# Hyytiälä Forest Station DMPS & APS Data Analysis (2010-2022)

## Project Overview

This project involves the collection, processing, and analysis of aerosol particle size distribution data from the Hyytiälä Forest Station. The dataset includes Differential Mobility Particle Sizer (DMPS) and Aerodynamic Particle Sizer (APS) measurements spanning a 10-year period (2010-2022). The objective is to extract, clean, and analyze this data to observe seasonal and yearly trends, as well as to combine DMPS and APS data for a comprehensive particle size distribution (PSD) analysis.

## Data Collection & Preprocessing

## Step 1: Data Retrieval

- The DMPS data was retrieved from the University of Helsinki Data Cloud.

(<https://datacloud.helsinki.fi/index.php/s/anfjq4A7WJ9Wt3c?path=%2FData>)

- The relevant dataset was located in `/Data/Hyytiala/DMPS\_sumfiles`.

(<https://datacloud.helsinki.fi/index.php/s/anfjq4A7WJ9Wt3c?path=%2FData%2FHyytiala%2FDMPS_sumfiles>)

- Data was downloaded for each individual year (e.g., 2011).

## Step 2: Extracting & Converting DMPS Data

- DMPS `.sum` files were extracted and combined into a single yearly Excel file (`.xlsx`).

- Data was grouped according to its respective channel/bin size.

- The combined dataset included all 365 days of data for a given year.

- A date column was added to maintain temporal alignment.

## Step 3: Cleaning & Filtering Data

- Duplicate rows caused by repeated channel information were removed using Excel’s filter function.

- After file conversion, manual verification was conducted to ensure data integrity.

- The final cleaned dataset represented a complete yearly summary.

## Step 4: Combining Multi-Year Data

- The yearly DMPS datasets were manually combined to create a full dataset spanning 2010-2022.

- Manual merging was preferred to ensure accuracy and data consistency.

## Step 5: Hourly Averaging of DMPS Data

- The dataset was converted into hourly averages for further analysis.

## Step 6: Processing APS Data

- The same procedure (steps 1-5) was applied to the APS dataset.

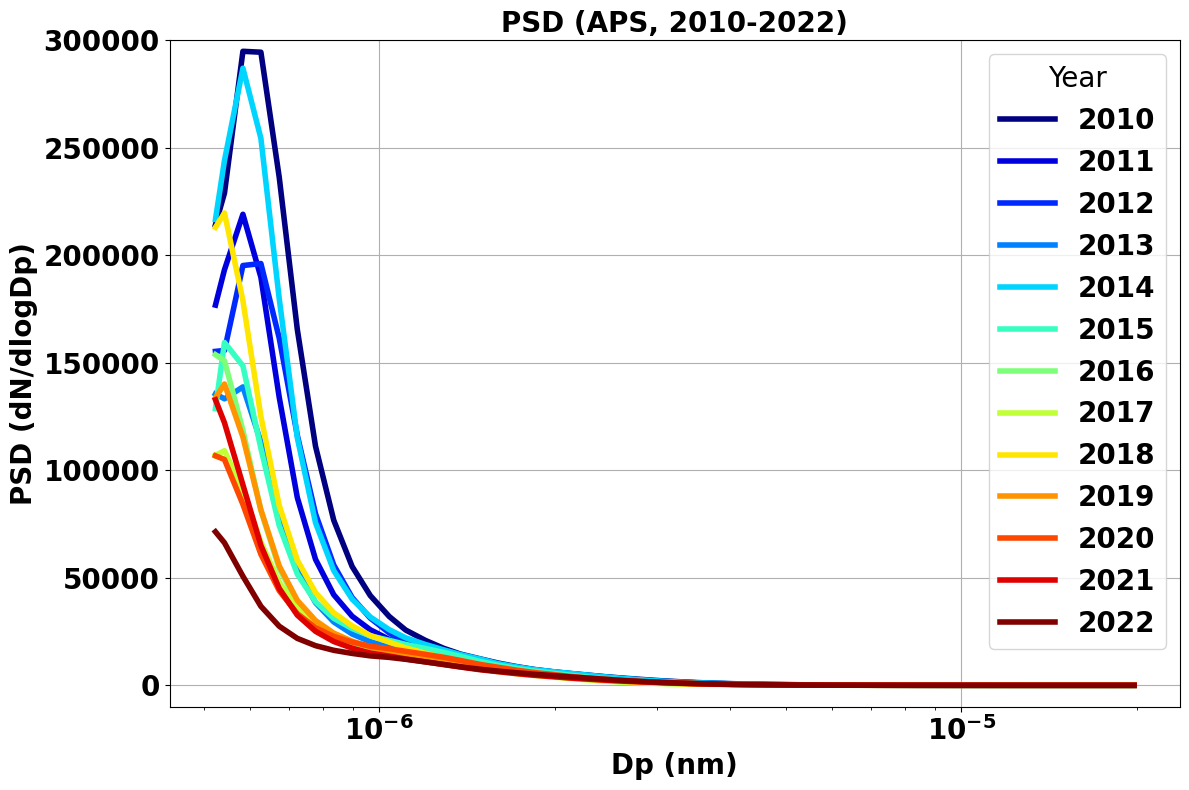
- APS data was also converted into hourly averages.

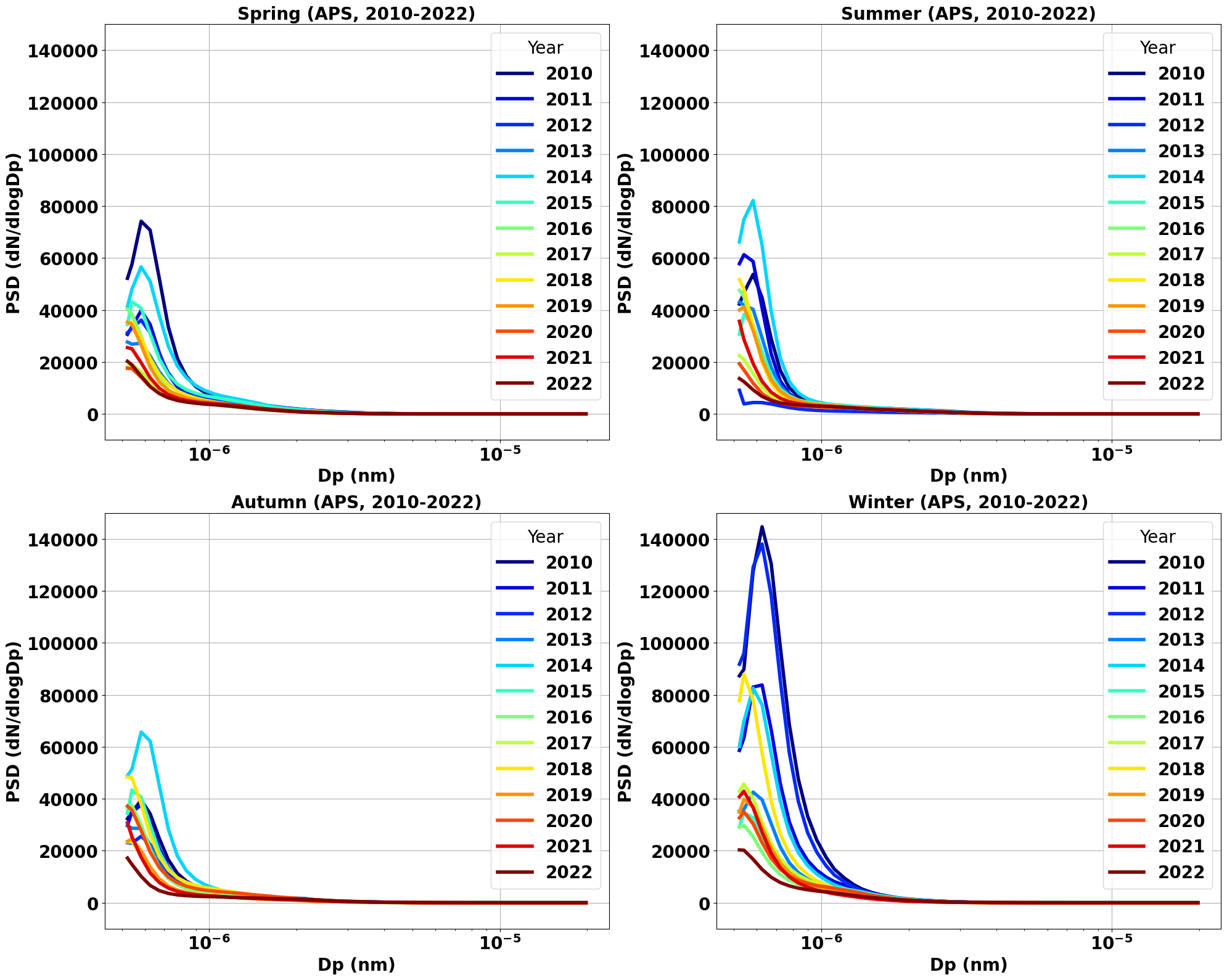
## Data Analysis & Visualization

## Step 7: Yearly Summation & Seasonal Analysis

- \*\*Yearly Sum Plots\*\*: Each year’s total data was plotted for APS (2010-2022).

- \*\*Seasonal Trends\*\*: Data was divided into four seasonal periods (Spring, Summer, Autumn, Winter) and plotted separately.





## Step 7: Yearly Summation & Seasonal Analysis

- \*\*Yearly Sum Plots\*\*: Each year’s total data was plotted for DMPS (2010-2022).

- \*\*Seasonal Trends\*\*: Data was divided into four seasonal periods (Spring, Summer, Autumn, Winter) and plotted separately.

|  |
| --- |
|  |
|  |

## Step 9: Combining DMPS & APS Data

- The DMPS and APS datasets were merged to form a complete Particle Size Distribution (PSD).

- \*\*Yearly Sum Plots\*\* and \*\*Seasonal Trends\*\* were analyzed for the merged dataset.

|  |
| --- |
|  |
|  |

## Project Deliverables

- \*\*Raw Data:\*\* Retrieved `.sum` files for DMPS and APS.

- \*\*Processed Data:\*\* Cleaned and formatted Excel files.

- \*\*Hourly Averaged Data:\*\* Aggregated hourly data for both instruments.

- \*\*Visualization Plots:\*\* Yearly and seasonal trend graphs.

- \*\*Final Combined Dataset:\*\* Merged DMPS and APS dataset for comprehensive PSD analysis.

- \*\*Code Scripts:\*\* Data extraction, cleaning, merging, and visualization scripts.

## Future Work & Enhancements

- Automate the data extraction and merging process using Python scripts.

- Implement data quality control checks.

- Develop interactive visualizations for enhanced data exploration.

- Publish findings in a research paper or scientific report.

## Repository Structure

```

📂 Hyytiälä-DMPS-APS-Analysis

│── 📂 Data

│ ├── 📂 Raw

│ ├── 📂 Processed

│ ├── 📂 Hourly\_Avg

│── 📂 Scripts

│── 📂 Visualizations

│── 📜 README.md

│── 📜 Data\_Processing\_Steps.md

│── 📜 Results.md

```