

Question 2: Motivation of Selected Error Threshold

The 5% error threshold was selected because:

Data Quality:

- A small percentage of failures is expected due to -
- Points exactly on hexagon boundaries
- Minor coordinate system discrepancies
- Edge cases in the H3 grid

Operational Impact:

- 5% represents a balance between strict quality control and practical reality
- Higher rates might indicate serious issues with data or processing
- Lower thresholds might be too strict for real-world geospatial data

Industry Standards:

- Similar thresholds are common in geospatial ETL processes
- Allows for known issues while catching unexpected problems