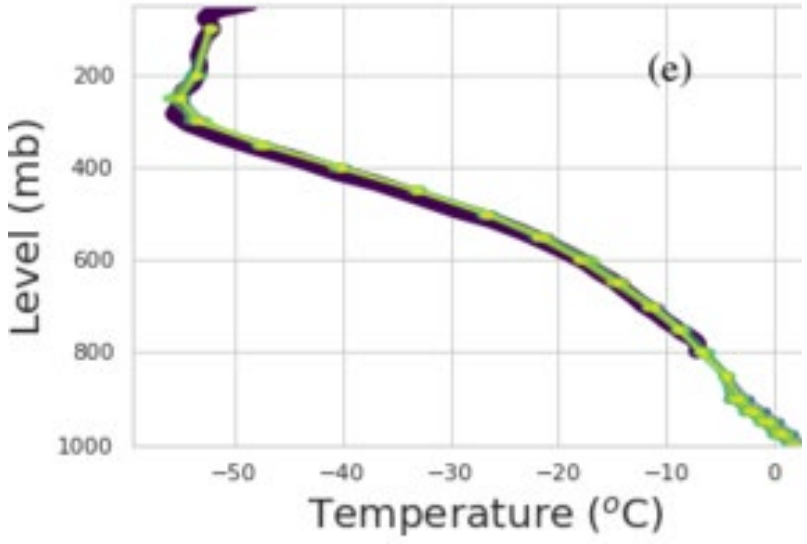
06/17/2022
HYSPLIT tutorial supplemental material
For sections 12 and 16

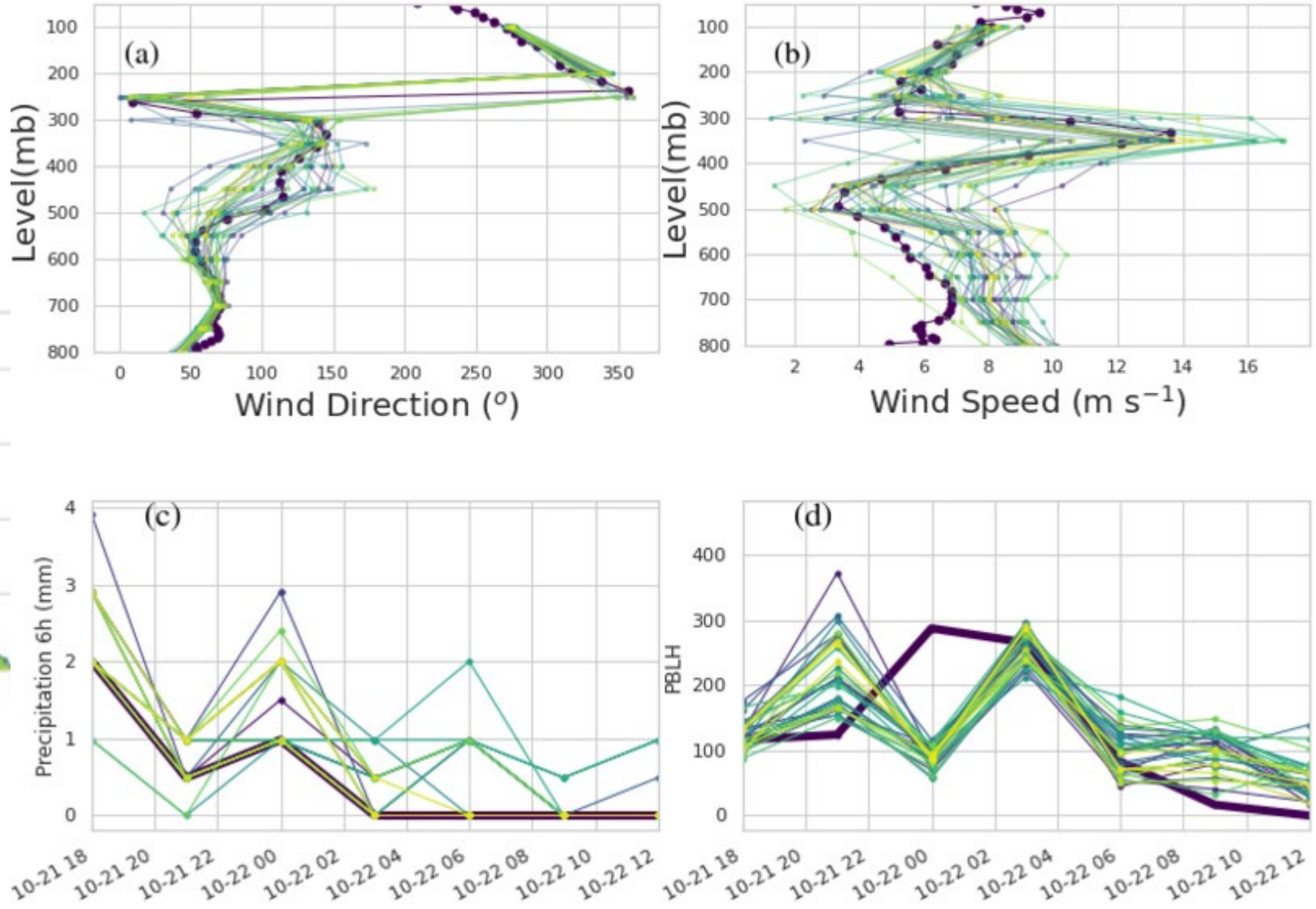
Contact: alice.crawford@noaa.gov

NWP ensemble datasets

- NOAA HREF high resolution ensemble forecast 10 members
- NOAA GEFS global ensemble forecast system 31 members. 0.25 degree.
- ECMWF ERA5 10 member (3h 63 km)
    - https://www.ecmwf.int/en/newsletter/159/meteorology/global-reanalysis-goodbye-era-interim-hello-era5
    - https://github.com/amcz/hysplit_metdata

Comparison of GEFS (thin lines) and GFS 0.25° (Thick line)
Near Bezymianny Volcano in Kamchatka
Profiles at 10 October 2020 at 21:00 UTC





From Crawford et. al.
Evaluation and Bias Correction Probabilistic Volcanic Ash Forecasts

https://doi.org/10.5194/egusphere-2022-290
Preprint. Discussion started: 11 May 2022

Conprob utility

- Ensemble relative frequency of exceedance that concentration >=
- 1% of maximum value
- 10% of maximum value
- Maximum value

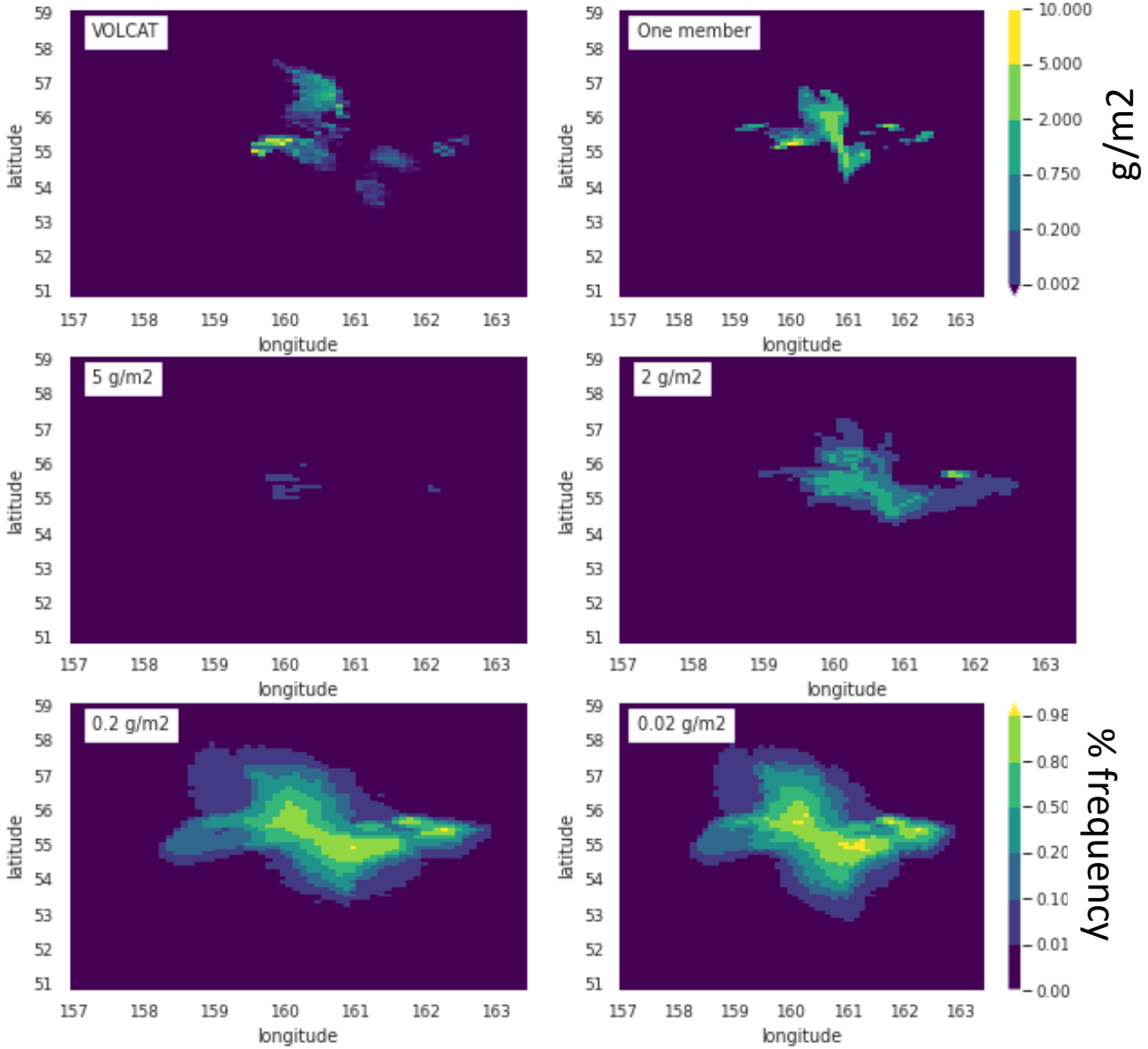
- Also called Agreement in Threshold Level (ATL)

Coefficient of Variation - relative standard deviation – standard deviation divided by mean

ATL Agreement in threshold level – Ensemble relative frequency of exceedance
Probability of concentration exceeding

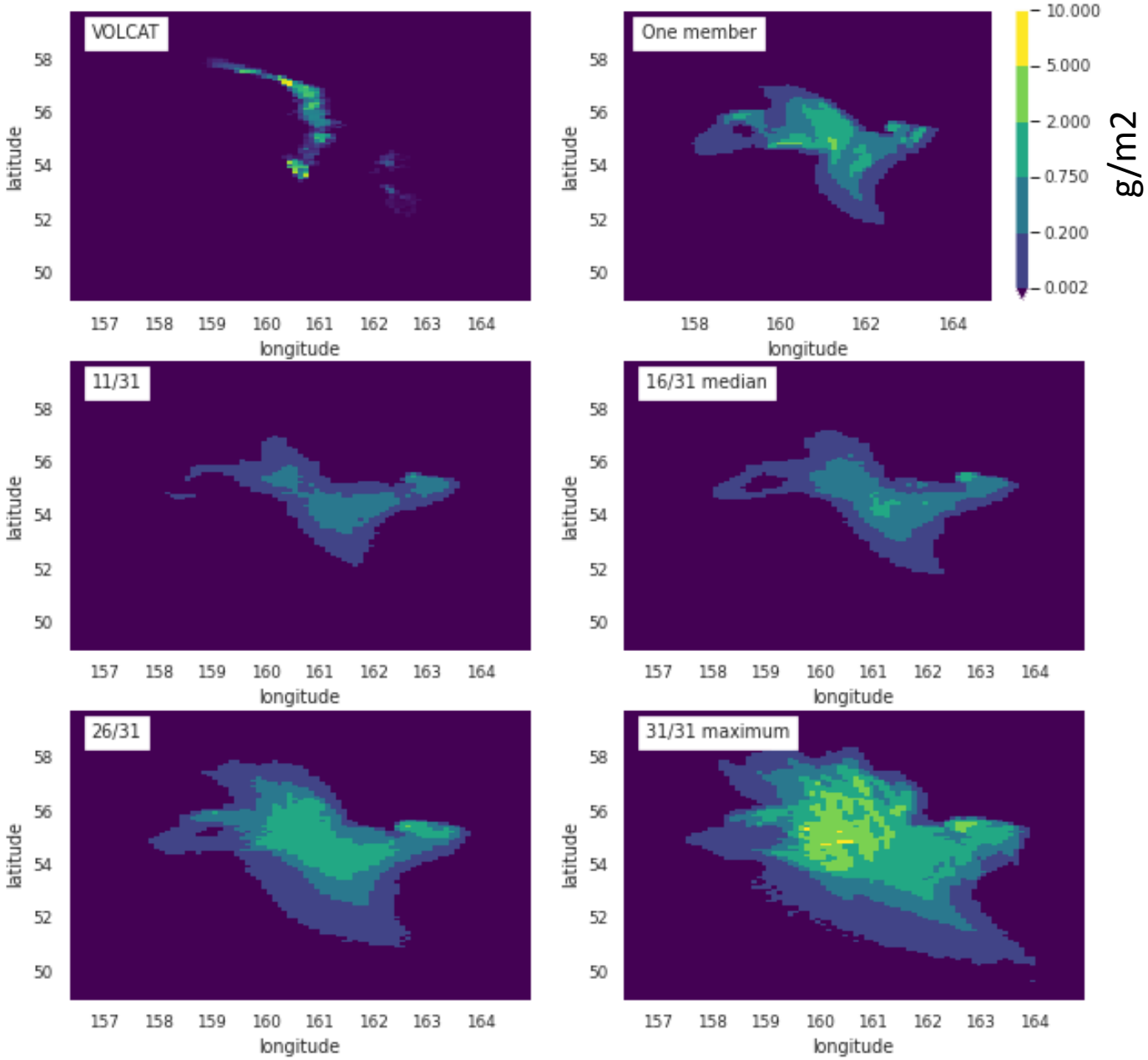
- Pick a concentration / mass loading threshold.
- Count how many ensemble members are above that threshold at each point.
- Each point becomes number of ensemble members above that threshold divided by total number of ensemble members.
- Values range from 0 to 100%.
- In order to use in risk based assessment, pick a probability threshold to decide which areas should be avoided.

All slides show a small/medium eruption of Bezymianny in Kamchatka in October 2020. HYSPLIT ensemble generated with GEFS and an inversion algorithm to determine source term. 31 members in ensemble.

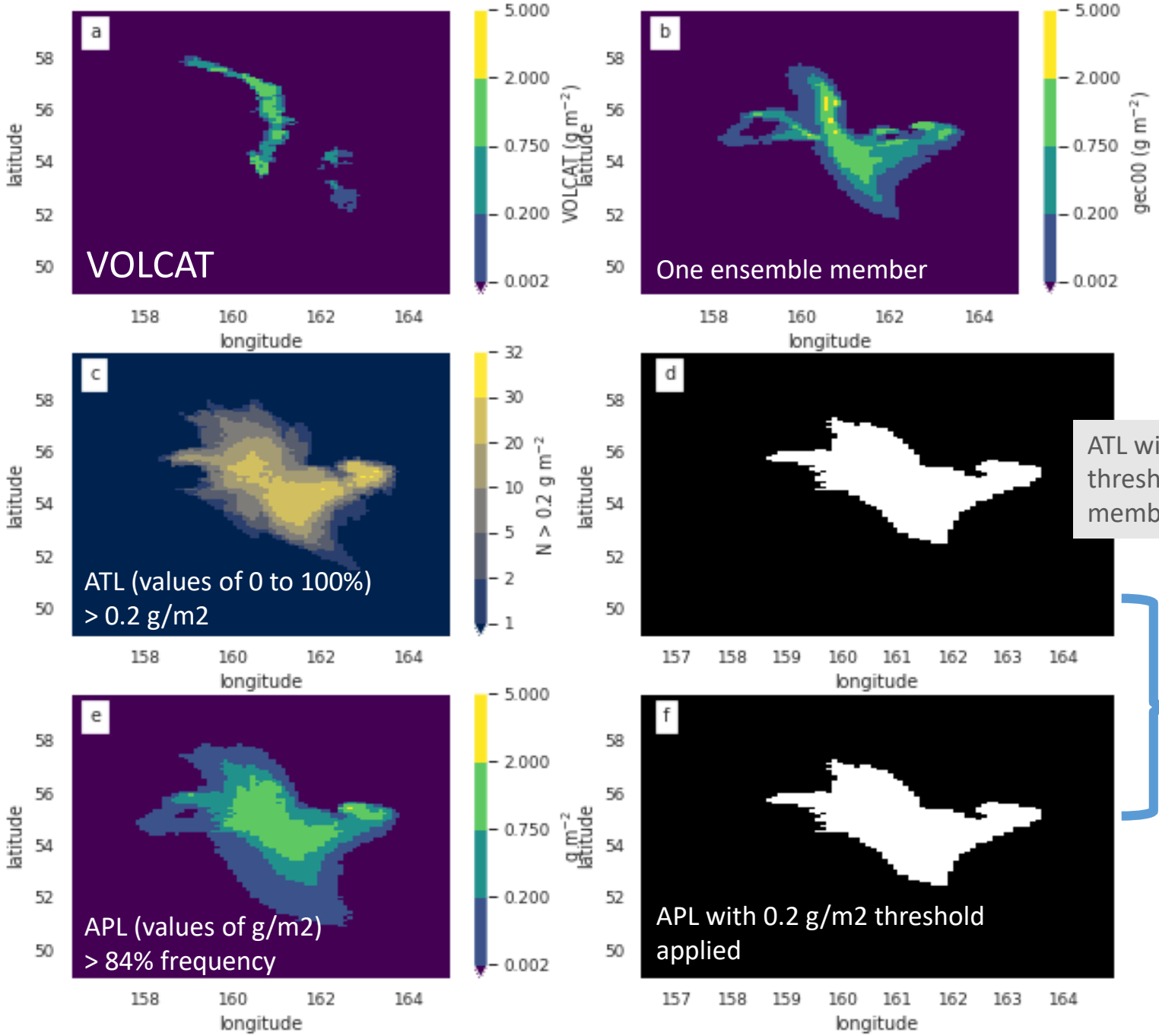


APL Agreement in Percentile Level – Concentration at percentile

- Pick a probability/frequency threshold , X%.
- Order values at each point from least to greatest and find value for which X% of ensemble members are smaller.
- Each point becomes concentration / mass loading for which X% of ensemble members are less than that value
- Values are in same units as individual ensemble members.
- APL at 50% is the ensemble median which is similar to ensemble mean but less sensitive to outliers.



How are ATL and APL related?



ATL with 16% probability threshold applied (> 5 members exceed 0.2mg/m3)

SAME

26 members had lower values than the value shown.

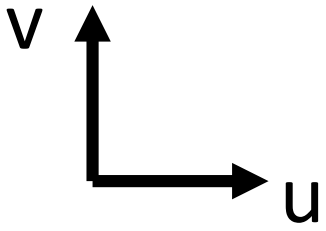
From Crawford et. al.

Evaluation and Bias Correction Probabilistic Volcanic Ash Forecasts

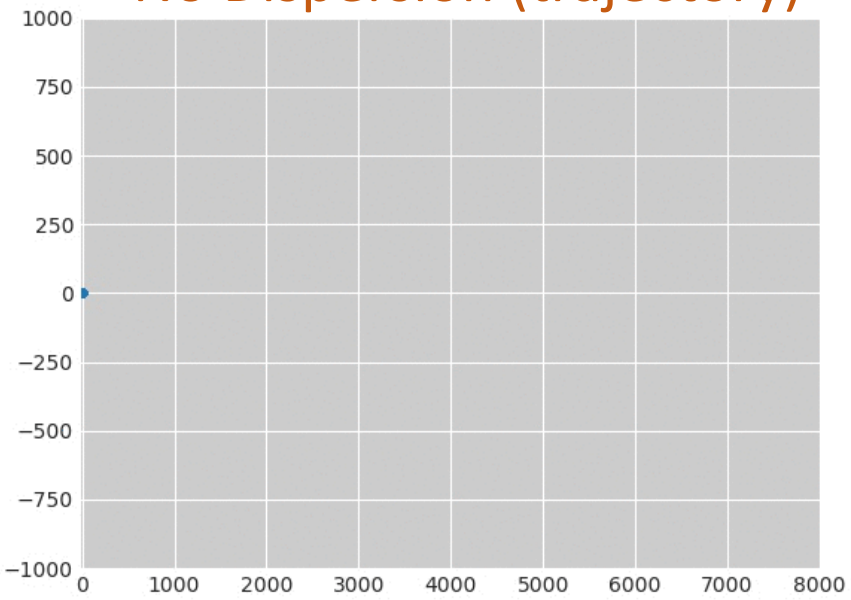
https://doi.org/10.5194/egusphere-2022-290
Preprint. Discussion started: 11 May 2022

2-D Dispersion example with toy model

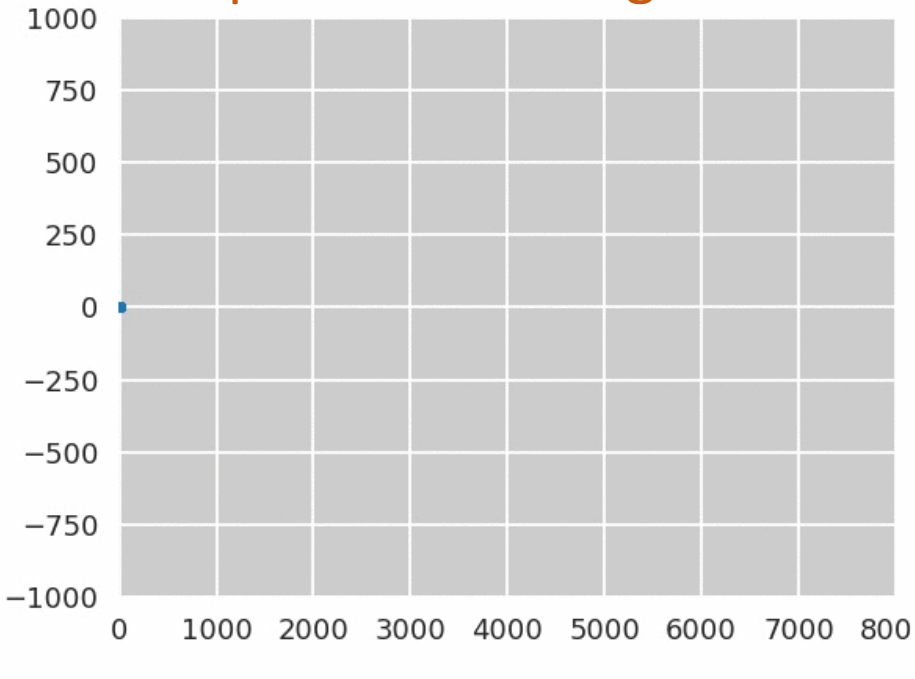
Mean wind 2 m/s in u direction
Instantaneous release of 100 particles from (0,0)



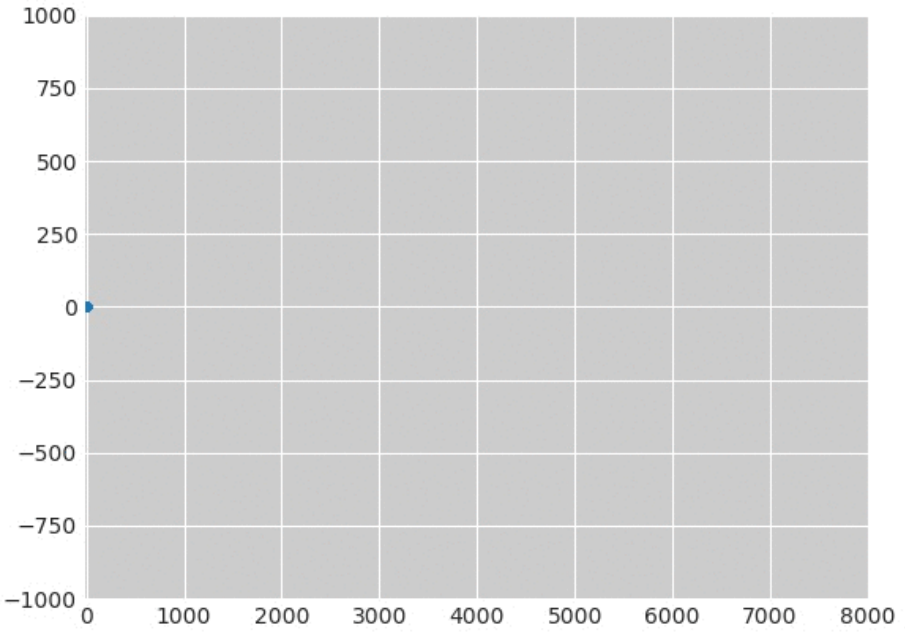
No Dispersion (trajectory)



Dispersion small sigma



Dispersion large sigma



Simplified Number of Particles?

Release 1 kg gram

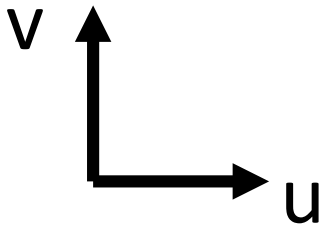
Consider :

Concentration grid size

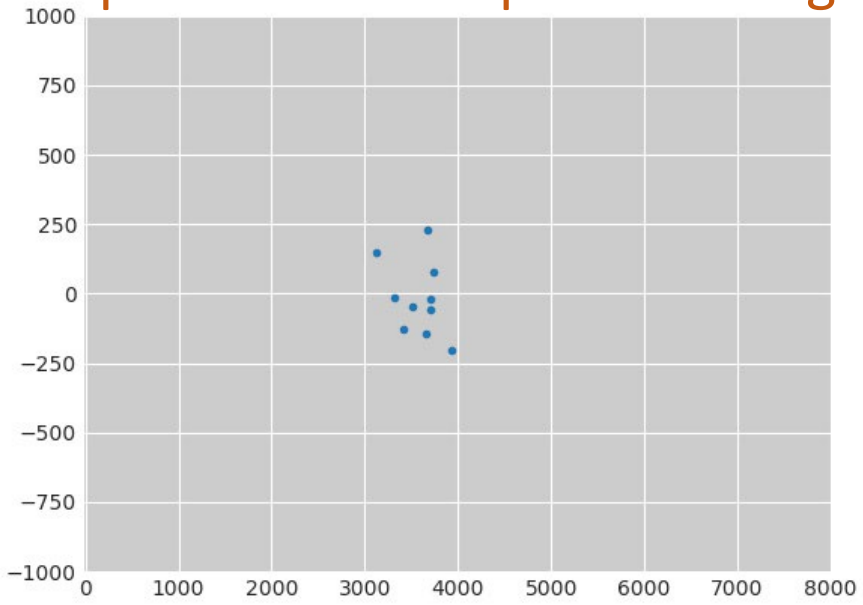
Averaging time

Length of time

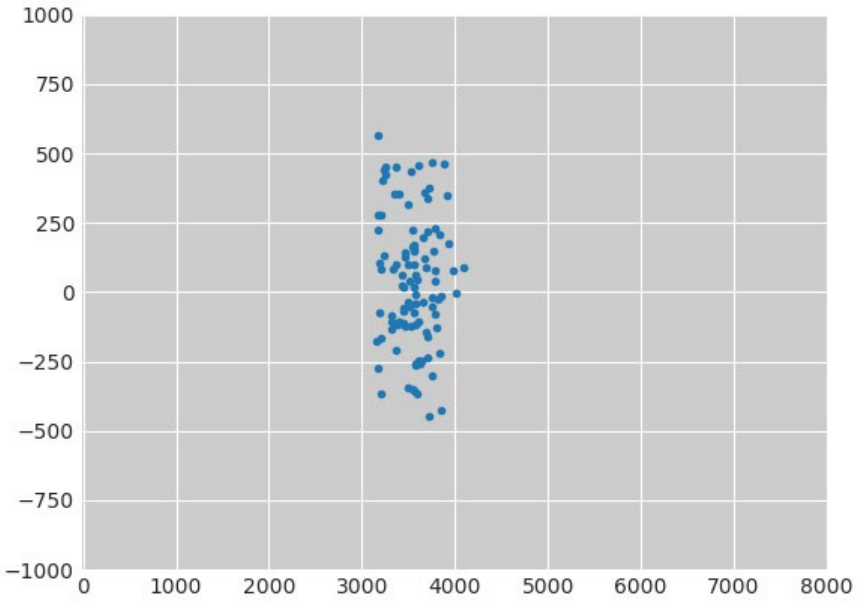
Minimum detection



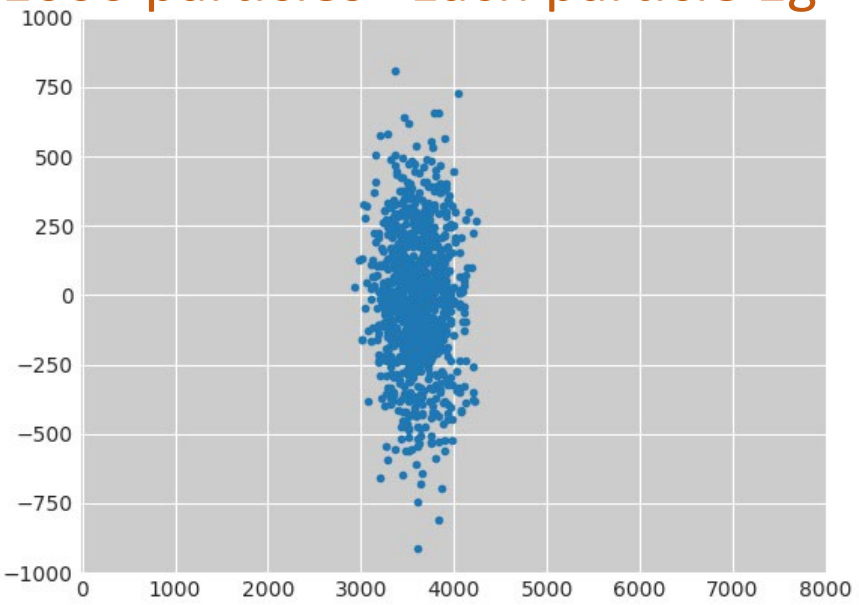
10 particles Each particle 100g



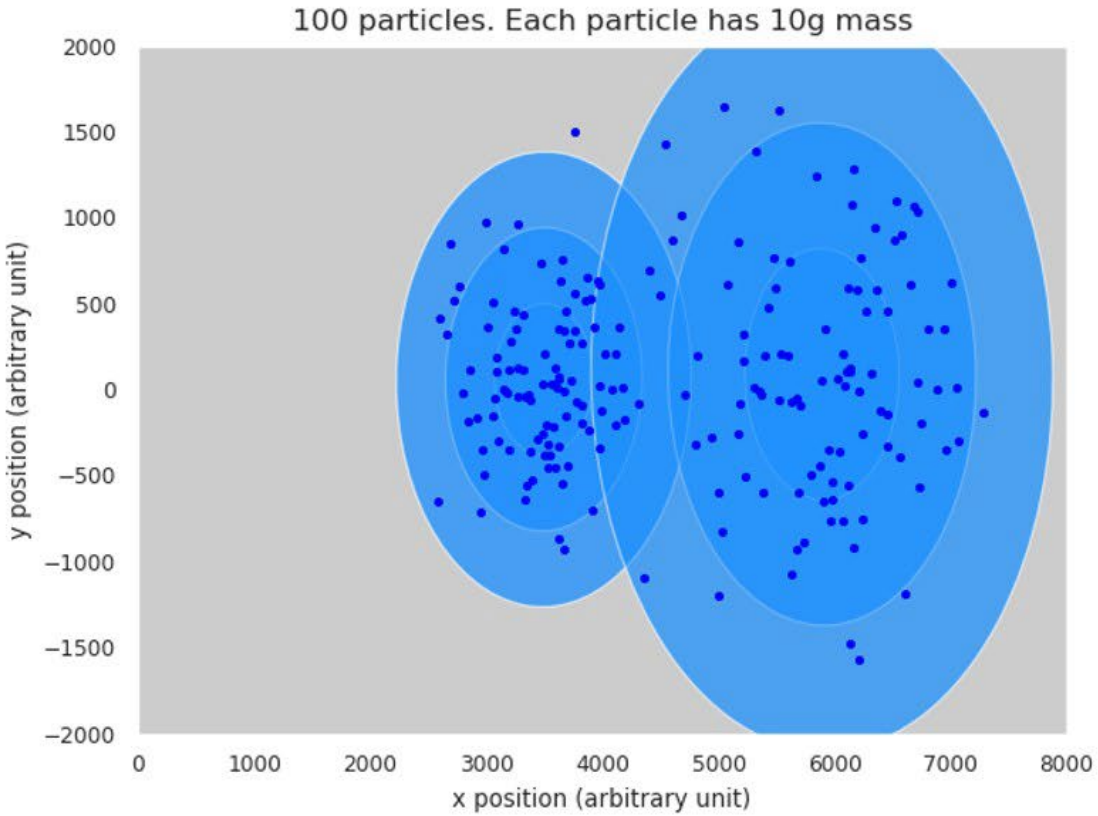
100 particles Each particle 10g

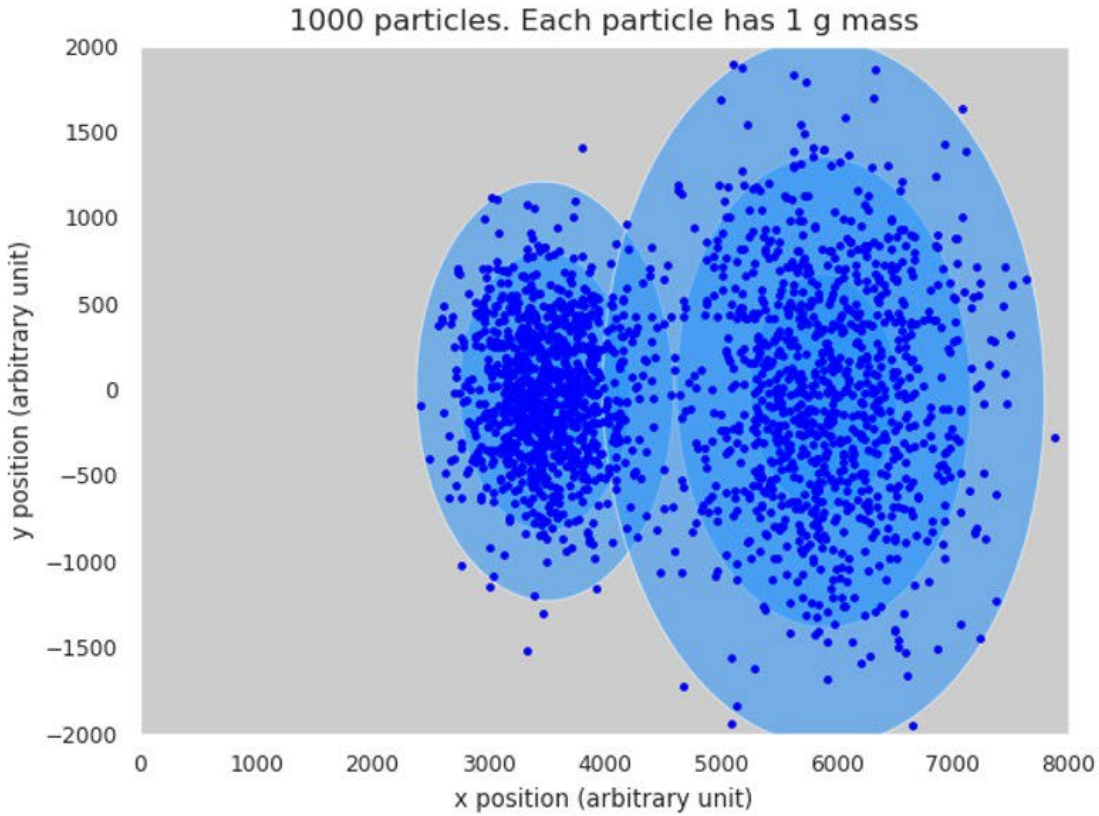


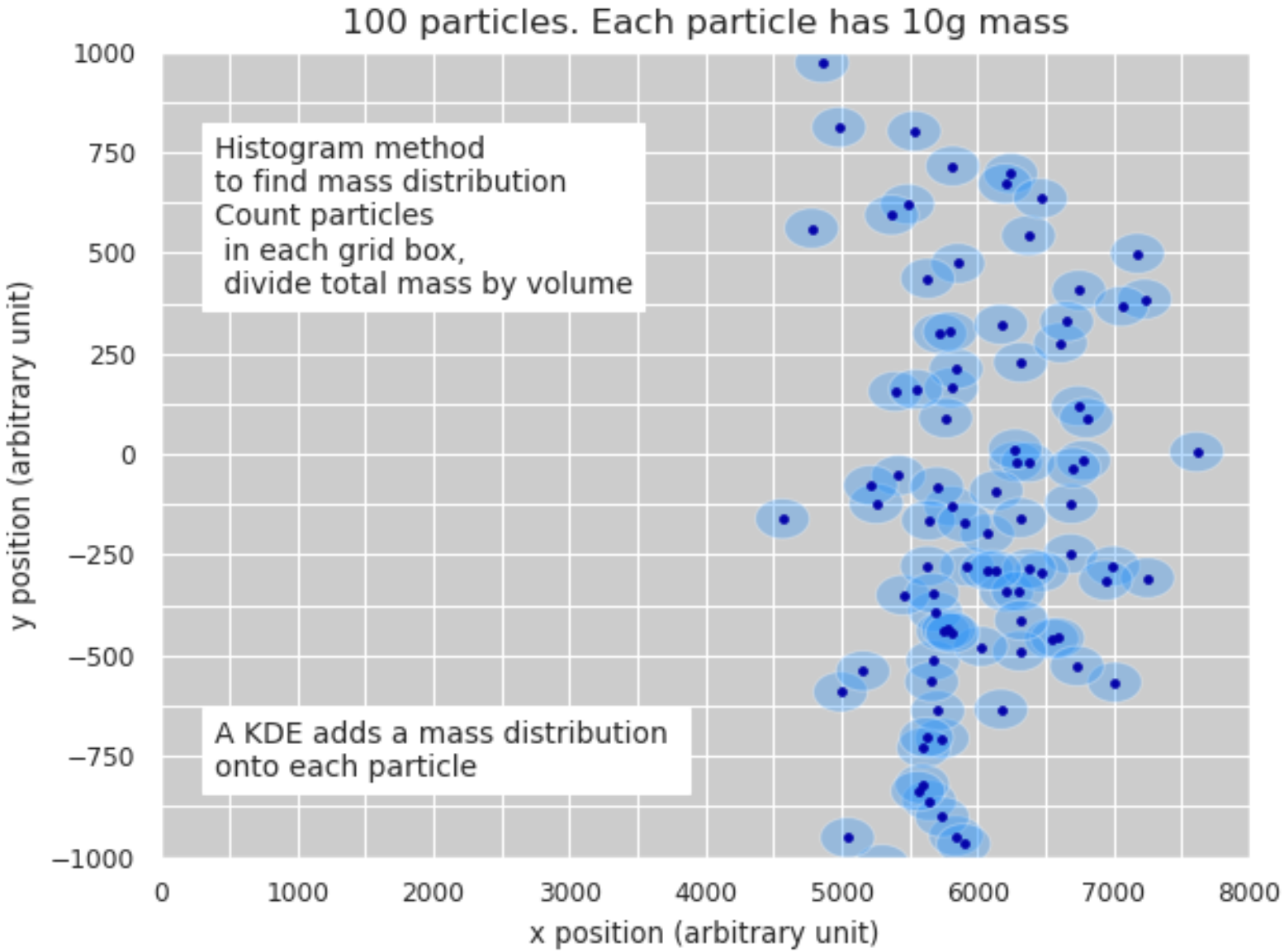
1000 particles Each particle 1g



Different numbers of particles representing distribution made up of two Gaussians.



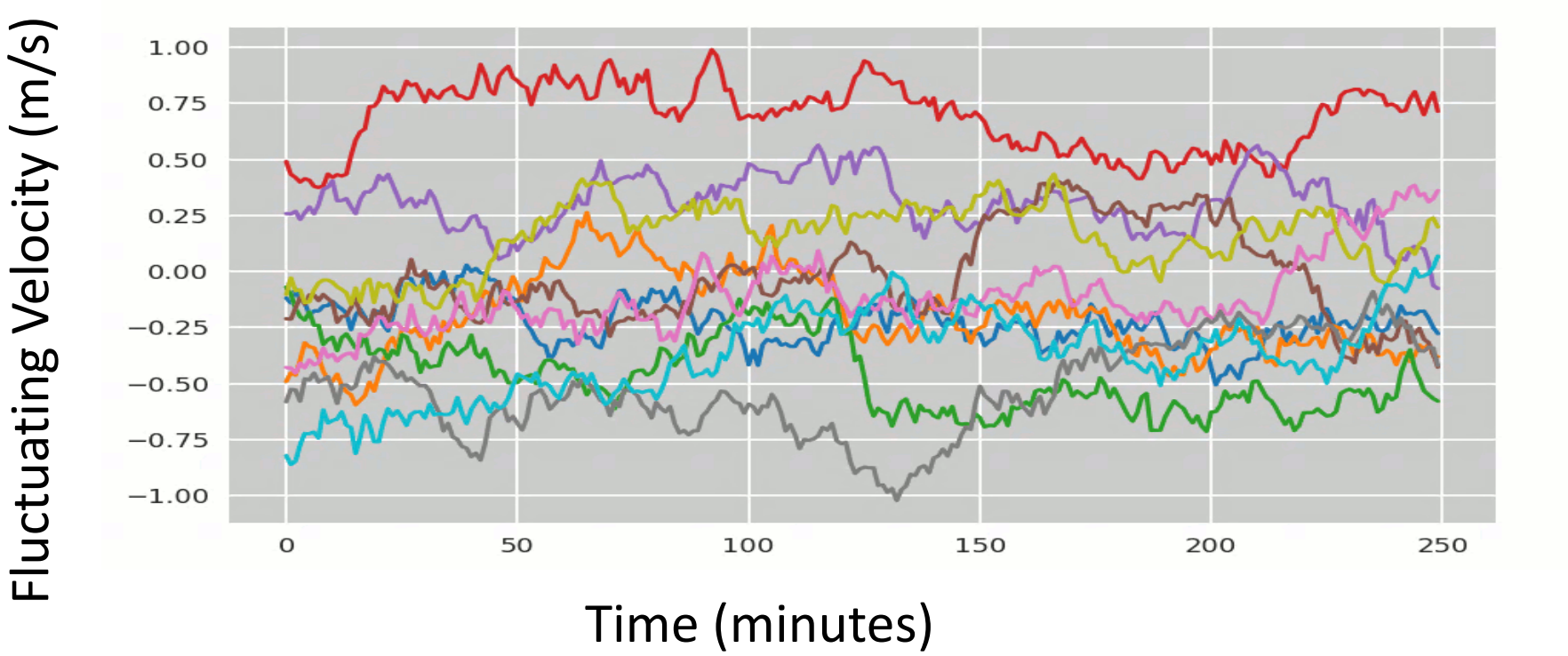




Guassian mixture model may also be used: https://www.mdpi.com/2073-4433/11/12/1369

Example: Fluctuating velocity for 10 particles.

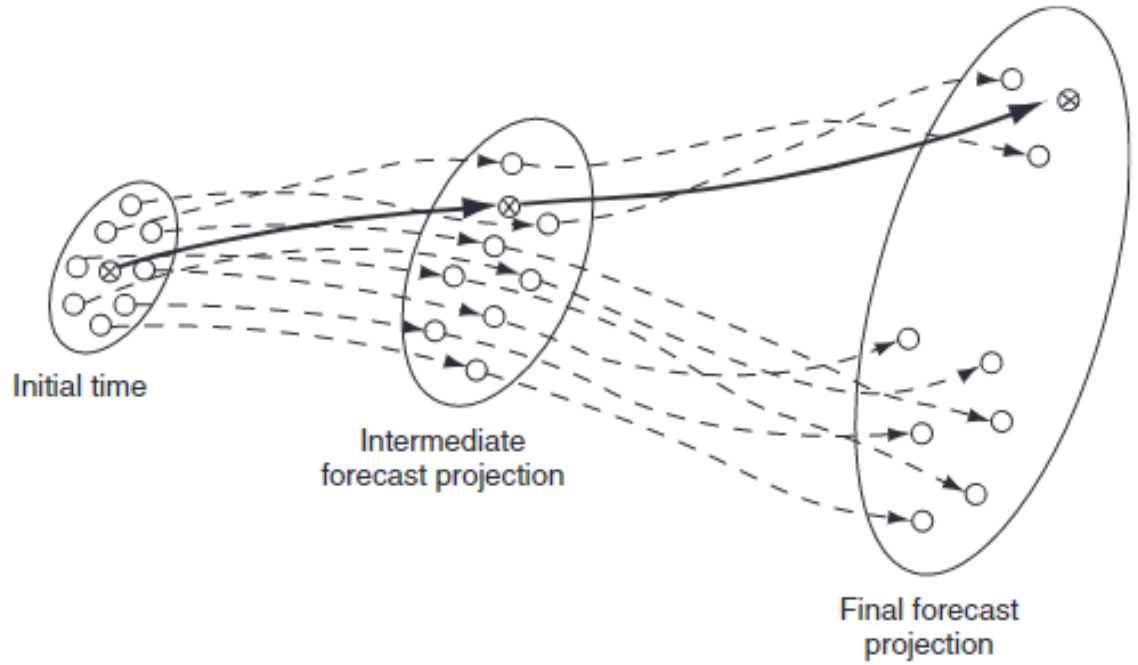
$$u(t+\Delta t) = R(\Delta t)\, u(t) + \sigma\, \lambda\left(1 - R(\Delta t)^2\right)^{0.5}$$



DNS and experiments show turbulent velocity PDF of fluid particles in a turbulent flow is very close to Gaussian

Velocity autocorrelation is represented well. Velocity is correlated over the Lagrangian time scale. (Time scale of largest scales in flow).

From: Statistical Methods in Atmospheric Sciences by Daniel S. Wilks



This is Volume 91 in the
INTERNATIONAL GEOPHYSICS SERIES

A series of monographs and textbooks
Edited by RENATA DMOWSKA, DENNIS HARTMANN, and H. THOMAS ROSSBY

FIGURE 6.24 Schematic illustration of some concepts in ensemble forecasting, plotted in terms of an idealized two-dimensional phase space. The heavy line represents the evolution of the single best analysis of the initial state of the atmosphere, corresponding to the more traditional single deterministic forecast. The dashed lines represent the evolution of individual ensemble members. The ellipse in which they originate represents the probability distribution of initial atmospheric states, which are very close to each other. At the intermediate projection, all the ensemble members are still reasonably similar. By the time of the final projection some of the ensemble members have undergone a regime change, and represent qualitatively different flows. Any of the ensemble members, including the solid line, are plausible trajectories for the evolution of the real atmosphere, and there is no way of knowing in advance which will represent the real atmosphere most closely.

- https://hysplitbbs.arl.noaa.gov/viewtopic.php?f=3&t=2313
- Stacked emissions example – one particle size.
- Each emission cycle MUST have same number of lines

```
YYYY MM DD HH    DURATION(hhhh) #RECORDS #spnum P001
YYYY MM DD HH MM DURATION(hhmm) LAT LON HGT(m) RATE(/h) AREA(m2) HEAT(w)
2020 10 21 19  0001 12
2020 10 21 19 00 0100 55.9780 160.5870 2880.0 0.00e+00 7.85e+05 0.00e+00
2020 10 21 19 00 0100 55.9780 160.5870 3880.0 0.00e+00 7.85e+05 0.00e+00
2020 10 21 19 00 0100 55.9780 160.5870 4880.0 0.00e+00 7.85e+05 0.00e+00
2020 10 21 19 00 0100 55.9780 160.5870 5880.0 0.00e+00 7.85e+05 0.00e+00
2020 10 21 19 00 0100 55.9780 160.5870 6880.0 1.54e+07 7.85e+05 0.00e+00
2020 10 21 19 00 0100 55.9780 160.5870 7880.0 0.00e+00 7.85e+05 0.00e+00
2020 10 21 19 00 0100 55.9780 160.5870 8880.0 3.01e+09 7.85e+05 0.00e+00
2020 10 21 19 00 0100 55.9780 160.5870 9880.0 1.22e+09 7.85e+05 0.00e+00
2020 10 21 19 00 0100 55.9780 160.5870 10880.0 2.52e+08 7.85e+05 0.00e+00
2020 10 21 19 00 0100 55.9780 160.5870 11880.0 0.00e+00 7.85e+05 0.00e+00
2020 10 21 19 00 0100 55.9780 160.5870 12880.0 0.00e+00 7.85e+05 0.00e+00
2020 10 21 19 00 0100 55.9780 160.5870 13880.0 0.00e+00 7.85e+05 0.00e+00
2020 10 21 20  0001 12
2020 10 21 20 00 0100 55.9780 160.5870 2880.0 1.15e+08 7.85e+05 0.00e+00
2020 10 21 20 00 0100 55.9780 160.5870 3880.0 0.00e+00 7.85e+05 0.00e+00
2020 10 21 20 00 0100 55.9780 160.5870 4880.0 4.16e+08 7.85e+05 0.00e+00
2020 10 21 20 00 0100 55.9780 160.5870 5880.0 1.26e+09 7.85e+05 0.00e+00
2020 10 21 20 00 0100 55.9780 160.5870 6880.0 0.00e+00 7.85e+05 0.00e+00
2020 10 21 20 00 0100 55.9780 160.5870 7880.0 1.84e+09 7.85e+05 0.00e+00
2020 10 21 20 00 0100 55.9780 160.5870 8880.0 3.36e+09 7.85e+05 0.00e+00
2020 10 21 20 00 0100 55.9780 160.5870 9880.0 3.94e+09 7.85e+05 0.00e+00
2020 10 21 20 00 0100 55.9780 160.5870 10880.0 1.15e+09 7.85e+05 0.00e+00
2020 10 21 20 00 0100 55.9780 160.5870 11880.0 1.38e+08 7.85e+05 0.00e+00
2020 10 21 20 00 0100 55.9780 160.5870 12880.0 6.14e+08 7.85e+05 0.00e+00
2020 10 21 20 00 0100 55.9780 160.5870 13880.0 0.00e+00 7.85e+05 0.00e+00
2020 10 21 21  0001 12
```

1 h emission cycle

Emissions of 1.54e7 from 6880 to 7880
Emissions of 0 from 7880 to 8880
Emissions of 3.0e9 from 8880 to 9880
Emissions of 1.22e9 from 9880 to 10880

1 h emission cycle

CONTROL file should have 12 location lines
And one species defined.

- https://hysplitbbs.arl.noaa.gov/viewtopic.php?f=3&t=2313
- Stacked emissions example – one particle size.
- Each emission cycle MUST have same number of lines

```
YYYY MM DD HH    DURATION(hhhh) #RECORDS #spnum P001
YYYY MM DD HH MM DURATION(hhmm) LAT LON HGT(m) RATE(/h) AREA(m2) HEAT(w)
2020 10 21 19  0001 12
2020 10 21 19 00 0100 55.9780 160.5870 2880.0 0.00e+00 7.85e+05 0.00e+00
2020 10 21 19 00 0100 55.9780 160.5870 3880.0 0.00e+00 7.85e+05 0.00e+00
2020 10 21 19 00 0100 55.9780 160.5870 4880.0 0.00e+00 7.85e+05 0.00e+00
2020 10 21 19 00 0100 55.9780 160.5870 5880.0 0.00e+00 7.85e+05 0.00e+00
2020 10 21 19 00 0100 55.9780 160.5870 6880.0 1.54e+07 7.85e+05 0.00e+00
2020 10 21 19 00 0100 55.9780 160.5870 7880.0 0.00e+00 7.85e+05 0.00e+00
2020 10 21 19 00 0100 55.9780 160.5870 8880.0 3.01e+09 7.85e+05 0.00e+00
2020 10 21 19 00 0100 55.9780 160.5870 9880.0 1.22e+09 7.85e+05 0.00e+00
2020 10 21 19 00 0100 55.9780 160.5870 10880.0 2.52e+08 7.85e+05 0.00e+00
2020 10 21 19 00 0100 55.9780 160.5870 11880.0 0.00e+00 7.85e+05 0.00e+00
2020 10 21 19 00 0100 55.9780 160.5870 12880.0 0.00e+00 7.85e+05 0.00e+00
2020 10 21 19 00 0100 55.9780 160.5870 13880.0 0.00e+00 7.85e+05 0.00e+00
2020 10 21 20  0001 12
2020 10 21 20 00 0100 55.9780 160.5870 2880.0 1.15e+08 7.85e+05 0.00e+00
2020 10 21 20 00 0100 55.9780 160.5870 3880.0 0.00e+00 7.85e+05 0.00e+00
2020 10 21 20 00 0100 55.9780 160.5870 4880.0 4.16e+08 7.85e+05 0.00e+00
2020 10 21 20 00 0100 55.9780 160.5870 5880.0 1.26e+09 7.85e+05 0.00e+00
2020 10 21 20 00 0100 55.9780 160.5870 6880.0 0.00e+00 7.85e+05 0.00e+00
2020 10 21 20 00 0100 55.9780 160.5870 7880.0 1.84e+09 7.85e+05 0.00e+00
2020 10 21 20 00 0100 55.9780 160.5870 8880.0 3.36e+09 7.85e+05 0.00e+00
2020 10 21 20 00 0100 55.9780 160.5870 9880.0 3.94e+09 7.85e+05 0.00e+00
2020 10 21 20 00 0100 55.9780 160.5870 10880.0 1.15e+09 7.85e+05 0.00e+00
2020 10 21 20 00 0100 55.9780 160.5870 11880.0 1.38e+08 7.85e+05 0.00e+00
2020 10 21 20 00 0100 55.9780 160.5870 12880.0 6.14e+08 7.85e+05 0.00e+00
2020 10 21 20 00 0100 55.9780 160.5870 13880.0 0.00e+00 7.85e+05 0.00e+00
2020 10 21 21  0001 12
```

This could be setup differently with Emission cycles covering different amounts Of time.

Example – 2 h emission cycles could be setup. Then the two cycles shown here would be combined and there would be 24 lines per emission cycle.

CONTROL file would need 24 location lines.

NUMPAR is number of particles emitted per Emission cycle.

EMIT file for multiple particle sizes

If you have two pollutants defined in your CONTROL file then in the EMITIMES file you need to have one line per pollutant. If one of the pollutants has 0 emissions for that time period then you need to add a dummy line with 0 emissions. See below. For this case you should have 4 start locations and two species defined in the CONTROL file.

The start locations and number of species in the CONTROL file are used to define how big some arrays should be. If the array sizes defined when the CONTROL file is read don't match what is needed by the information in the EMITIMES file, you will end up with segmentation faults.

CODE: SELECT ALL

```
YYYY MM DD HH DURATION(hhhh) #RECORDS
YYYY MM DD HH MM DURATION(hhmm) LAT LON HGT(m) RATE(/h) AREA(m2) HEAT(w)
1981 05 09 00 2688 8
1981 05 09 00 00 0600 39.024 -120.365 0.0 2.120 40000 12409.83  # species one
1981 05 09 00 00 0600 39.024 -120.365 0.0 0 .000 40000 12409.83  # species two
1981 05 15 00 00 0600 38.772 -120.529 0.0 2.766 40000 14753.93  # species one
1981 05 15 00 00 0600 38.772 -120.529 0.0 5.000 40000 14753.93  # species two
1981 06 07 00 00 1100 38.822 -119.997 0.0 0.000 40000 11580.51  # species one
1981 06 07 00 00 1100 38.822 -119.997 0.0 2.532 40000 11580.51  # species two
1981 08 27 00 00 0600 39.058 -120.285 0.0 2.500 40000 13278.11  # species one
1981 08 27 00 00 0600 39.058 -120.285 0.0 1.775 40000 13278.11  # species two
```

Number of lines = Number of location lines in CONTROL file
X Number of species in CONTROL file.

CONTROL file for this run should have 4 locations.