a popular approach for the routine stakeholder question, is water quality getting better or worse? You often see a map similar to this one from Denis Helsel, formerly with USGS, with stations that have significant increases point up, significant decreases pointing down and stations without a likely trend represented as circles. We can use single watersheds or multiple watresheds with this approach. If you have a "minimally distrubed" reference watershed, it is useful to include.



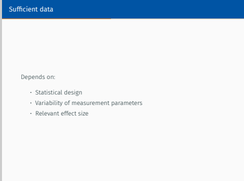
Big advantages of this approach are the interpretable answers for stakeholders, typically reported as percent change. You can use a single watershed outlet, and you are not neccesarily attributing change to a particular single BMP or group of BMPs. There are also methods for seperating the impact of changes in flow, if you have flow data.



The biggest disadvantage is that trends aren't really that useful unless you have multiple years of data, often 10 or more years are reccomended. Over time there will be changes in both the watershed, possible lab methods or even measured parameters. Depending on the method you choose, trends can be highly sensitve to starting and ending values, and the time period you choose. In this graph you can see the non-linear smoothed trend in discharge with periods of time that we see clear negative and clear positive trends. If we fit a liner trend test to ten years of data, there is no trend, but 20 years of data show a negative trend. Defining the relevant period of time is important.



We are almost done, I have a couple of slides left on data sufficiency, but we aren't going into too much detail. Before I start, any questions?



How do you know if you have enough data to meet your objective? This is going to depend largely on the particular statistical design used, the variability of the water quality paramater (high variability requires more data), and the relevant effect size (percent change or difference in water quality for example) that you need to reliably identify. Large changes or effects can be identified with less data, but small changes need more data.

Sufficient data

Things to consider:

- Adequate number of sites
- Adequate number of samples
- Frequency of samples
- Monitoring duration

In addition to the number of samples, consider if you have enough sites to identify spatial patterns or changes, or influences from known sources, are you sampling frequently enough to capture seasonal or (maybe daily) variations? Are you monitoring long enough to capture long term trends and not just short term variation.

Sufficient data

Available tools:

- TCEQ assessment criteria (for ambient water quality)
- Statistical power analysis: **helselStatisticalMethodsWater2020<empty citation>** Chapter 13

We could have a whole data on how much data you need, but I'm going to point you to two different resources. First, TCEQ's assessment procedures are very good for general characterization or reconasiance type approaches. They provide clear guidance for duration and sample size needed to make statisically based decisions according to the state's numeric criteria. If you are using a watershed or plot trial, chapter 13 of the USGS stastical methods in water reources handbook, is excellent. This discusses how to do something called power analysis to estimate the statstical power of your approach based on the number of samples and expected variance. Stastical power is essentially how likely your stasstical test is to correctly idenfity an effect when there actually is an effect. We typically aim for a power of 80% but that is not a hard and fast rule.

Here is an example of a power analysis I conducted for detecting 10-year trends in E. coli concentrations. If you pick one of these plots, the x-axis shows the number of annual samples needed to achieve the desired stastical power on the y-axis. Each line represents a different effect size, so the light yellow lines represent if we wanted to detect a 10% reduction in bacteria, even at 12 samples a year, our power is only around 30-40%. But an 80% difference can be detected with only 4 samples a year with nearly 100% power. There are lots of methods for doing power analysis depending on your goals, I don't see it done in our field very often, but I highly reccomend it when you are designing a project. Take a look at the resources at the end of this presentation or reach out to discuss with me.