# Identifying and Removing Data Records Influenced by Fog

*This document covers a general overview of fog and fog-affected data and how to remove them.*

## Table of Contents



## Introduction

Particulate Matter (PM) measurements made using light scattering-based sensors (e.g., nephelometers and optical particle counters including the MODULAIR and MODULAIR-PM[1]) are affected during periods of high relative humidity due to the hygroscopic growth of aerosols as well as the presence of fog[2,3]. Often, the measurement of fog leads to an overestimation of $PM_{10}$ and TSP readings. The increase in PM mass readings can be corrected if the hygroscopic growth factor of the measured aerosol is known or can be estimated[4]. However, the presence of fog is more challenging to deal with. This document outlines an approach for flagging and removing data records where fog interference is present.

## Fog and its impact on laser-based particle measurements

Fog is defined as water droplets or ice crystals formed near the surface of the Earth that lead to poor visibility (< 1km)[5]. There are varied conditions under which fog forms (e.g., radiation fog, advection fog, valley fog) making the exact meteorologic conditions under which it forms challenging to identify using sensor measurements alone[3].

The method and degree to which fog particles interfere with PM measurements depends on the measurement method (e.g., nephelometer vs. optical particle counter) and the PM size range you are trying to measure. Please read Hagan and Kroll (2020)[2] for more information on the fundamental physics of why this is the case. Generally, air sensors comprised of a nephelometer (e.g., Plantower PMS5003) will not be affected by fog in the same way that those with optical particle counters (e.g., Alphasense OPC-N3) will. Fog particles are large (up to 50 μm[6]) and comprised of mostly water. Due to their size, most nephelometers cannot detect fog particles as many nephelometers cannot measure particles larger than around 1 μm[7]. Optical Particle Counters (OPCs) can count and size particles much larger than 1 μm and thus can measure fog, though the sizing accuracy may be quite poor due to the difference in refractive index between liquid water and the sensors' calibrant.

When fog is present, the geometric mean diameter of the particles measured by the OPC will increase rapidly, as will the mass measured. The top panel of Figure 1 shows how $PM_C$ ($PM_C = PM_{10} - PM_{2.5}$) increases rapidly during known fog events. During this period, the $PM_1$ and $PM_{2.5}$ remain relatively consistent as the fog particles are mostly larger than 3 μm (Figure 3).

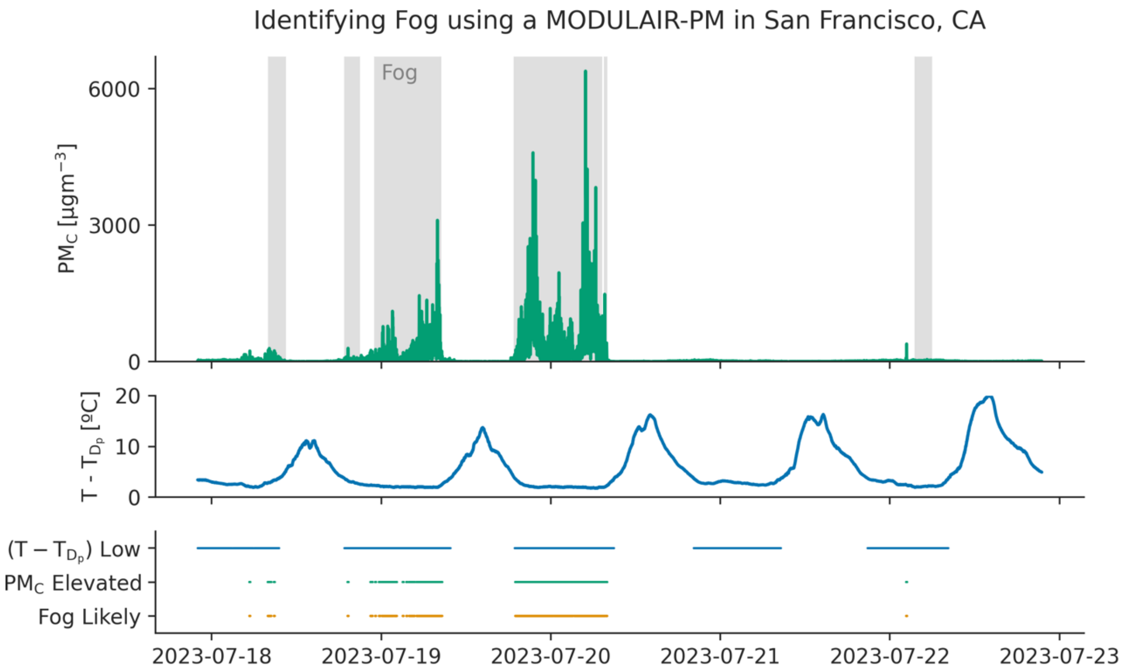


Figure 1. Fog can be identified by using the spread in dew point temperature and elevated $PM_C$ as an approximate heuristic. The top subplot shows the $PM_C$ values in green with gray boxes indicating periods where fog was identified using a trail camera. The middle subplot shows the spread in dew point temperature. The bottom subplot shows when each heuristic was true.


Figure 2. An image captured during a known fog event in San Francisco, CA. Data during this period is shown in Figure 1 during a several-hour fog event.

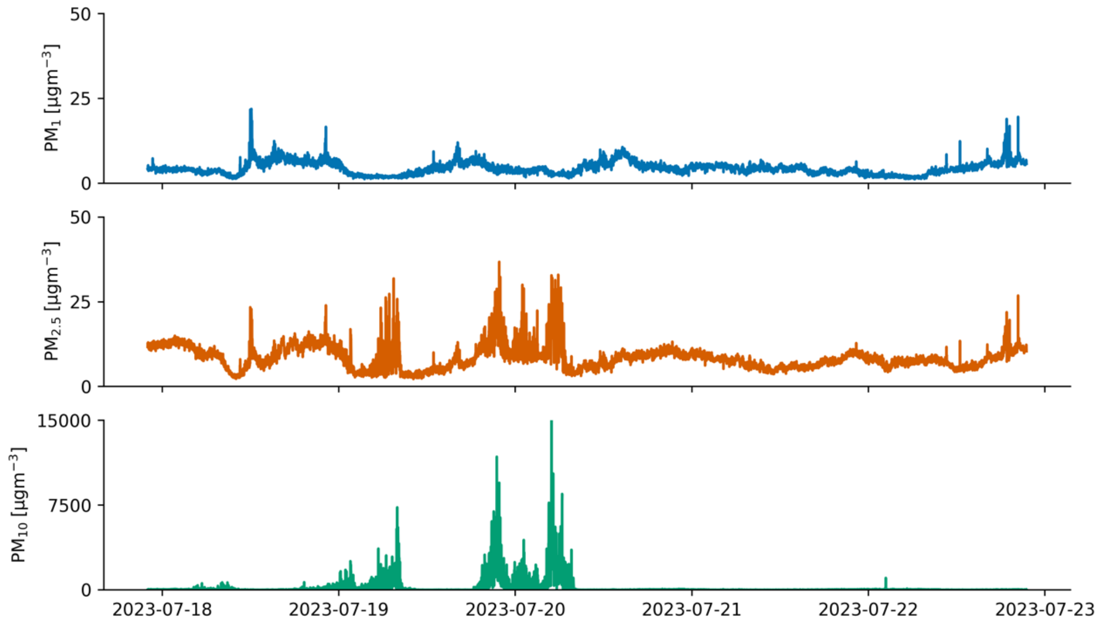


Figure 3. Fog affects $PM_{10}$ with slight implications for $PM_{2.5}$ measurements. $PM_1$ is not impacted. Please note the y-axes are different for the top and middle panels than for the bottom panel.

## Why is measuring fog an issue?

While fog is an aerosol, we are interested in measuring the dry mass of the aerosol. Regulations are based on the dry mass and most air sensor use is in support of regulatory compliance. If you are unable to account for – or correct for – the presence of fog in your data, you may see an increase in exceedances due to the overestimation of PM mass.

## How is fog interference typically addressed?

Instruments that are certified to comply with US EPA Federal Reference Method (FRM) or Federal Equivalent Method (FEM) typically include a pre-conditioning step to dry the incoming air using either a diffusion drier or thermal drier built into a heated inlet system. This step is designed to remove excess water including fog to ensure the incoming air is at or below 35% relative humidity for $PM_{2.5}$ monitoring of between 30-40% for $PM_{10}$ monitoring.

The power requirements and cost for implementing a heated inlet are often prohibitive for including with air sensors. There are few, if any, examples of air sensors using active heating to remove excess water before measurement, though it is an active area of research[8].

## How can affected data records be flagged or removed?

We have developed a simple heuristic that can be easily applied in post-processing to identify when fog is present so that it can be removed if desired. This is designed to be a temporary fix that is easy to implement. More robust approaches are being explored, as documented in the Future Research section below.

We have found that fog is most likely to form when the difference between the measured ambient temperature and the dew point is less than 3.75°C. We refer to this as the "dew point spread". We recommend flagging and removing data when the following two conditions are met:

1. The dew point spread is less than 3.75°C
2. The coarse particles ($PM_C$) exceed 200 $\mu gm^{-3}$

Since fog events tend to occur for longer than a 1-minute window, we recommend smoothing the data using a rolling maximum over an eleven-minute window to ensure you don't see instantaneous drops in flagged data. This means that for an observation window at time *t*, we recommend using the maximum value seen from between *t-5* to *t+5* minutes. Below are sample implementations in both Python and Microsoft Excel.

### Python Implementation

To follow this approach, you should use the pandas python library. The code here assumes you are using the raw data downloaded from the QuantAQ Cloud. If you have modified column names or otherwise modified the data, you may need to adjust the code accordingly.

```python
import math

def dew_point_c(temp_c, rh_pct):
    """Compute the Dew Point in degrees Celcius.

    See Psychrometry and Psychrometric Charts (1974),
        https://en.wikipedia.org/wiki/Dew_point#cite_ref-20

    Parameters
    ----------
    temp_c : float
        The temperature in degrees celcius.
    rh_pct : float
        The relative humidity as a percentage.

    Returns
    -------
    temp_dew_point : float
        The estimated dew point temperature in degrees celcius.
    """
    A = 17.27
    B = 237.7

    alpha = ((A * temp_c) / (B + temp_c)) + math.log(rh_pct / 100.0)

    return (B * alpha) / (A - alpha)


def add_fog_flag(data, temp_col="sample_temp", rh_col="sample_rh", **kwargs):
    """Return a modified DataFrame with heuristic fog interference information.

    Parameters
    ----------
    data : pd.DataFrame
    temp_col : str
    rh_col : str

    Returns
    -------
    data : pd.DataFrame
        A modified DataFrame with a column labeled 'fog_flag'.
    """
    cpy = data.copy()

    dew_point_spread_threshold = kwargs.pop("dew_point_threshold", 3.75)
    pmc_threshold = kwargs.pop("pmc_threshold", 200.0)

    # Compute the dew point and dew point spread
    cpy.loc[:, "dew_point"] = cpy.apply(lambda r: dew_point_c(r[temp_col], r[rh_col]), axis=1)
    cpy.loc[:, "dew_point_spread"] = cpy[temp_col] - cpy["dew_point"]

    # Compute PMc
    cpy.loc[:, "PMc"] = cpy['pm10'] - cpy['pm25']

    # Add a flag column
    cpy["fog_flag"] = (
        (cpy["dew_point_spread"] < dew_point_spread_threshold) &
        (cpy["PMc"].rolling(11, center=True, min_periods=0).max() > pmc_threshold)
    )

    # Remove the added columns
    if kwargs.pop("remove_added_columns", True):
        for col in ("dew_point", "dew_point_spread", "PMc"):
            del cpy[col]

    return cpy
```

Figure 4. A python implementation of our recommended fog heuristic.

### Microsoft Excel Implementation

Please see the linked Excel workbook as an example. This uses minute-level data exported from the QuantAQ Cloud as a CSV and adds a fog flag (Column U) and supporting calculations (columns Q through T). Note that if you're auto-filling the **pmc_rolling** column, start from cell T7 – cells T1:T6 use a slightly different formula since there aren't enough rows before them for an eleven-cell window.

### Limitations of this approach and potential sources of error

The method proposed above is a crude first attempt and may be susceptible to errors (both false positives and false negatives). To be more conservative and avoid false negatives (i.e., flag more than needed), you can change the threshold values used to flag the data.

The biggest risk with this approach is to improperly flag data where there could be fog but also real dust. There are special locations and circumstances (e.g., a mine alongside the water) where this may be more likely. It is difficult to distinguish between dust and fog using the particle sensor alone.

### Future Research

QuantAQ is currently collecting data across a variety of foggy environments alongside trail cameras to conduct further research. We intend to develop a more robust approach for fog identification in the future.

## Citing this document

If you would like to reference this document, please use the citation format listed below. For more information, please visit the direct link on Zenodo.

> David McClosky & David H. Hagan. (2024). Identifying and Removing Data Records Influenced by Fog. (2024.03). https://doi.org/ 10.5281/zenodo.10793534

## References


(1) Hagan, D. H.; Cross, E. S. *(QAN 001) Introduction to the MODULAIR-PM*; 2024.02; Zenodo, 2024. https://zenodo.org/records/10688216 (accessed 2024-02-21).

(2) Hagan, D. H.; Kroll, J. H. Assessing the Accuracy of Low-Cost Optical Particle Sensors Using a Physics-Based Approach. *Atmospheric Meas. Tech.* **2020**, *13* (11), 6343–6355. https://doi.org/10.5194/amt-13-6343-2020.

(3) Nurowska, K.; Mohammadi, M.; Malinowski, S.; Markowicz, K. Applicability of the Low-Cost OPC-N3 Optical Particle Counter for Microphysical Measurements of Fog. *Atmospheric Meas. Tech.* **2023**, *16* (9), 2415–2430. https://doi.org/10.5194/amt-16-2415-2023.

(4) Crilley, L. R.; Shaw, M.; Pound, R.; Kramer, L. J.; Price, R.; Young, S.; Lewis, A. C.; Pope, F. D. Evaluation of a Low-Cost Optical Particle Counter (Alphasense OPC-N2) for Ambient Air Monitoring. *Atmospheric Meas. Tech.* **2018**, *11* (2), 709–720. https://doi.org/10.5194/amt-11-709-2018.

(5) Gultepe, I.; Zhou, B.; Milbrandt, J.; Bott, A.; Li, Y.; Heymsfield, A. J.; Ferrier, B.; Ware, R.; Pavolonis, M.; Kuhn, T.; Gurka, J.; Liu, P.; Cermak, J. A Review on Ice Fog Measurements and Modeling. *Atmospheric Res.* **2015**, *151*, 2–19. https://doi.org/10.1016/j.atmosres.2014.04.014.

(6) Liu, Q.; Wu, B.; Wang, Z.; Hao, T. Fog Droplet Size Distribution and the Interaction between Fog Droplets and Fine Particles during Dense Fog in Tianjin, China. *Atmosphere* **2020**, *11* (3), 258. https://doi.org/10.3390/atmos11030258.

(7) He, M.; Kuerbanjiang, N.; Dhaniyala, S. Performance Characteristics of the Low-Cost Plantower PMS Optical Sensor. *Aerosol Sci. Technol.* **2020**, *54* (2), 232–241. https://doi.org/10.1080/02786826.2019.1696015.

(8) Samad, A.; Melchor Mimiaga, F. E.; Laquai, B.; Vogt, U. Investigating a Low-Cost Dryer Designed for Low-Cost PM Sensors Measuring Ambient Air Quality. *Sensors* **2021**, *21* (3), 804. https://doi.org/10.3390/s21030804.


## Changelog

| | |
|---|---|
| 2024.03.07 | This is the first release of QAN 002. |