NETLAKE Guidelines for automated monitoring system development

011 Measurement Frequency

Objective

The purpose of this fact sheet is to provide some advice on the frequency at which data should be measured and saved from automated monitoring stations (AMS).

Considerations

AMS used for lake studies can measure meteorological parameters, water temperature, water chemistry, and indicators of biological processes. All of these vary along a continuum of time scales, which can potentially be studied. Here we provide some guidance on the measurement frequency and storage frequency, based on our collective experience in aquatic monitoring.

Fundamental Considerations

The place to start when designing an automated monitoring strategy is to first define the question(s) that one wishes to answer with data from the AMS, the processes that affect that question, and the time scale(s) over which these processes are expected to vary. The later should set the lower limit on the frequency at which data is collected. As an example, water temperature is a fundamental property of lakes that can affect many other processes, and can therefore be measured at a variety of frequencies. To obtain information on the lake heat budgets or patterns of thermal stratification daily data collection could be sufficient, while to obtain information on levels of water turbulence, measurements need to be made at multiple times per second. While the highest frequency measurements can always be aggregated to longer frequencies, there are costs and tradeoffs that must be made when collecting data at higher frequency. Therefore, measuring at the highest frequency is not always the obvious choice.

Planning for the future

One of the great advantages of AMS is that they provide long-term consistently measured data records. These can be used to evaluate such things as long-term changes in climate or changes in lake inputs. All uses of the data may not be evident when a monitoring station is first established. Future, perhaps unanticipated, use of the data might benefit from storing data at high frequencies. Therefore, despite the sound advice above, there are good reasons to measure at higher frequencies than needed if the costs are not prohibitive. Using the example of Erken Laboratory water temperature, if one was measuring using a temperature sensor chain of 20 sensors and storing data as 4 byte numbers, then one year of daily measurement would only require approximately 30 KB of memory, well below the storage capacity of modern data loggers. When Lake Erken's automated monitoring program first began in 1986, state of the art data loggers had 16 KB of memory, making the storage of too frequent data a real issue. Today however, data storage is typically multiple MB, and it is even possible to store up to 16 GB, a 6 order of magnitude increase in storage capacity. Storing those 20 x 4 byte temperature measurements every 5 min would use about 8.4 MB of storage over an entire year. In most cases there are few reasons to not store the temperature data at frequencies in the minute range: Storage is not limiting; costs are not prohibitive; automated quality control can be just as easily preformed on the higher frequency data. Only the lower frequency data of interest need to be analyzed, but the higher frequency data is always available for future use.

Measurement vs Storage

The frequency of measurements is not necessarily the same as that of data storage. With many data loggers it is possible to measure frequently, and store measurement averages (and/or other statistics) at lower frequencies. In most cases this is a good idea, since averaging high frequency measurements over a longer storage interval generally provides more representative data, but again logger processing should be linked to the questions being asked and the purpose of the monitoring program. The Campbell data loggers in use at Lake Erken measure every second and store mean data values at 5 min, 60 min and 24 hour intervals. In the early days of the monitoring program only the hourly and daily data storage were practical. The 5 min storage interval was added latter as logger memory increased, but the hourly and daily means were retained for the sake of consistency. Today one can consider if it is worth internally processing data in the logger at all. Some advocate storing data at as high a frequency as possible, and preforming all post processing separately from the logger. This has the advantage of providing maximum flexibility in data processing, but can increase data storage and transmission costs. No matter what the choice of internal logger processing it is important that the processing method, or lack of it, is well documented in station metadata descriptions. This information can be very helpful when processing and comparing data from stations run by different institutions - such as data collected across the NETLAKE network.

Limitations on the frequency of measurement and storage

Finally there are properties of the measurement system which will set upper limits on the frequency of measurement. One of the most fundamental properties is the response time of the sensor itself. There is no value in measuring a sensor at a rate that is faster than it can be expected to respond to a change in the environment. Optical sensors such as fluorometers and under water light sensors have nearly instantaneous responses. The response of a temperature sensor is also nearly instantaneous from the electrical point of view, but in practice the response will be slower due to the thermal mass of the sensor itself: large robust sensors will take longer to change temperature than smaller more fragile sensors. Likewise, sensors that rely on the measurement of gradients across a membrane will be dependent on the rates of equilibration across the gradient. Sensor response time will also be related to the steepness of gradient through which the sensor is passing. Luckily in most cases the response time of aquatic sensors is generally faster than the frequency that would need to be measured The response time of most commercial sensors is well documented by their manufactures, and this information should considered when setting the upper limit on measurement frequency. The second limitation on measurement frequency is with the data logger.

While theoretically there are limitations on the frequency of measurement imposed by the data logger electronics, this will rarely be a practical limitation. However, data logger storage capacity, in conjunction with data communications and data transmission capacity can limit the frequency of measurement, and/or the frequency of storage. Going back to the Lake Erken temperature example above of 20 sensors measured at a 5min frequency, but in this case storing the data as higher resolution 8 byte numbers (the equivalent of a single precision floating point value) would require 38.4 KB of data logger storage per day. This is no problem if the data can be collected at daily or weekly intervals, but could become an issue if data can only be collected on a monthly basis. The second potential limitation is data transmission. Knowing that there are 8 bits in a byte the above daily data collection (369 Kbits) would require a data transmission time of 4.5 min/week using a GSM modem with a transmission speed of 9600 Kbits s^{-1} This would not be a prohibitively long transmission time, but GSM connections are not free, and this does illustrate that data transmission rates, can in some cases set limits on the frequency of data storage. In general given the large storage capacity (MB to GB) of modern data loggers it is less likely that on-logger storage will limit measurement storage frequency, and more likely data transmission rates could act as a bottleneck that limits the frequency of data storage.

Likely Problems

- Not properly matching data measurement and collection frequencies to the purpose of the monitoring project.
- Not collecting data that could add to the value of the long term monitoring program at little additional cost.
- Collecting so much data that the project staff are not able to properly analyze and use it

Suggested citation: Pierson, D., de Eyto, E., Laas, A., and Jennings, E. 2016. Measurement frequency (Factsheet 011). In: Laas, A., de Eyto, E., Pierson, D. and Jennings, E. (Eds.) *NETLAKE Guidelines for automatic monitoring station development*. Technical report. NETLAKE COST Action ES1201. pp 44-46. http://eprints.dkit.ie/id/eprint/518

Acknowledgement
This factsheet is based upon work from the NETLAKE COST Action (ES1201), supported by COST (European Cooperation in Science and Technology).