RainfallFussion - Multi-Station Imputation Tool

Daftar Isi

Dokumentasi lengkap dengan penjelasan fitur, instalasi, metodologi, dan troubleshooting.

Bagian Utama

- Deskripsi Proyek
- Fitur Utama
- Teknologi yang Digunakan
- Instalasi dan Penggunaan
- Metodologi Imputasi
- Panduan Penggunaan
- Format Data Input
- Output dan Hasil
- Metrik Validasi
- Struktur File

Deskripsi Proyek

RainfallFussion adalah alat web-based yang dirancang untuk melakukan imputasi (pengisian) data curah hujan yang hilang dari multiple weather stations menggunakan pendekatan **Spatial Correlation Ensemble**.

Proyek ini dikembangkan untuk **FFWS SURAKARTA** (Flood Forecast Warning System Surakarta) dan menggunakan kombinasi dari empat metode statistik berbeda dalam voting ensemble untuk menghasilkan estimasi yang akurat dan reliable.

Lokasi Proyek:

```
d:\000 manajemen aku\FFWS SURAKARTA BARU\DATA\final RainfallFussion\
```

Fitur Utama

1. **Input Data Multi-Stasiun**

- Upload data dari tiga sumber stasiun:
  - **Stasiun Target** (stasiun dengan data yang akan diisi)
  - **Stasiun Referensi 1 & 2** (stasiun pembanding untuk imputasi)
- Format: File Excel (.xlsx) dengan struktur data tertentu

2. **Perhitungan Koordinat & Jarak**

- Input koordinat geografis (latitude/longitude) setiap stasiun

- Otomatis menghitung jarak antar stasiun menggunakan formula Haversine
- Menampilkan visualisasi jarak dalam kilometer

3. **Multi-Method Ensemble Imputation**

Menggunakan 4 metode statistik dalam voting ensemble:

- **Normal Ratio Method (NRM)** - berdasarkan rasio rata-rata tahunan
- **Inverse Distance Weighting (IDW)** - berdasarkan bobot jarak inverse
- **Linear Regression Method** - menggunakan model regresi linier
- **Correlation Weighting** - berdasarkan korelasi spasial antar stasiun

4. **Metrik Validasi Komprehensif**

Menghitung berbagai metrik untuk mengukur kualitas imputasi:

- RMSE (Root Mean Squared Error)
- MAE (Mean Absolute Error)
- MAPE (Mean Absolute Percentage Error)
- NSE (Nash-Sutcliffe Efficiency)
- r (Pearson Correlation Coefficient)

5. **Visualisasi Data**

- Grafik perbandingan data asli vs data imputasi
- Tabel detail fill-out locations
- Statistik visual dalam bentuk metric cards

6. **Export Hasil**

- Download hasil imputasi dalam format Excel (.xlsx)
- Tersistem dengan format yang mudah dibaca

🛠️ Teknologi yang Digunakan

Teknologi	Fungsi
HTML5	Struktur halaman web
CSS3	Styling modern dengan gradient dan responsive design
Vanilla JavaScript	Logika imputasi dan manipulasi data tanpa framework
XLSX Library	Pembacaan dan penulisan file Excel
Chart.js	Visualisasi grafik dan data
JSZip	Handling file zip jika diperlukan

💻 Instalasi dan Penggunaan

Persyaratan

- Browser modern (Chrome, Firefox, Edge, Safari)
- Koneksi internet (untuk mengunduh library CDN)
- File Excel dengan format sesuai spesifikasi

Cara Menjalankan

1. **Buka file HTML**

```
Buka file index.html di browser
```

1. **Atau gunakan local server** (recommended)

```
# Menggunakan Python 3
python -m http.server 8000

# Atau menggunakan Node.js dengan http-server
npx http-server
```

1. **Akses aplikasi**

```
http://localhost:8000/index.html
```

Metodologi Imputasi

A. Normal Ratio Method (NRM)

Formula yang digunakan:

```
NRM = Σ(Rata-rata Target,m / Rata-rata Ref,m) × Nilai Ref,i,m
```

Langkah:

1. Hitung rata-rata tahunan untuk setiap bulan dari stasiun target dan referensi
2. Hitung rasio rata-rata target terhadap referensi
3. Kalikan rasio dengan nilai observasi dari stasiun referensi
4. Rata-rata hasil dari semua stasiun referensi yang tersedia

Kelebihan: Mengpertahankan karakteristik musiman data asli

B. Inverse Distance Weighting (IDW)

Formula yang digunakan:

dimana:

- Y = nilai target yang diestimasi
- X = nilai dari stasiun referensi
- a, b = koefisien regresi yang dihitung dari data observasi

Langkah:

1. Bangun model regresi linier menggunakan data yang tersedia
2. Gunakan model untuk memprediksi nilai yang hilang
3. Rata-rata hasil dari semua stasiun referensi

Kelebihan: Memperhitungkan hubungan linier antar stasiun

D. Correlation Weighting

Formula yang digunakan:

```
CW = (|r| × nilai_1 + |r| × nilai_2) / (|r|_1 + |r|_2)
```

dimana r = Pearson Correlation Coefficient

Langkah:

1. Hitung korelasi Pearson antara stasiun target dan referensi
2. Gunakan nilai absolut korelasi sebagai bobot
3. Rata-rata tertimbang nilai observasi

Kelebihan: Berdasarkan kekuatan hubungan statistik antar data

---

E. Voting Ensemble

Format kombinasi akhir:

```
Nilai Imputasi Akhir = (NRM + IDW + LR + CW) / n
```

dimana n = jumlah metode yang menghasilkan nilai valid (≥ 0)

Keuntungan:

- Mengurangi bias dari satu metode individual
- Meningkatkan robustness hasil
- Lebih akurat untuk data dengan karakteristik kompleks

🕮 Panduan Penggunaan

Step 1: Input Koordinat Stasiun

1. Klik setiap card "Koordinat Stasiun"
2. Masukkan **Latitude** (°) - nilai antara -90 hingga 90
3. Masukkan **Longitude** (°) - nilai antara -180 hingga 180
4. Klik tombol "Simpan Koordinat"

```
Contoh:
Stasiun Target - Lat: -7.5515, Long: 110.8243
Stasiun Ref 1  - Lat: -7.5600, Long: 110.8100
Stasiun Ref 2  - Lat: -7.5400, Long: 110.8400
```

Catatan: Jarak akan otomatis dihitung setelah koordinat disimpan.

Step 2: Upload File Data

1. Klik tombol **"Pilih File"** di setiap card stasiun
2. Pilih file Excel (.xlsx) yang sesuai format
3. Tunggu notifikasi "File berhasil dimuat"
4. Ulangi untuk semua 3 stasiun (Target, Ref1, Ref2)

Indikator Kesuksesan:

- Semua 3 file dimuat
- Tombol "Process Data" akan aktif (tidak lagi disabled)
- Status menunjukkan "☑ Semua file berhasil dimuat!"

Step 3: Proses Data

1. Klik tombol **"PROCESS DATA"** (akan aktif setelah semua file dimuat)
2. Tunggu progress bar mencapai 100%
3. Proses meliputi:
    - Perhitungan korelasi spasial
    - Imputasi dengan 4 metode
    - Perhitungan metrik validasi
    - Pembuatan visualisasi

Timeline Estimasi: 5-30 detik tergantung ukuran data

Step 4: Lihat Hasil

Setelah proses selesai, akan menampilkan:

A. Metrik Validasi (Metric Cards)

- **RMSE** - Root Mean Squared Error (semakin kecil semakin baik)
- **MAE** - Mean Absolute Error (error rata-rata absolut)
- **MAPE** - Mean Absolute Percentage Error (error persentase)
- **NSE** - Nash-Sutcliffe Efficiency (1.0 = sempurna)
- **r** - Correlation Coefficient (1.0 = perfect correlation)

B. Tabel Fill-Out Locations

Menampilkan detailed info:

- Year, Day, Bulan
- Asal nilai (Asli/Imputasi)
- Nilai asli
- Nilai imputasi
- 4 Metode individual (NRM, IDW, LR, CW)

C. Grafik Visualisasi

- Grafik perbandingan data sebelum dan sesudah imputasi
- Trend line untuk analisis lebih lanjut

Step 5: Download Hasil

1. Klik tombol **"DOWNLOAD HASIL"** (aktif setelah proses selesai)
2. File Excel akan otomatis diunduh dengan nama:

```
RainfallFussion_Results_[timestamp].xlsx
```

1. Buka file di Excel atau aplikasi spreadsheet untuk analisis lebih lanjut

---

📊 Format Data Input

Struktur File Excel

File Excel harus memiliki format presisi berikut:

Row 1: Header

```
Year | Day | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec
```

Row 2 dst: Data

```
Year | Day | nilai_jan | nilai_feb | ... | nilai_des
2020 |  1  |     45.2  |    52.3   | ... |    38.5
2020 |  2  |     23.4  |    null   | ... |    42.1
...
```

Spesifikasi Detail

- **Kolom A (Year):** Tahun dalam format numerik (2020, 2021, dll)
- **Kolom B (Day):** Hari dalam tahun (1-365/366)
- **Kolom C-N (Jan-Dec):** Nilai curah hujan harian dalam mm
  - **Format:** Numerik dengan desimal (contoh: 25.5)
  - **Nilai hilang:** Gunakan `null`, kosong, atau tidak ada nilai
  - **Nilai valid:** Harus ≥ 0 (tidak ada curah hujan negatif)

Contoh Data Valid

```
Year | Day | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec
2020 |  1  | 12.5| 15.3| 20.1| 18.5| 22.3| 25.6| 28.9| 26.3| 23.1| 19.8| 16.5|
14.2
2020 |  2  | null| 14.2| 19.8|     | 21.5| 24.3| 27.6| 25.1| 22.8| 18.9| 15.3|
13.1
2020 |  3  | 11.2| 13.5| 18.9| 17.3| 20.1| 23.5| 26.4| 24.2| 21.5| 17.6| 14.8|
12.5
```

Tips Format Data

☑ Lakukan:

- Simpan dalam format .xlsx (Excel 2007+)
- Gunakan urutan bulan: Januari-Desember
- Pastikan semua sel berisi numerik atau kosong
- Gunakan separator desimal sesuai regional settings

❌ Hindari:

- Format .xls lama atau CSV
- Nilai teks seperti "N/A", "missing", "—"
- Nilai negatif atau tidak logis
- Data yang tidak terstruktur atau dengan row kosong di tengah

---

Output dan Hasil

File Output Excel

File download berisi sheet-sheet berikut:

Sheet 1: Imputed Data

Year	Day	Jan_Original	Jan_Imputed	Jan_Method	Feb_Original	...
2020	1	12.5	12.5	Original	15.3	...
2020	2	null	14.85	Ensemble	14.2	...

Keterangan:

- `_Original`: Nilai asli dari data
- `_Imputed`: Nilai hasil imputasi
- `_Method`: Sumber data (Original/Ensemble)

Sheet 2: Validation Metrics

Ringkasan metrik validasi:

```
Metric         | Value
---------------|----------
RMSE           | 2.34
MAE            | 1.89
MAPE           | 8.5%
NSE            | 0.876
Correlation    | 0.923
Items Imputed | 145
```

Sheet 3: Method Comparison

Detail untuk setiap data yang diimputasi:

Year	Day	Month	NRM_Value	IDW_Value	LR_Value	CW_Value	Ensemble_Final
2020	2	Jan	14.5	15.2	14.8	14.9	14.85

📊 Metrik Validasi

1. RMSE (Root Mean Squared Error)

$$RMSE = \sqrt{(\Sigma(y_i - \hat{y}_i)^2 / n)}$$

- **Unit:** mm (sama dengan data curah hujan)
- **Interpretasi:** Semakin kecil RMSE semakin baik
- **Rentang ideal:** < 5 mm untuk data curah hujan
- **Contoh:** RMSE = 2.34 berarti error rata-rata 2.34 mm

2. MAE (Mean Absolute Error)

```
MAE = Σ|yᵢ - ŷᵢ| / n
```

- **Unit:** mm
- **Interpretasi:** Error rata-rata tanpa memperhatikan tanda
- **Kelebihan:** Lebih robust terhadap outliers dibanding RMSE
- **Contoh:** MAE = 1.89 mm

3. MAPE (Mean Absolute Percentage Error)

```
MAPE = (100/n) × Σ|yᵢ - ŷᵢ| / yᵢ
```

- **Unit:** Persen (%)
- **Interpretasi:** Persentase error relatif terhadap nilai aktual
- **Rentang ideal:** < 10% untuk model yang baik
- **Contoh:** MAPE = 8.5% berarti error rata-rata 8.5%

4. NSE (Nash-Sutcliffe Efficiency)

```
NSE = 1 - (Σ(yᵢ - ŷᵢ)² / Σ(yᵢ - ȳ)²)
```

- **Rentang:** -∞ hingga 1.0
- **Interpretasi:**
  - NSE = 1.0 → Prediksi sempurna
  - NSE = 0.0 → Prediksi sama dengan mean observasi
  - NSE < 0 → Model lebih buruk dari mean
- **Rentang ideal:** NSE > 0.75 untuk model yang sangat baik

5. r (Pearson Correlation Coefficient)

```
r = Σ(xᵢ - x̄)(yᵢ - ȳ) / √[Σ(xᵢ - x̄)² × Σ(yᵢ - ȳ)²]
```

- **Rentang:** -1.0 hingga 1.0
- **Interpretasi:**
  - r ≈ 1 → Korelasi positif kuat
  - r ≈ 0 → Tidak ada korelasi
  - r ≈ -1 → Korelasi negatif kuat
- **Rentang ideal:** r > 0.8 menunjukkan korelasi kuat

🗀 Struktur File

```
d:\000 manajemen aku\FFWS SURAKARTA BARU\DATA\final RainfallFussion\
├── index.html                # File aplikasi utama (1239 baris)
├── desktop.ini               # Konfigurasi folder Windows
├── README.md                 # Dokumentasi (file ini)
└── [data input files]        # Folder untuk input Excel files
    ├── StasiuntTarget.xlsx   # Data stasiun target
    ├── Referensi1.xlsx       # Data stasiun referensi 1
    └── Referensi2.xlsx       # Data stasiun referensi 2
```

Deskripsi File Utama

index.html (1239 baris)

Aplikasi monolith web yang berisi:

- **HTML Structure** (lines 1-320): Elemen UI dan layout
- **CSS Styling** (lines 12-280): Design responsif dengan gradient modern
- **JavaScript Logic** (lines 312-1239): Semua fungsi imputasi dan validasi

Ukuran File: ~45 KB (uncompressed)

README.md

Dokumentasi lengkap (file ini) yang menjelaskan:

- Overview dan fitur proyek
- Instalasi dan cara penggunaan
- Metodologi statistik
- Format data input/output
- Metrik validasi

🔧 Detail Implementasi Teknis

Fungsi-Fungsi Penting

1. `readExcelFile(file, station)`

Membaca file Excel dan mengonversi ke format JSON

```
reader.readAsArrayBuffer(file);
// Result: dataFiles[station] = array of {year, day, months[]}
```

2. `calculateCorrelation()`

Menghitung korelasi Pearson antar stasiun

```
const correlation = {
    ref1: pearsonCoefficient(...),
    ref2: pearsonCoefficient(...)
}
```

3. `multiStationVotingEnsemble(targetData, refData1, refData2, distances, correlation)`

Fungsi utama imputasi menggunakan ensemble

```
for each missing value:
    predictions = [
        normalRatioMethod(...),
        inverseDistanceWeighting(...),
        linearRegressionMethod(...),
        correlationWeighting(...)
    ]
    imputedValue = average(predictions)
```

4. `calculateValidationMetrics()`

Menghitung RMSE, MAE, MAPE, NSE, r

```
metrics = {
    rmse: calculateRMSE(...),
    mae: calculateMAE(...),
    mape: calculateMAPE(...),
    nse: calculateNSE(...),
    correlation: calculatePearson(...)
}
```

5. `createCharts(data)`

Membuat visualisasi menggunakan Chart.js

```
chart = new Chart(context, {
    type: 'line',
    data: chartData,
    options: chartOptions
})
```

Variabel Global Penting

```

```

```
const dataFiles = {};           // Menyimpan data dari 3 stasiun
const imputedData = [];         // Hasil imputasi
const distances = {};           // Jarak antar stasiun (km)
const validationMetrics = {};   // Metrik validasi
```

Event Listeners

- `#targetFile.change` → Upload file target
- `#ref1File.change` → Upload file referensi 1
- `#ref2File.change` → Upload file referensi 2
- `#processBtn.click` → Trigger imputasi
- `#downloadBtn.click` → Download hasil Excel

📝 Contoh Kasus Penggunaan

Skenario: Imputasi Data Pos Hujan Surakarta

Kondisi Awal:

- Pos hujan Stasiun Target memiliki 150 hari dengan data hilang dalam setahun
- Data Referensi 1 & 2 dari stasiun terdekat tersedia lengkap
- Jarak target ke ref1 = 8.5 km, jarak ke ref2 = 12.3 km

Proses:

1. Upload 3 file data dengan struktur yang benar
2. Masukkan koordinat: Target (-7.55, 110.82), Ref1 (-7.56, 110.81), Ref2 (-7.54, 110.84)
3. Klik PROCESS DATA
4. Sistem menjalankan voting ensemble dengan 4 metode
5. Download hasil dengan 150 nilai baru yang terimputasi

Hasil yang Diharapkan:

- RMSE: 2-4 mm (tergantung variabilitas data)
- NSE: 0.80-0.95 (model sangat baik)
- Correlation: 0.85-0.98 (korelasi kuat)

❓ Troubleshooting

Problem: "Error membaca file"

Solusi:

- Pastikan format file adalah .xlsx (bukan .xls atau .csv)
- Pastikan kolom pertama adalah Year
- Hilangkan row kosong atau header tambahan

Problem: "Tombol PROCESS masih disabled setelah upload 3 file"

Solusi:

- Refresh halaman dan coba upload ulang
- Buka Developer Console (F12) untuk melihat error detail
- Pastikan semua file format sesuai spesifikasi

Problem: "Preview data not showing"

Solusi:

- Pastikan data memiliki minimal 5 baris data
- Cek tidak ada merged cells di file Excel

Problem: "Hasil imputasi bernilai 0 atau tidak logis"

Solusi:

- Pastikan stasiun referensi memiliki data pada bulan-tahun yang sama
- Cek outliers di data mentah
- Pertimbangkan menggunakan stasiun referensi yang berbeda

Referensi Metodologi

Referensi untuk Normal Ratio Method

- Xia, Y., et al. (2002). "Objective evaluation of gridded precipitation data sets...."
- Dile, Y. T., & Srinivasan, R. (2014). "Evaluation of CFSR climate data..."

Referensi untuk IDW

- Shepard, D. (1968). "A two-dimensional interpolation function for irregularly-spaced data"
- Oliver, M. A., & Webster, R. (1990). "Kriging: a method of interpolation for geographical information systems"

Referensi untuk Linear Regression

- Draper, N. R., & Smith, H. (1981). "Applied Regression Analysis"

Referensi untuk Ensemble Methods

- Solomatine, D. P., & Shrestha, D. L. (2009). "AdaBoost.RT: a boosting algorithm for regression problems"

Informasi Pengembang

Proyek: RainfallFussion - Multi-Station Imputation Tool **Organisasi:** FFWS SURAKARTA (Flood Forecast Warning System) **Tujuan:** Imputasi data curah hujan yang hilang untuk forecasting banjir

Fitur Utama:

- ☑ Spatial Correlation Analysis

- ☑ Multi-Method Ensemble
- ☑ Comprehensive Validation Metrics
- ☑ Excel Integration
- ☑ Responsive Web UI

Browser Compatibility:

- ☑ Chrome 90+
- ☑ Firefox 88+
- ☑ Edge 90+
- ☑ Safari 14+

Lisensi & Penggunaan

Dokumentasi ini dibuat untuk kebutuhan internal FFWS SURAKARTA. Untuk pertanyaan atau saran pengembangan, silakan hubungi tim yang bersangkutan.

Last Updated: Februari 2025 **Version:** 1.0 **Status:** Production Ready

---

Support & Contact

Untuk bantuan teknis atau pertanyaan mengenai penggunaan aplikasi:

1. Baca troubleshooting section di atas
2. Buka Developer Console (F12) untuk melihat error log
3. Periksa format data input sesuai spesifikasi
4. Hubungi admin sistem FFWS SURAKARTA

Selamat menggunakan RainfallFussion!