Axial Seamount Differential Uplift Analysis

Research Context

This project calculates differential uplift between two Bottom Pressure Recorders (NANO-BPRs) at Axial Seamount. The analysis subtracts depth at MJ03F (Central Caldera) from MJ03E (Eastern Caldera) to show relative vertical displacement.

Data Sources

Station	Location	Path
MJ03E	Eastern Caldera	`/home/jovyan/ooi/kdata/RS03ECAL-MJ03E-06-BOTPTA302-streamed-botpt_nano_sample_15s/`
MJ03F	Central Caldera	`/home/jovyan/ooi/kdata/RS03CCAL-MJ03F-05-BOTPTA301-streamed-botpt_nano_sample_15s/`

- **Format**: netCDF files
- **Time period**: 2015-01-01 to 2026-01-16

Technical Environment

- **Language**: Python 3.12
- **Key packages**: xarray, pandas, matplotlib
- **Working directory**: `/home/jovyan/repos/specKitScience/my-analysis_botpt/`

Pressure-to-Depth Conversion

```
depth_m = (pressure_psia - 14.7) * 0.670
```

Quality Control: Spike Removal

Spikes are removed using Median Absolute Deviation (MAD), which is more robust to outliers than standard deviation.

Algorithm

1. Calculate rolling median over 24-hour centered window
2. Compute absolute deviation from rolling median for each point
3. Calculate rolling MAD (median of absolute deviations)
4. Scale MAD by 1.4826 to approximate standard deviation (for normal distributions, std ≈ 1.4826 × MAD)
5. Flag points where deviation > threshold × scaled_MAD
6. Replace flagged points with NaN

Thresholds

Data	Threshold	Rationale
Individual station depth	5.0	Conservative; catches obvious sensor glitches
Differential signal	3.5	More aggressive; catches glitches that appear when one sensor spikes but not the other

Why MAD over Standard Deviation?

- Standard deviation is sensitive to extreme values—a single large spike inflates the std, making other spikes harder to detect
- MAD uses medians, so outliers have minimal influence on the threshold calculation

Outputs

All outputs in the `outputs/` directory:

Data Products (`outputs/data/`)

File	Description
`differential_uplift_hourly.parquet`	Hourly cleaned data (92,222 rows) with `depth_mj03e_m`, `depth_mj03f_m`, `differential_m`
`differential_uplift_daily.parquet`	Daily averaged data (4,034 rows) with same columns

Figures (`outputs/figures/`)

File	Description
`depth_mj03e.png`	Time series of Eastern Caldera depth (meters)
`depth_mj03f.png`	Time series of Central Caldera depth (meters)
`differential_uplift.png`	Time series of MJ03E - MJ03F depth difference, with red horizontal line marking 2015 high

Reproducible Notebook (`outputs/notebooks/`)

File	Description
`differential_uplift_analysis.ipynb`	Fully annotated Jupyter notebook mirroring the analysis
`environment.yml`	Conda environment specification
`README.md`	Quick start guide for reproducing the analysis

The notebook is self-contained and allows someone without access to Claude Code to clone the folder and reproduce the analysis from scratch.

Usage

```
import pandas as pd
bpr = pd.read_parquet('outputs/data/differential_uplift_daily.parquet')
```

```
other = pd.read_parquet('other_instrument.parquet')
merged = bpr.join(other, how='inner')
```

Notes

- Swap xarray dimensions: `ds.swap_dims({'obs': 'time'})`