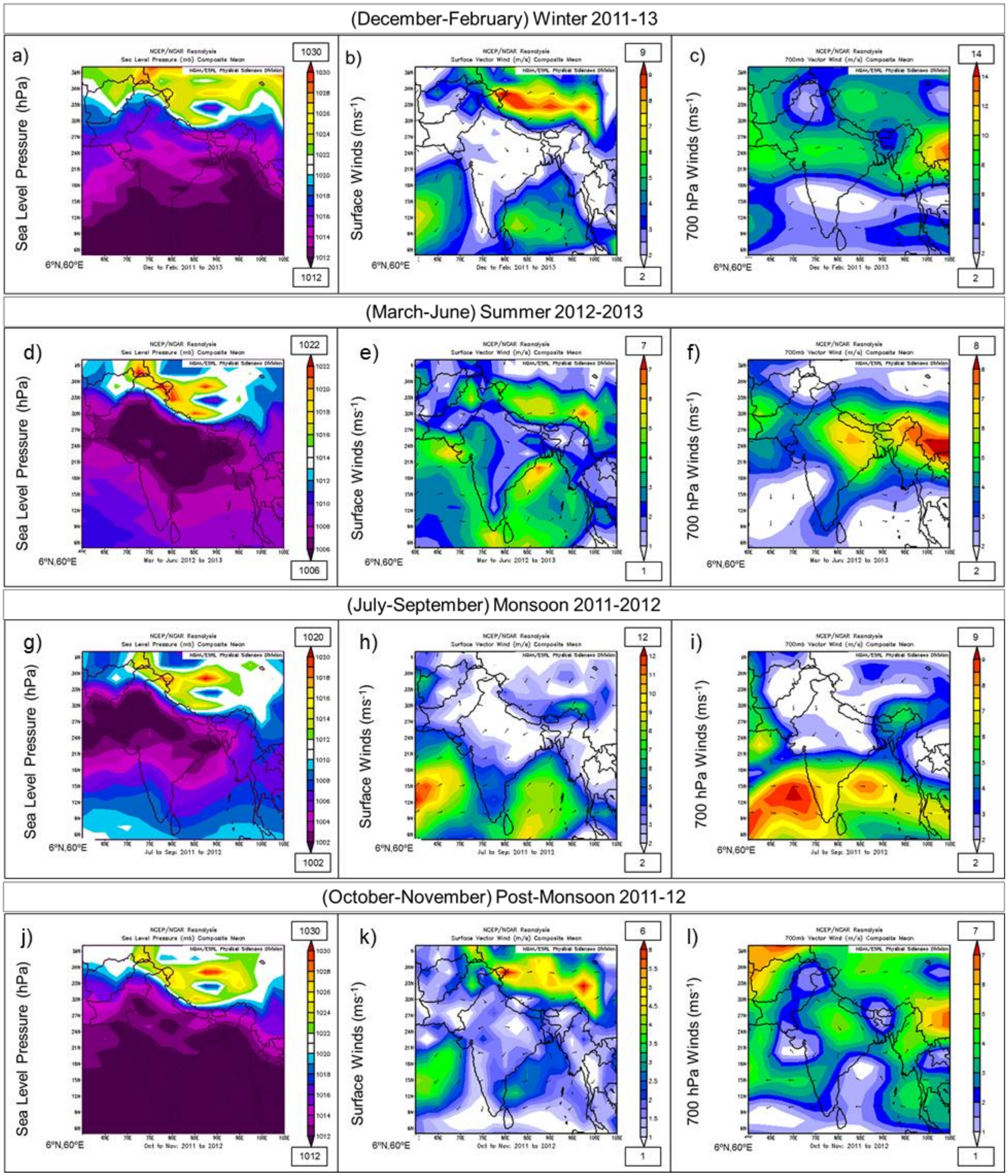
Reviewer comment:
This manuscript describes the measurements of particulate matter over two years at a site on the Indo Gangetic Plain in India, and the interpretation of those measurements using clusters of atmospheric back trajectories. The manuscript contains some important data quantifying the contributions of long range transport to the air quality exceedances observed in this important and understudied part of the world. These are my comments:
Authors' response:
We thank reviewer #1 for her/his helpful comments which have improved the clarity of the presentation further.

Reviewer comment:
Page 11412 line 28: The wording used in this sentence is repeated on page 11413 line 6. Please rephrase.
Authors' response: We have rephrased Page 11412 line 28
Modification in text:
"Statistical analysis of large datasets–has been a popular tool for identifying source regions of particulate matter..."

Reviewer comment:
Page 11415 to 11417– The description of the meteorology of the region is very long and hard to follow without an accompanying figure. Please consider shortening this section and replacing the text with a diagram in the supporting information that shows the seasonality of the prevailing winds and pressure gradients.
Authors' response: We understand that the text is difficult to follow without accompanying figures. We have, therefore, included a figure showing sea level pressure, surface winds and 700 hPa winds for each season in the Supplement. (Supplementary Fig. 1) using the NCEP Reanalysis Derived data provided by the NOAA Physical Science Division, Boulder, Colorado, USA, from their web site at http://www.esrl.noaa.gov/psd.
However, it should be noted, that this is the first paper from IISER Mohali Atmospheric Chemistry Facility which presents measurements from all seasons. Therefore, a detailed description of the meteorology of the site is warranted. Moreover, there are substantial discrepancies regarding the surface winds, mostly pertaining to the winds speed, between the model output and our measured wind speed. The measured wind speeds (Table 1) are much higher than the modelled wind speeds. In fact the measured wind speed and wind direction close to the surface is frequently better represented by the 700 hPa winds, than it is by the surface or 850 hPa winds of the model output. This problem is not unique to this particular model and meteorological dataset but also applies to e.g. Figure 6 in Lawrence and Lelieveld., (2010). We, therefore, consider a detailed description of the meteorology necessary and have revised the text such that it now refers to and discusses the new figure in Supplement 1.
Modification in text:
P11415 line 12 after "During winter season, weak northerlies or north-westerlies and a weak, low-level anti-cyclonic circulation prevails in the NW-IGP." We have inserted "The surface pressure map, the surface winds and 700 hPa winds of NCEP Reanalysis derived data provided by the NOAA Physical Science Division, Boulder, Colorado, USA are shown in supplementary Figure 1 panel a) to c)."
P11415 line 14 after "Wintertime fog occurs frequently and is favoured by subsidence of air masses over the IGP, low temperatures, high relative humidity and low wind speeds (5 ms^{-1})" we inserted "However, ground level wind speeds at our site are generally not as low as the surface wind speeds of the NCAR Reanalysis dataset would suggest. The seasonally averaged experimentally observed wind speed is 4.4 ms^{-1} and not < 2 ms^{-1}. This underestimation of the surface wind speeds in the NW-IGP is not unique to this particular model and meteorological dataset but also applies to e.g. Figure 6 in Lawrence and Lelieveld (2010)."
P11416 line 3 after "These loo-winds are extremely hot and dry" we inserted "and are not adequately resolved by the NCEP Reanalysis meteorological dataset (Supplementary Figure 1, panel d-f), which shows a very moderate average surface wind speed of 1.5-2.5 ms^{-1} over the NW-IGP that stands in stark contrast to the observed average wind speed of 5.6 ms^{-1}."



Supplementary Figure 1: Seasonal composites of sea level pressure, surface winds and 700hPa winds. The values on the top and bottom of the legend in each figure shows the maximum and minimum respectively. (Images provided by the NOAA-ESRL Physical Sciences Division, Boulder Colorado from their Web site at http://www.esrl.noaa.gov/psd/)

P11417 line 3 after "Break spells are associated with lower-tropospheric inversions, dusty winds and lower troposphere anti-cyclonic vorticity over the IGP (Sikka, 2003; Rao and Sikka, 2005; Bhat, 2006)." We inserted: "Both the observed wind direction and wind speeds of the surface winds during monsoon season are poorly resolved by the meteorological dataset. While the model suggests low wind speed (2 ms^{-1}) where south westerly winds dominate (Supplementary Figure 1, panel g-i), actual observations show that north westerly winds dominate during break spells and south easterly winds during active spells. The average wind speed is 5.1 ms^{-1}."

P1417 line 9 after "Winds are generally weak; the wind speed is less than 5ms^{-1} for more than 80% of the time." We inserted: "Post monsoon season shows the lowest discrepancy between modelled surface winds 1.5-2 ms^{-1} (Supplementary Figure 1, panel j-l) and observed average wind speeds 3.4 ms^{-1}. Surface wind vectors represent the observed night time flow at our site better, while the 700 ha Pa wind vectors represent daytime observations better. "

Reviewer comment: Figure 5 seems to provide an unnecessary level of detail.
Authors' response: While we agree that Figure 5 provides an extra dimension of detail, it is still relevant to the discussion. We have shifted Figure 5 to the Supplement. (Supplementary Figure 3)
Modifications in the text: Shifted Fig. 5 to the Supplement and modified the references to this figure in the text accordingly.

Reviewer comment: The paper is very long and details about the clusters might be best left to the supplemental.
Authors' response: We have shifted this information to the supplement.

Reviewer comment: There are too many figures in general. The number of figures should be less than 10.
Authors' response: We shifted Fig. 4 and Fig. 5 to the Supplement as Supplementary Figure 2 and Supplementary Figure 3 respectively. The information shown in figure 6 and figure 7 has now been shown on only one image and the text has been modified in accordance with this change. The total number of figures in the paper is now 9.

Reviewer comment: One suggestion is to add the average trajectory shown in Figure 6 onto the plot of all trajectories in Figure 7. The average trajectory could be shown by a black line amongst all the individual trajectories for each cluster.
Authors' response: We thank the reviewer for the suggestion. We have incorporated the average cluster trajectory onto the individual trajectories.

Reviewer comment: Figure 11: It is customary to include p values when presenting linear regression statistics. Also, the use of different colors in this figure is not explained. Also two regression equations are given. Why two? Figure caption needs more explanation.
Authors' response: We have only shown one regression line. The coefficient "a" represents the intercept and "b" represents the slope in the equation: $PM_{2.5}$ = a + b × [Specie] where the specie maybe CO, NO_2, benzene or acetonitrile depending on the plot. With each plot, we have provided slope, error in slope and intercept and error in intercept.
We thank the reviewer for pointing it out that the information pertaining to what the coefficients "a" and "b" stand for in each of the linear regression plot was missing from the figure caption and we have included the same in the figure caption.
We did not mention p value rather we presented r^2 values as a goodness of fit parameter as is customary. Nevertheless, we did check the p value. It was less than 0.05 in each fit thus rendering the null hypothesis false.
We thank the reviewer for highlighting the missing colour index in the figure. The different colours represent different clusters.
Modifications in text: After incorporating the above discussed changes the figure caption is as follows:
"Figure 11: Scatter plots of fine mode PM with CO, NO2, acetonitrile and benzene and for winter, summer, monsoon and post-monsoon season. "a" stands for intercept , "b" stands for slope in the linear regression equation."

Reviewer comment: Figure 12: Caption includes the phrase: "during the daytime/nighttime low" which "low" is being referred to here?
Authors' comment: We thank the author for highlighting the missing information. The daytime and nighttime lows refer to the two minima observed in the diel-profile of coarse and fine mode PM. (Fig. 3); daytime: 12:00 to

16:00 LT and nighttime: 03:00 to 06:00 LT. We have now elaborated on the daytime/nighttime low in Figure 3 and its caption and lines 10-13, page 11417. The modifications have been shown below.

Modifications in text:

Modified Figure 3 and its caption

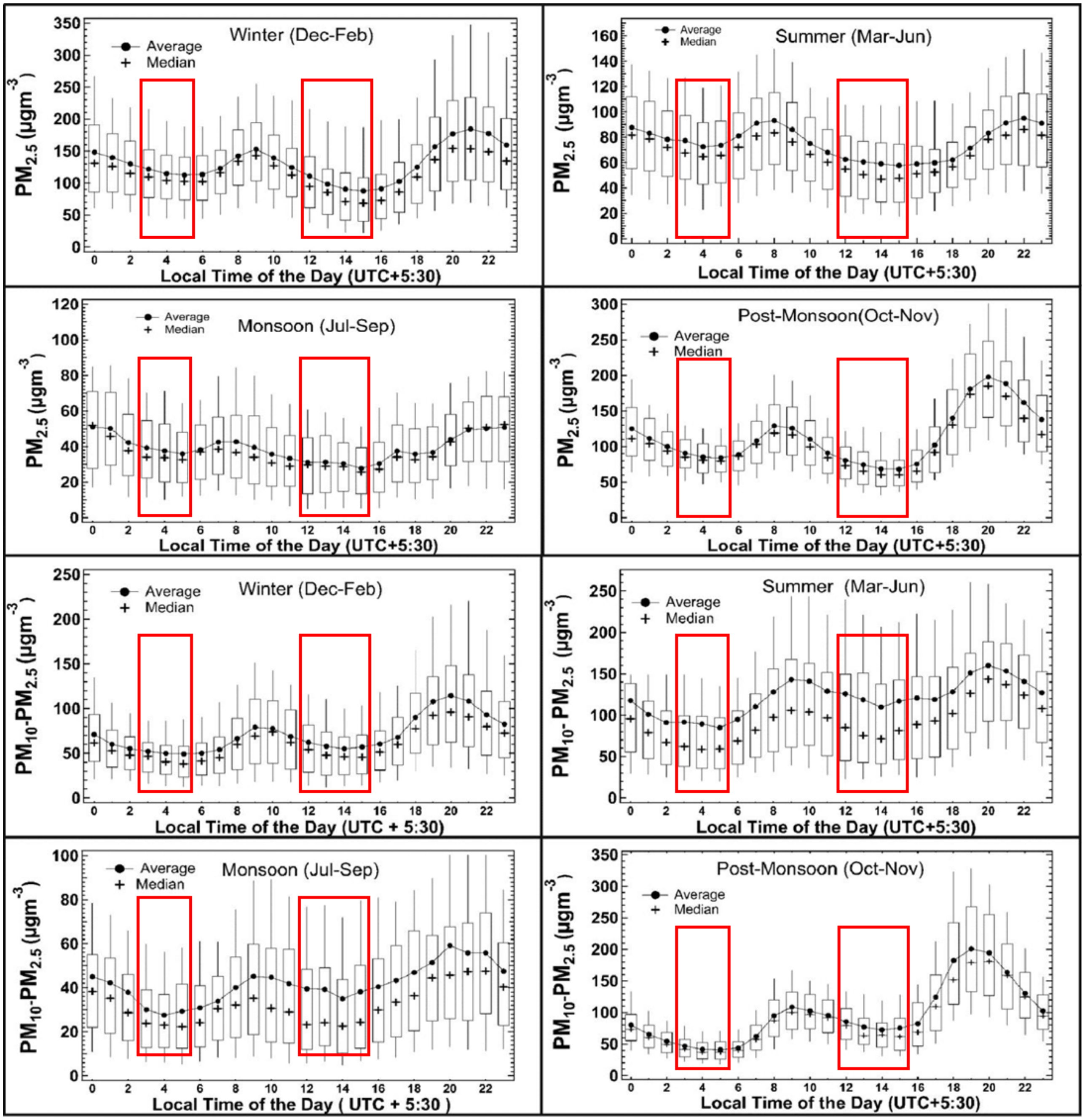


Figure 3: Diel box and whisker plots for fine mode (top four panels) and coarse mode (bottom four panels) particulate matter for winter, summer, monsoon and post monsoon season for the period November 2011 to August 2013 respectively. The box indicates the upper and lower quarter value; the cross indicates the median and the dots connected by lines provide the mean. The whiskers indicate the 5th and 95th percentile respectively. Periods of calm (WS <$1 ms^{-1}$) have been excluded while preparing the graph. The interval highlighted in red shows the daytime low (12 to 4 p.m. LT) and nighttime low (3 to 6 a.m. LT).

Modified lines 10-13, page 11417

"Since the purpose of this study is to investigate the contribution of long range transport to PM pollution we restricted our analysis to measurements obtained between 12 noon to 4 p.m. LT (UTC+05:30) during the day and 3 to 6 a.m. LT (UTC+05:30) at night as depicted in Fig. 3 and consider calm conditions with wind speeds of less than $1 ms^{-1}$ separately."

Reviewer comment: Also, it would be nice to know how many exceedance days there are in each season in addition to the percentages listed for each cluster.

Authors' comment: We thank the reviewer for asking this information as this prompted us to re-evaluate the number of exceedance days season-wise and while doing so we realized that using the average of the 4 hour time window of the minima instead of the 24 hour average while calculating the number of exceedance days is strictly speaking not correct. We have now used 24 hour averages and have included the information regarding the number of exceedance days in the text. We have also corrected the figure and the corresponding text. While the percentage of exceedance days may have changed, the main conclusion that the long range transport does not control % of exceedance days remains unchanged.

Modifications:

Modified Figure 12 and its caption with revised numbers and corrected the corresponding numbers in the text

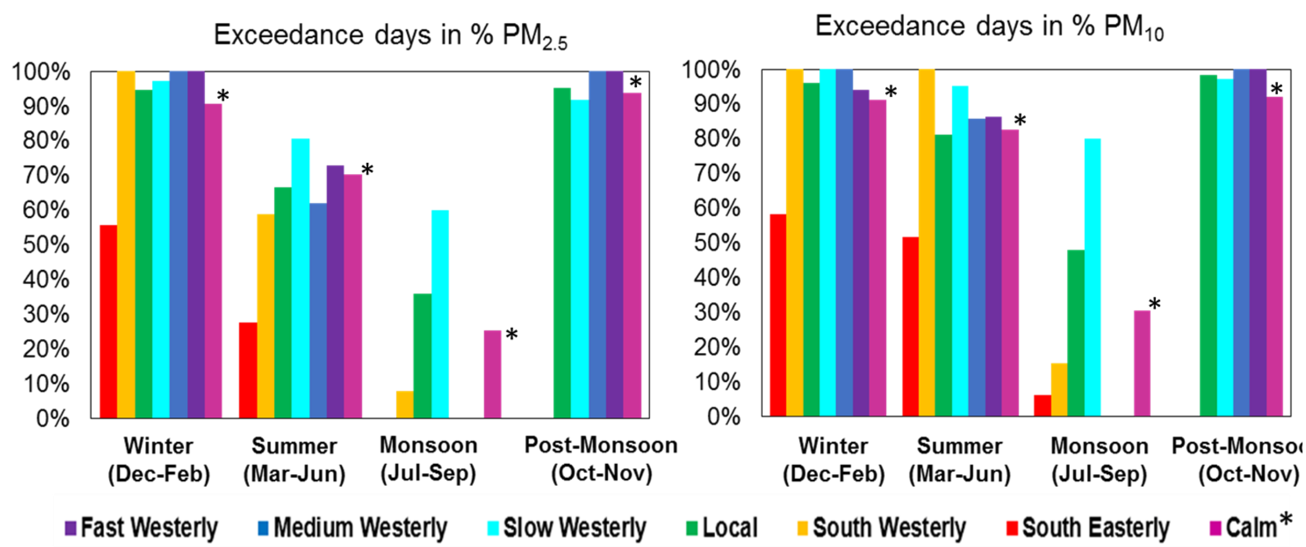


Figure 12. Percentage of days where the 24 hour average $PM_{2.5}$ or PM_{10} mass loading (in μgm^{-3}) exceeds the national ambient air quality standard for each air mass cluster and season.

* For calm conditions the $PM_{2.5}$ or PM_{10} were averaged for all times during a 24 hour interval that had WS < $1 ms^{-1}$ only. The fraction of exceedance days is calculated based on this average, and not a genuine 24 hour average, as wind speeds do not remain low continuously.

Modified section 3.4 (page 11436-11439) to include % of exceedance days:

"3.4 Impact of air mass transport on Particulate Matter (PM) exceedance events

The mean PM mass loadings of an air mass cluster represent a poor proxy for the number of exceedance days that is the number of days on which 24 hour average PM_{10} or $PM_{2.5}$ exceeded the NAAQS of 100 μgm^{-3} or 60 μgm^{-3} respectively. During winter season on ~160 days (out of 180 days) NAAQS of both PM_{10} and $PM_{2.5}$ was exceeded. For summer season, exceedance days associated with PM_{10} were more frequent (203 out of 243 days) that those associated with $PM_{2.5}$ (157 out of 243). During monsoon season, the receptor site received cleaner air masses with only 13 and 21 exceedance days (out of 92) of $PM_{2.5}$ and PM_{10} respectively. Post-monsoon season had frequent exceedance days with NAAQS of PM_{10} and $PM_{2.5}$ being exceeded ~ 110 out of 114 days.

While individual pollution episodes with extremely high PM mass loadings such as dust storms can profoundly influence the cluster mean, they barely affect the number of exceedance days as such events are rare. "

References

Lawrence, M. G. and Lelieveld, J.: Atmospheric pollutant outflow from southern Asia: a review, Atmos. Chem. Phys., 10, 11017-11096, doi: 10.5194/acp-10-11017-2010, 2010.