Precipitation Stretch Optimizer with Offsets - Documentation

Overview

The `stretch_precipitation_with_offsets()` function applies a three-stage transformation to daily precipitation and temperature time series data:

1. **Monthly Offset Application**: Applies monthly temperature offsets and precipitation multipliers to adjust baseline conditions
2. **Percentile-Based Stretching**: Stretches precipitation values based on their position in the cumulative distribution
3. **Mass Balance Optimization**: Automatically optimizes parameters to ensure total precipitation is conserved

This tool is designed for climate change impact studies where you need to:

- Adjust precipitation amounts and temperature by month (e.g., wetter winters, drier summers)
- Intensify extreme precipitation events while maintaining water balance
- Generate modified climate scenarios for hydrological modeling

Input Files

1. Daily Climate Data File

Required columns:

- `date`: Date in YYYY-MM-DD format (or other standard formats)
- `precipitation`: Daily precipitation amount (mm or other units)
- `air_temperature`: Daily air temperature (°C or other units)

Example format (short.csv):

```
date,precipitation,air_temperature
1990-01-01,5.2,2.3
1990-01-02,0.0,-1.5
1990-01-03,12.8,3.7
```

Notes:

- Column names are case-insensitive
- Handles common spelling variants (e.g., "precipitaition", "temp")
- Date formats are auto-detected

2. Monthly Offset File

Required columns:

- `Month`: Month number (1-12)
- `PPctChange`: Precipitation percent change for the month (%)
- `Toffset`: Temperature offset for the month (°C)

Example format (MonthlyDeltaShifts.csv):

```
Month,PPctChange,Toffset
1,10,2.0
2,5,1.8
3,-5,1.5
...
12,15,2.2
```

Notes:

- Must contain exactly 12 rows (one per month)
- PPctChange can be positive (wetter) or negative (drier)
- Toffset can be positive (warmer) or negative (cooler)

Processing Workflow

Stage 1: Apply Monthly Offsets

Temperature Offset: All days in each month receive the same temperature adjustment:

```
shifted_temperature = original_temperature + Toffset[month]
```

Precipitation Multiplier: Only days with precipitation > 0 are adjusted:

```
shifted_precipitation = original_precipitation × (100 + PPctChange[month]) / 100
```

Example:

- If PPctChange = 10 for January, all January precipitation values are multiplied by 1.10 (10% increase)
- If PPctChange = -20 for July, all July precipitation values are multiplied by 0.80 (20% decrease)
- Days with 0 precipitation remain 0

Stage 2: Calculate Percentiles

For all days with precipitation > 0 (using shifted values):

1. **Rank** all non-zero precipitation values
2. **Calculate empirical CDF**: `z = ((rank - 0.5) / n) × 100`
    - Where n = total number of days with precipitation
    - z ranges from ~0 to ~100 (percentile)

Stage 3: Apply Stretching Formula

The stretching formula has two regimes based on the threshold percentile `t`:

For days with z ≥ t (above threshold):

$$x = (z - t) / (100 - t)$$
$$p' = p \times ((100 + s) / 100) \times x^d$$

For days with z < t (below threshold):

$$x = z / t$$
$$p' = p \times c \times (x^a \times (1-x)^b)$$

Where:

- `p` = shifted precipitation value
- `p'` = stretched precipitation value
- `t` = threshold percentile (user-specified)
- `s` = stretch factor (user-specified)
- `a, b, c, d` = parameters optimized to maintain mass balance

Stage 4: Mass Balance Optimization

The function automatically optimizes parameters `a`, `b`, `c`, and `d` using the Nelder-Mead algorithm to ensure:

$$\text{sum(stretched\_precipitation)} \approx \text{sum(shifted\_precipitation)}$$

This preserves total precipitation while redistributing it according to the stretching formula.

Function Parameters

Required Parameters

- **`input_file`**: Path to CSV file with daily climate data
- **`threshold`**: Percentile threshold (0-100)
  - Typical values: 90-99
  - Higher values = stretch only the most extreme events
- **`stretch_factor`**: Percentage increase for extreme events
  - Example: 50 means extreme events increase by up to 50%

Optional Parameters

- **`offset_file`**: Path to monthly offset CSV (default: "MonthlyDeltaShifts.csv")
- **`output_file`**: Output file path (default: "stretched_precipitation.csv")
- **`date_format`**: Date format string (default: "%Y-%m-%d", auto-detected if different)
- **`tolerance`**: Convergence tolerance for optimization (default: 0.01 = 1%)
- **`max_iter`**: Maximum optimization iterations (default: 1000)

Output Files

1. Main Output CSV

Columns:

- `date`: Date
- `original_precipitation`: Original input precipitation
- `shifted_precipitation`: After applying monthly multipliers
- `stretched_precipitation`: After stretching (final output)
- `stretch_factor`: Ratio of stretched/shifted for each day
- `original_temperature`: Original input temperature
- `shifted_temperature`: After applying monthly offsets

2. Metadata File

Automatically created with `_metadata.txt` suffix containing:

- Input/output file paths
- Parameters used
- Optimized parameter values
- Summary statistics
- Convergence information

Usage Examples

Example 1: Basic Usage

r

```r
result <- stretch_precipitation_with_offsets(
  input_file = "short.csv",
  offset_file = "MonthlyDeltaShifts.csv",
  output_file = "stretched_precipitation.csv",
  threshold = 95,
  stretch_factor = 50
)
```

This will:

- Apply monthly offsets from MonthlyDeltaShifts.csv
- Stretch events above the 95th percentile
- Increase extreme events by up to 50%

- Maintain mass balance

Example 2: More Conservative Stretching

r

```
result <- stretch_precipitation_with_offsets(
  input_file = "short.csv",
  offset_file = "MonthlyDeltaShifts.csv",
  threshold = 90,
  stretch_factor = 30,
  tolerance = 0.001  # Higher precision
)
```

Example 3: Custom File Paths

r

```
result <- stretch_precipitation_with_offsets(
  input_file = "data/climate_1990_2020.csv",
  offset_file = "scenarios/rcp85_offsets.csv",
  output_file = "results/scenario_rcp85_stretched.csv",
  threshold = 98,
  stretch_factor = 75
)
```

Interpreting Results

Console Output

The function prints detailed information:

1. **Offset Application Summary**
   - Shows PPctChange and Toffset for each month
   - Reports total precipitation before/after offsets
   - Reports mean temperature before/after offsets
2. **Optimization Progress**
   - Displays optimized parameters (a, b, c, d)
   - Shows convergence error
3. **Final Summary**
   - Complete transformation: Original → Shifted → Stretched
   - Mass balance verification
   - Date range and record counts

Checking Quality

Good indicators:

- Convergence error < tolerance (typically < 1%)
- Mass balance ratio very close to 1.0
- Reasonable parameter values (typically 0.1 to 10)

Warning signs:

- High convergence error (> 1%)
- Extreme parameter values (< 0.01 or > 100)
- Missing stretch ratios in output

Common Use Cases

Climate Change Scenario 1: Wetter Winters, Drier Summers

csv

```
Month,PPctChange,Toffset
1,15,2.5
2,15,2.5
3,10,2.0
4,5,2.0
5,0,2.5
6,-10,3.0
7,-15,3.5
8,-15,3.5
9,-5,3.0
10,5,2.5
11,10,2.5
12,15,2.5
```

With threshold=95, stretch_factor=50:

- Winters get 15% more precipitation
- Summers get 15% less precipitation
- All months get warmer
- Extreme events intensify by up to 50%

Climate Change Scenario 2: Uniform Warming, Intense Extremes

csv

```
Month,PPctChange,Toffset
1,0,3.0
2,0,3.0
...
12,0,3.0
```

With threshold=99, stretch_factor=100:

- No change to monthly precipitation totals
- All months warm by 3°C
- Only top 1% of events stretch
- Most extreme events can double in intensity

Technical Details

Percentile Calculation Method

Uses empirical CDF with mid-rank adjustment:

```
percentile = ((rank - 0.5) / n) × 100
```

This avoids assigning 0% or 100% to any observation.

Optimization Algorithm

- Method: Nelder-Mead simplex algorithm
- Objective: Minimize relative error between stretched and shifted totals
- Constraints: All parameters must be positive (enforced via abs())

Mass Balance Verification

The stretching preserves mass balance on the **shifted** precipitation:

```
sum(stretched_precipitation) ≈ sum(shifted_precipitation)
```

Note: If monthly multipliers change total precipitation, the stretched total will match the shifted total, not the original total. This is intentional.

Troubleshooting

Issue: High Convergence Error

Possible causes:

- Extreme stretch_factor values
- Very high or low threshold values
- Complex precipitation distribution

Solutions:

- Reduce stretch_factor
- Increase max_iter
- Decrease tolerance requirement

Issue: Missing Stretch Ratios

Possible causes:

- Days with precipitation in input but not in daily stretch file
- Date format mismatch

Solutions:

- Verify date formats match
- Check for gaps in time series

Issue: Unrealistic Temperature/Precipitation Values

Possible causes:

- Incorrect offset file values
- Wrong units

Solutions:

- Verify offset file contents
- Check that PPctChange is in percent (not decimal)
- Check that Toffset matches temperature units

References

This methodology combines:

1. Monthly delta change approach (common in climate impact studies)
2. Percentile-based stretching (precipitation intensification)
3. Mass balance preservation (hydrological consistency)

Version History

- **v2.0** (Current): Added monthly offset functionality
- **v1.0**: Original percentile-based stretching only

Contact & Support

For questions about this code or methodology, please refer to the project documentation or contact the development team.