Geographic Correlation Analysis
User's Guide

Christopher Stubbs

Analysis developed with Claude Code

January 2026

Contents

1 Overview

The `wwv_geographic_correlation.py` script analyzes correlations in signal fading and phase stability between geographically separated receivers observing the same WWV frequency simultaneously. This reveals the **spatial coherence scale** of ionospheric disturbances.

1.1 Key Questions Addressed

1. Are fades correlated between receivers at the same frequency?
2. Is phase stability correlated between receivers?
3. How do correlations depend on frequency (2.5, 5, 10 MHz, etc.)?

1.2 Scientific Background

Two KiwiSDR receivers separated by ∼30 km observe the same WWV transmissions. If ionospheric structures are large compared to the receiver separation, fades should be correlated. If structures are small (on the order of the separation distance), fades will be independent.

This spatial coherence information is critical for understanding:

- Whether adaptive ionospheric correction at one site would help nearby sites
- The scale of ionospheric turbulence structures
- How propagation effects decorrelate over distance

2 Installation and Requirements

2.1 Dependencies

```
pip install numpy matplotlib scipy
```

2.2 File Location

The script should be run from a directory containing KiwiSDR IQ recordings in the standard filename format (see Section 4).

3 Command Line Usage

3.1 Basic Syntax

```
python wwv_geographic_correlation.py [options] [timestamp]
```

Argument/Option	Description
`timestamp`	Optional. Specific timestamp to analyze (e.g., `20260101.1148`)
`--date DATE`	Analyze all pairs from a specific date (e.g., `20260101`)
`--list`	List available receiver pairs without analysis
`--no-plot`	Skip plot generation (text output only)

3.2 Arguments and Options

3.3 Usage Examples

```
# Analyze all available receiver pairs
python wwv_geographic_correlation.py

# Analyze a specific timestamp
python wwv_geographic_correlation.py 20260101.1148

# Analyze all pairs from a specific date
python wwv_geographic_correlation.py --date 20260101

# List available receiver pairs
python wwv_geographic_correlation.py --list

# Analyze without generating plots
python wwv_geographic_correlation.py --no-plot
```

4 Input File Requirements

4.1 Filename Format

The script expects KiwiSDR IQ WAV files with the standard naming convention:

```
YYYYMMDD.HHMM.RXID.proxy.kiwisdr.com.freqXXXXX.SRATE.wav
```

Example filenames:

```
20260101.1148.22350.proxy.kiwisdr.com.freq5000.20000.wav  (Cambridge)
20260101.1148.22463.proxy.kiwisdr.com.freq5000.20000.wav  (Sudbury)
```

4.2 Required File Structure

For geographic correlation analysis, you need:

- At least two receivers (different `RXID` values)
- Same timestamp (`YYYYMMDD.HHMM`)
- Same frequency (`freqXXXXX`)
- Stereo I/Q format WAV files

4.3 Complete Sets

A "complete set" requires both receivers to have recordings at the same frequencies. The script automatically identifies which timestamps have complete sets for analysis.

5 Receiver Configuration

5.1 Default Receiver Definitions

The script has built-in definitions for known receivers:

```
RECEIVER_INFO = {
    '22350': {'name': 'Cambridge', 'location': 'Cambridge, MA'},
    '22463': {'name': 'Sudbury', 'location': 'Sudbury, MA'},
}
BASELINE_KM = 32  # Approximate distance between receivers
```

5.2 Adding New Receivers

To add new receiver locations, edit the `RECEIVER_INFO` dictionary at the top of the script. Also update `BASELINE_KM` if the separation distance differs.

6 Configuration Parameters

Key configurable parameters at the top of the script:

Parameter	Default	Description
`T_TRIM_START`	2.0 s	Seconds to trim from recording start
`SNR_THRESHOLD_DB`	12.0 dB	Threshold below 90th percentile for fade detection
`CARRIER_FILTER_BW`	500 Hz	Lowpass filter bandwidth for carrier extraction
`PHASE_JUMP_THRESHOLD`	1.5 rad	Threshold for phase splice detection
`FADE_MARGIN_SAMPLES`	100	Samples to expand fade regions
`TAU_ROUNDTRIP`	0.020 s	Round-trip time for adaptive correction assessment
`XCORR_MAX_LAG_SEC`	10.0 s	Maximum lag for cross-correlation computation

7 Output Files

7.1 Generated Plots

For each analyzed timestamp, the script generates multiple diagnostic plots:

Filename	Content
`geo_TIMESTAMP_amplitude.png`	Amplitude time series comparison between receivers
`geo_TIMESTAMP_fading.png`	Fade pattern comparison with timeline and breakdown
`geo_TIMESTAMP_xcorr.png`	Amplitude cross-correlation functions
`geo_TIMESTAMP_scatter.png`	Amplitude scatter plots (receiver 1 vs receiver 2)
`geo_TIMESTAMP_phase.png`	Phase correlation scatter plots
`geo_TIMESTAMP_XMHz_phase_stability.png`	Detailed phase stability per frequency
`geo_TIMESTAMP_structure_function.png`	Structure function $D(\tau)$ comparison
`geo_TIMESTAMP_spectrum.png`	Carrier power spectrum comparison
`geo_TIMESTAMP_fade_stats.png`	Amplitude CDF and fade statistics

7.2 Data Output File

A tab-delimited text file `geo_TIMESTAMP_data.txt` containing:

- Metadata header with analysis parameters
- Summary table by frequency (correlation metrics, fade statistics)
- Structure function $D(\tau)$ values for all time lags
- Detailed fade statistics for each receiver
- Amplitude percentile tables

7.3 Batch Summary

When analyzing multiple timestamps, an additional summary is generated:

- `geo_correlation_summary.png` — Box plots and scatter plots showing correlation statistics across all analyzed timestamps, broken down by frequency

8 Key Metrics and Interpretation

8.1 Amplitude Correlation

The Pearson correlation coefficient r between amplitude time series from both receivers:

- $r > 0.7$: Strong correlation — fading is highly correlated
- $0.3 < r < 0.7$: Moderate correlation
- $r < 0.3$: Weak correlation — ionospheric structures are small relative to baseline

8.2 Fade Correlation

Binary fade mask correlation:

- **Pearson correlation** of fade masks (1 = faded, 0 = clear)
- **Jaccard similarity**: $J = \frac{|A \cap B|}{|A \cup B|}$ where A and B are the sets of faded samples at each receiver

  Higher Jaccard values indicate fades occur simultaneously at both receivers.

8.3 Phase Correlation

During jointly valid (non-faded) periods:

- **Phase correlation**: Pearson r of detrended phase
- **Phase derivative correlation**: Correlation of $d\phi/dt$, which indicates whether phase fluctuations track together

8.4 Structure Function

The phase structure function $D(\tau) = \langle [\phi(t+\tau) - \phi(t)]^2 \rangle$ measures phase variance at different time lags:

- **RMS @ 20 ms** is the critical metric for adaptive correction feasibility
- < 0.5 rad: Good — suitable for closed-loop correction
- 0.5–1.0 rad: Marginal
- > 1.0 rad: Poor — correction difficult

8.5 Fade Statistics

For each receiver:

- **Fade fraction**: Percentage of time below threshold
- **Fade rate**: Number of fade events per minute
- **Fade duration**: Mean and maximum fade length
- **Fade depth**: How far signal drops below threshold

9 Example Analysis Workflow

9.1 Step 1: Check Available Data

```
python wwv_geographic_correlation.py --list
```

Example output:

```
Available timestamps with complete receiver pairs:
------------------------------------------------------------
  20260101.1148: ['Cambridge', 'Sudbury'] @ [2.5, 5.0, 10.0] MHz
  20260101.1415: ['Cambridge', 'Sudbury'] @ [2.5, 5.0, 10.0] MHz
```

9.2 Step 2: Analyze Specific Timestamp

```
python wwv_geographic_correlation.py 20260101.1148
```

9.3 Step 3: Review Output

The script prints a summary table and generates plots. Example console output:

```
======================================================================
GEOGRAPHIC CORRELATION ANALYSIS: 20260101.1148
Receivers: ['Cambridge', 'Sudbury']
Baseline: ~32 km
======================================================================
Common frequencies: [2.5, 5.0, 10.0] MHz

  Analyzing 2.5 MHz: Cambridge vs Sudbury
    Cambridge: 87% valid, RMS@20ms=0.312 rad
    Sudbury: 82% valid, RMS@20ms=0.298 rad
    Amplitude correlation: r=0.652 at lag=0.0 ms
    Fade correlation: r=0.445, Jaccard=0.312
    Joint valid: 75%
    Phase correlation: r=0.234
    Phase derivative correlation: r=0.187
```

9.4 Step 4: Batch Analysis

For comprehensive study across multiple timestamps:

```
python wwv_geographic_correlation.py --date 20260101
```

This generates all individual plots plus a summary showing trends across timestamps.

10 Interpretation Guidelines

10.1 Highly Correlated Fading (Jaccard > 0.5)

- Ionospheric structures are large compared to the 32 km baseline
- Adaptive correction at one site may benefit nearby sites
- Often seen at lower frequencies where ionospheric effects are stronger

10.2 Uncorrelated Fading (Jaccard < 0.2)

- Small-scale ionospheric turbulence dominates
- Each receiver experiences independent propagation conditions
- Adaptive correction would need to be site-specific

10.3 Frequency Dependence

Expect to see:

- **Lower frequencies** (2.5, 5 MHz): Higher absorption, potentially more correlated large-scale structure
- **Higher frequencies** (10, 15, 20 MHz): Less ionospheric effect, may show different correlation patterns depending on whether single or multiple modes are present

11 Troubleshooting

11.1 No Complete Receiver Pairs Found

- Check that files follow the expected naming convention
- Verify timestamps match exactly between receivers
- Ensure both receivers have recordings at the same frequencies

11.2 Insufficient Joint Valid Data

If phase correlation shows "Insufficient joint valid data":

- Both receivers have extensive fading at the same time
- Try a different timestamp or frequency with better propagation

11.3 Unexpected Correlation Values

- Check that recordings truly cover the same time period
- Verify receiver clocks are synchronized
- Large timing offsets appear in the cross-correlation peak lag

12 References

- WWV Phase Analysis User's Guide — detailed phase stability analysis
- HF Polarization Physics — magnetoionic propagation background
- HF 3D Propagation Simulator Manual — ray tracing simulations

---

Analysis developed with Claude Code, January 2026.