Forest Health Index (FHI) and Stand Health Index (SHI) Calculation and Visualization Manual for Duncan Woods

---

1. Introduction & Rationale

Forests are vital ecological systems providing climate regulation, biodiversity preservation, and timber production. Monitoring their health is essential, especially with threats from pests, diseases, and climate change. Traditional survey methods are labor-intensive and often subjective. The Forest Health Index (FHI) and Stand Health Index (SHI) provide quantitative, reproducible measures derived from remote sensing data, allowing timely, large-scale monitoring.

This project integrates multispectral data from Sentinel-2, structural metrics from GEDI LiDAR, and stability analysis over multiple years to produce a robust forest health assessment. The process combines Geographic Information System (GIS) tools, Google Earth Engine (GEE) scripting, Python-based data processing, and Kepler.gl visualization.

---

2. GIS Preparation in QGIS

1. **Install QGIS**: Download from qgis.org and install.
2. **Load a Basemap**: Use Google Satellite or OpenStreetMap via XYZ Tiles.
3. **Create Polygon Layer**:
4. Layer → Create Layer → New Shapefile Layer → Polygon.
5. CRS: EPSG:4326.
6. **Digitize AOI**: Trace the Duncan Woods boundary.
7. **Attribute Table**: Add fields like `id` or `name`.
8. **Save Shapefile**: Name it `Duncan_woods_1acre_grid.shp`.
9. **Upload to GEE**: In GEE, Assets → New → Shapefile Upload → Import all `.shp`, `.shx`, `.dbf`, `.prj` files.

---

3. GEE Data Acquisition

We use Sentinel-2 Level-2A imagery for NDVI.

- **Bands**: B8 (NIR) and B4 (Red) are used.
- **NDVI Formula**: `(B8 - B4) / (B8 + B4)`.
- **Cloud Masking**: QA60 band is used to remove clouds/cirrus.

Key GEE Code ([from GEE_Code.txt]):

```
var s2 = ee.ImageCollection('COPERNICUS/S2_SR_HARMONIZED')
  .filterBounds(aoi)
  .filterDate(startDate, endDate)
  .filter(ee.Filter.lt('CLOUDY_PIXEL_PERCENTAGE', 60))
  .map(maskS2clouds)
  .map(function(img) {
    return img.normalizedDifference(['B8','B4']).rename('NDVI');
  });
```

Monthly Aggregation: The script loops over each month, calculates median NDVI, and exports a table to Google Drive.

---

4. Downloading and Local Data Handling

- In Google Drive, download the exported CSV.
- Place it in your working directory.
- Load in Python:

```
import pandas as pd
df = pd.read_csv('DuncanWoods_NDVI_mean_2018_2025_fixed.csv')
df['NDVI_mean'] = df['NDVI_mean'].replace(-9999, np.nan)
```

---

5. Metric Normalization & FHI Calculation

NDVI Normalization:

```
ndvi_min = df['NDVI_mean'].min()
ndvi_max = df['NDVI_mean'].max()
df['NDVI_norm'] = (df['NDVI_mean'] - ndvi_min) / (ndvi_max - ndvi_min)
```

Canopy Height & Cover: Normalized using min–max scaling. **FHI Formula**: $FHI = (NDVI_{norm} + Height_{norm} + Cover_{norm})/3$

---

6. Stability Integration

Compute yearly NDVI std per cell:

```
df['year'] = pd.to_datetime(df['date']).dt.year
ndvi_std = df.groupby(['cell_id','year'])['NDVI_mean'].std().reset_index()
```

Normalize inverse std to get `Stability_norm`.

---

7. Updated_FHI Creation

We combine `FHI` and `Stability_norm` using the harmonic mean: $Updated\_FHI = \frac{2}{(1/FHI_{norm})+(1/Stability_{norm})}$ Rescale to original FHI range for interpretability.

---

8. Classification Thresholds

Calculate mean (μ) and standard deviation (σ) of `Updated_FHI`. Categorize:

- $\leq$ μ−1.5σ: Severely Stressed
- μ−1.5σ to μ−1.0σ: Stressed
- μ−1.0σ to μ: Slightly Stressed
- μ to μ+1.0σ: Moderately Healthy
- ## μ+1.0σ to μ+1.5σ: Healthy

    μ+1.5σ: Very Healthy Override: FHI < 0.4 → Severely Stressed.

---

9. Filtering for Specific Dates

```
df['date'] = pd.to_datetime(df['date'])
df_june2025 = df[(df['date'].dt.year == 2025) & (df['date'].dt.month == 6)]
```

---

10. Kepler.gl Visualization

1. Prepare CSV with `cell_id`, `.geo`, `Updated_FHI`, `FHI_Category`.
2. Upload to kepler.gl.
3. Add Polygon Layer.
4. Color by `FHI_Category`.
5. Adjust legend and style.
6. Export as HTML or PNG.

11. Full Code Commentary

Every Python and GEE block is annotated to explain variable roles, functions, and transformations.

12. Final Outputs & Interpretation

June 2025 ranges:

Category	Min	Max
Severely Stressed	0.5609	0.5825
Stressed	0.5842	0.6093
Slightly Stressed	0.6115	0.6603
Moderately Healthy	0.6609	0.7115
Healthy	0.7121	0.7366
Very Healthy	0.7377	0.7804

The process creates a reproducible, scalable framework for monitoring forest health with high spatial and temporal resolution.