Methods for Assessing Historical Water Level Data Quality for the Texas Coast

N. Cail, R. Procious, C. Baca, P. E. Tissot, N. Durham, J. Womack, L. Huynh, and D. Martin

Texas A&M University-Corpus Christi, Conrad Blucher Institute, Corpus Christi, Texas

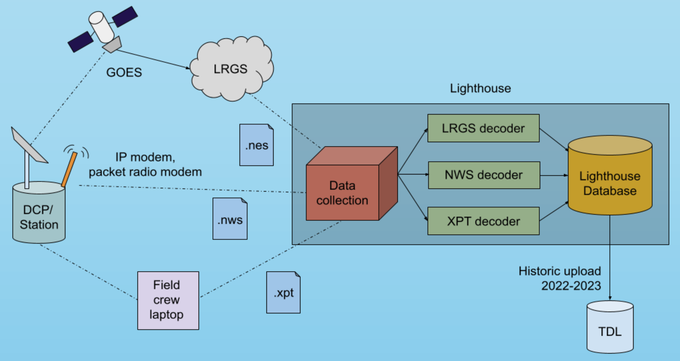
ai2es.org

Problem Statement

High quality water level data are important for many uses. This research shares a methodology and initial results assessing the quality of historical data from the Texas Coastal Ocean Observation Network. The method is based on comparing the data of four NWLON stations processed by both NOAA and TCOON systems.

Background

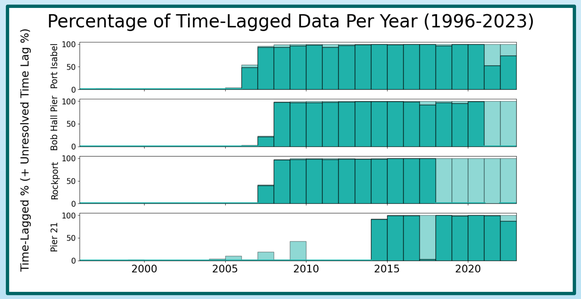
- High quality water level data is needed for sea level rise assessment, analysis of past extreme events, training AI to develop QA/QC methods, predict water levels and flooding.
- TCOON has a combined 90 current and historical tide stations along the Texas coast with data starting around 1993 [1].
- Historical summary data were reviewed by NOAA, including tidal datums. This research focuses on high frequency data.
- Water levels are measured/archived in 6-minute increments with each value equal to the average of 181 measurements at 1 Hz.
- Measurements are microwave radar, sonar, or pressure sensors.
- Data is transmitted by:
  - NOAA's NESDIS Data Collection System (DCS) satellites (GOES) every 6 minutes. The CBI downloads it from NOAA's DCS Administration and Data Distribution System (DADDS).
  - Direct connection to the station via cellular network modem (previously local radio networks).
- Data is decoded with CBI software dependent on the communication path. The software changed to adapt to evolving communication protocols and operational needs.
- This research data was retrieved as publicly available in the TDL [1]. It is raw DCP data, measured relative to station datum, without spike removal or gap-filling processing.

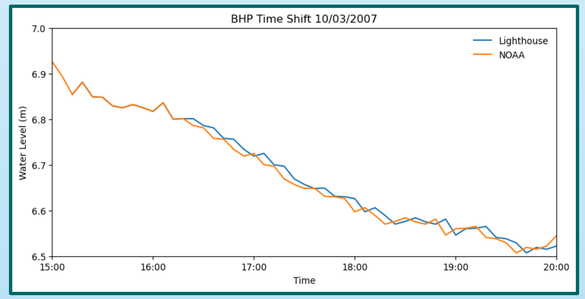


Water level data collection schematic. *DCP (Data Collection Platform), GOES (Geostationary Operational Environmental Satellite), LRGS (Local Readout Ground Station), .nes (NESDIS LRGS messages), .nws (national ocean service formatted messages), .xpt (Sutron DCP manual downloaded messages), TDL (Texas Digital Library).*

Methodology

- Data from four NWLON stations [2] were selected as the standard to evaluate the past processing of data through Lighthouse: Bob Hall Pier (8775870), Pier 21 (8771450), Port Isabel (8779770), Rockport (8774770).
- Data from the same sensors had been concurrently processed by both systems. Differences came from different software and, at times, different communication paths. NOAA data [2] is considered as the 'true/accepted' values and differences are used to assess the historical lighthouse system using visual and custom algorithmic inspection of 1996-2023 data.
- Main component of the analysis code [3] is the temporal correction algorithm, which assumes the data has varying time-shifts and iteratively tests guess values of the shift. If time alignment with NOAA is possible, it adjusts the data accordingly.





Analysis Findings

- Discrepancies: temporal shifts (TSs), vertical offsets (VOs), gaps, flatlines, and spikes (all assumed to be independent issues).

Time Shifts

- All four stations begin having a 6-minute time lag (behind NOAA) at some point in the year range 1996 – 2023.
- The code attempts to discern a TS despite missing data (it is possible to consider intermittent missing data as time-shifted).
- Some data may have a TS *unresolvable* by our algorithm, preventing alignment of the data with NOAA's and determination the TS.

Vertical Offsets

- We correct TSs algorithmically, then analyze the duration and magnitude of VOs.
- Most of the VOs are less than 2 cm (typically acceptable for NOAA), They range from 1 mm – 121 mm (lasting at least 1 hour). When not filtering by duration, the range can be on the order of meters due to flatlines and spikes in the TCOON data.
- Many VOs are briefly interrupted by other VOs or correct values. The overall lifetime of VOs can be hours to years. The longest uninterrupted VO was 0.01 m lasting 300 days 23 hours 24 minutes at Pier 21 from 6/29/2022 8:12 PM to 4/26/2023 7:36 PM.
- The station with the most vertically-offset data is Port Isabel: 19.71 – 20.94%.

Discussion

Time lag hypothesis and correction

- Water level data are the result of different communication and decoding paths including the Local Readout Ground Station (LRGS) which decodes data transmitted from satellite.
- During 2006-2007, the LRGS format switched from an hourly message to every 6 minutes. The Lighthouse decoder was updated but was incorrectly extracting the sample time as the time of signal reception, causing some of the 6-minute lags.
- Data retrieval from Lighthouse defers to a hierarchy of communication paths to select which source to provide data from. The satellite data is lower in the hierarchy. The time lags identified for each station correlate to how much LRGS data is used.
- By coincidence, until 2014, the Pier 21 data satellite transmit time was close enough that the correct sample time could be extracted (calculating the floor of the transmit time to the nearest 6-minute time).

Solution

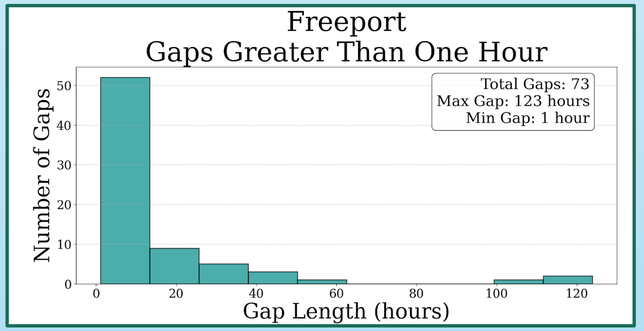
- Our team corrected the LRGS decoder for test years 2007 and 2008 for the 4 stations. The corrected data was merged with the rest of the time series and evaluated and was found to be without any time lags. The remaining of the four test stations data is being evaluated. With confirmation, the correction will be implemented for all 90 TCOON stations historical data.

Conclusions & Future Work

- 28 years of historical water level data was reviewed, and methods developed to assess quality and identify potential problems.
- 6-min time shifts and their cause were identified and corrected.
- The team is continuing to analyze the data for vertical offsets.
- Automated gap filling and anomaly detection methods following IOOS standards [4] and prior work [5] are being tested.

Gaps and back-up water levels

- For each TCOON station, an analysis of the number and distribution of the gaps was also conducted. These results will guide the gap filling efforts.
- Additionally, the stations and respective time-period for which backup water level data existed was documented.



Nomenclature

TCOON: Texas Coastal Ocean Observation Network
NOAA: National Oceanic and Atmospheric Administration
NWLON: National Water Level Observation Network
TGLO: Texas General Land Office
NESDIS: National Environmental Satellite, Data, and Information Service

References

[1] *Texas Coastal Ocean Observation Network (TCOON).* Texas Digital Library data repository. https://hdl.handle.net/1969.6/89444
[2] NOAA. *Tides and Currents.* https://tidesandcurrents.noaa.gov/
[3] https://github.com/ncail/noaa-lighthouse-problem
[4] Bushnell, M., & Worthington, H. (2016). Manual for real-time quality control of water level data: a guide to quality control and quality assurance for water level observations.
[5] Tissot, P. E., Zhu, W. B., Duff, S., Rink, M., Rizzo, J., & Martin, D. (2014, September). Development, assessment and implementation of an automated gap filling method for tide stations with dual water level sensors. In 2014 Oceans-St. John's (pp. 1-10). IEEE.

Acknowledgements

This work is supported by the Texas General Land Office and the NSF AI Institute for Research on Trustworthy AI in Weather, Climate, and Coastal Oceanography (AI2ES), under National Science Foundation Grant No. RISE-2019758 for the latter support.