Water quality monitoring of Lake Cluster in Pokhara Valley, Nepal with satellite images and machine learning

Out of 70% of the water covering Earth's surface only 0.1% is freshwater (Carpenter et al. 2011). This small portion of freshwater, is facing major stress from the growing population. While developed nations are shifting their interests toward sustainable use of water resources, developing nations are yet to realize the need for conservation (Capps et al. 2016). Thus, water bodies are facing unprecedented stress and insufficient effort has been undertaken for its revitalization. One such example is the lake cluster in Pokhara valley of Nepal, consisting of nine lakes, which has been enlisted as a Ramsar category wetland.

These lakes provide ecosystem services like irrigation, water supply, habitat for organisms, and aesthetic beauty for tourism and recreation. Lake Phewa, popular among all, for tourist destination is suffering from lake encroachment, invasion from alien species, etc. (Paudel et al. 2017). Begnas and Rupa, two more lakes are also facing the problem of rapid eutrophication, increase in organic pollutants from agricultural runoffs has been identified as serious threat (Pant et al. 2019). Remaining six lakes of the cluster are not getting much attention due to their smaller size and remote location. Thus, there is a dire need for conservation before the damages are irreversible.

Conservation starts with monitoring. Monitoring is the regular measurement of key characteristics to note the state of an ecosystem and its change over time. This scientific information provides the basis for making wise decisions on management of the local environment (Karpatne et al. 2016). Lake monitoring with traditional methods is a lengthy process. Lack of skilled manpower, budget constraint, difficult terrains, etc. adds more hurdles.

The monitoring programs done previously are also sporadic. These obstacles can be avoided with the use of satellite images.

Getting remote sensing data of an area from any part of the world has never been easy as today. All thanks to the satellites constantly orbiting and acquiring data. Information from earth observation data is like diamonds ready to be mined. Such information can be on various issues and one such issue is water resource conservation.

With the increase in spatial and temporal resolution of the satellite images (Toming et al. 2016), lake monitoring can be done regularly with more spatial detail. Additionally, the satellite provides images with different bands of wavelengths ranging from visible to infrared. Information across various bands can be used to estimate parameters that are crucial to determine the health of the water body. These advancements in space technologies have enabled users for more regular and sustainable sampling.

Fetching information from the satellite image starts with the step of relating the satellite image with the in situ water quality parameters of the lake. A relationship between the surface reflectance values across various bands of a satellite image with the water quality parameter is needed, that's where the power of machine learning can be leveraged. Machine learning specifically supervised learning, trains itself from labeled data to get a relationship between dependent and independent variables (Singh et al. 2016). For the lake monitoring, the dependent variables are water quality parameters (most often water temperature, chlorophyll-a, suspended solids, etc.), while the independent variables are the reflectance value of visible bands, near-infrared, infrared, and thermal bands. Machine learning can easily find the pattern that human mind couldn't see. For training and testing the model, field data of the same day of the satellite

overpass is needed. Once a reliable model is developed with an acceptable level of accuracy, the model can be used to estimate the water quality parameters without further need for sampling. Regularly captured satellite images enable frequent monitoring of the lake clusters.

While satellite data are readily available through various online platforms, the acquisition of a colossal amount of data and processing in a local personal computer can be a bottleneck. The personal computers normally used for other simple computing purposes may take too much of time to go through all the images and analyze them. That's where cloud computing comes into play. Various cloud-based platforms are available in which users process the information at a much faster rate. Information generation from this technology becomes much easier and faster. Additionally, cloud storage provides the additional benefit of not requiring data storage on a local computer. This makes the processing of a large amount of data much easier and faster, and therefore suitable for satellite image processing.

The information generated will be of great benefit only when they are disseminated among the stakeholders. Scientific findings are generally published in peer-reviewed literature, which people from non-research backgrounds do not have an access to. Thus, web GIS can be suitable platform for sharing information with a wide range of audiences. Just with a simple smartphone and internet connection, one can access the data. Also, time series data and basic statistics can be made available which can support decision-making processes. On top of that, data of all the lakes of the cluster can be made available through a single platform and can be updated frequently with minimal effort. This would be suitable for a wide range of audiences from people working at policy levels, conservation sectors, and research institutions to people from local communities who are using the lake for making a living.

In conclusion, water bodies from all over the world need to be conserved and conservation starts with monitoring. While developed countries have started various lake monitoring programs, developing countries are facing barriers on the implementation of monitoring programs. One such developing country Nepal is well known for its water resources and one of them is the lake cluster of Pokhara valley. Lack of resources and difficult terrain make it challenging to implement continuous monitoring programs with traditional methods. Thus, combining various new technologies and leveraging their benefits for water conservation is needed, not as a replacement but as a complement to traditional water quality monitoring methods to provide a scientific basis for freshwater conservation.

Bibliography

Capps, Krista A, Catherine N Bentsen, and Alonso Ramírez. 2016. "Poverty, urbanization, and environmental degradation: urban streams in the developing world." *Freshwater Science* no. 35 (1):429-435.

Carpenter, Stephen R, Emily H Stanley, and M Jake Vander Zanden. 2011. "State of the world's freshwater ecosystems: physical, chemical, and biological changes." *Annual review of Environment and Resources* no. 36:75-99.

Karpatne, Anuj, Ankush Khandelwal, Xi Chen, Varun Mithal, James Faghmous, and Vipin Kumar. 2016. "Global monitoring of inland water dynamics: State-of-the-art, challenges, and opportunities." *Computational sustainability*:121-147.

Pant, Ramesh R, Khadka Bdr Pal, Nanda L Adhikari, Subash Adhikari, and Akkal D Mishra. 2019. "Water Quality Assessment of Begnas and Rupa Lakes, Lesser Himalaya Pokhara, Nepal." *Journal of the Institute of Engineering* no. 15 (2):113-122.

Paudel, Namita, Subash Adhikari, and Ganesh Paudel. 2017. "Ramsar lakes in the foothills of Himalaya, Pokhara-Lekhnath, Nepal: An overview." *Janapriya Journal of Interdisciplinary Studies* no. 6:134-147.

Singh, Amanpreet, Narina Thakur, and Aakanksha Sharma. 2016. A review of supervised machine learning algorithms. Paper read at 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom).

Toming, Kaire, Tiit Kutser, Alo Laas, Margot Sepp, Birgot Paavel, and Tiina Nõges. 2016. "First experiences in mapping lake water quality parameters with Sentinel-2 MSI imagery." *Remote Sensing* no. 8 (8):640.