

Filling Gaps in the Cloud: Ensuring dataset integrity of CHS's water level network

Arfeen Khaleel¹, Sheward Mike¹, Ferrier Jake¹, MacAulay Phillip¹, Michael Rollings¹

¹Canadian Hydrographic Service, Department of Fisheries and Oceans Canada

Khaleel.Arfeen@dfo-mpo.gc.ca

Mike.Sheward@dfo-mpo.gc.ca

The Canadian Hydrographic Service (CHS) recently operationalized its new sensor-to-client water levels and currents data collection and delivery system. This includes the cloud-based Integrated Water Level System (IWLS) to host Canadian Hydrographic Service (CHS) data in a central database. This data is collected and distributed to the IWLS by the CHS's Permanent Water Level Network (PWLN). This network consists of 97 tidal and non-tidal stations spanning the country and collecting environmental data including water level, current and metadata. This system as a whole, provides nationally standardized data and metadata collection, quality control and delivery.

The IWLS houses: observations pushed in real time from PWLN stations, constituent predictions produced on-demand and several types of modeled water level forecasts. While many PWLN stations provide observational data via redundant communications, unfortunately, not all data is received by the IWLS in real-time. Network failures or hardware malfunctions occasionally result in missed messages and gaps in the datasets.

Resolving the data gaps will achieve complete and reliable real-time available data. These data gaps must be detected and filled as quickly as possible in order to provide stakeholders with the best possible dataset. This presentation will explore the design and strategies used by the CHS's new IWLS Gap Fill System. Discussion will include: gap identification, the multi-tiered data recovery approach, data quality metrics, systems architecture, downstream implications and potential future work.