Dear Editor,

Based on the suggestions of both referees #1 and #2, we have performed a lot of technical and grammatical corrections to refine the content of our manuscript. We have gone through quite thoroughly the manuscript and made many corrections/amendments. Now, the revised manuscript should be more lucid and reader-friendly.

Additionally, we have modified our figures and tables according to the suggestions of referee #1. A new table has been inserted to improve the clarity of our discussion. Meanwhile, several new references have been added into the manuscript as suggested by both referees. The title is modified as suggested by referee#2.

The content in the original manuscripts which was deemed unnecessary by referee #2 (and some by referee #1) was trimmed and appropriately improved. You may find the new version in the attachment.

For the sake of your convenience, we have highlighted the modified sections in different colors:

Yellow → Grammatical problems. Sentences are replaced by more suitable words suggested by reviewers.

Turquoise (light blue) → New discussion and amendments for clarification.

Grey → Additional citations and references.

Green → New addition information and restructuring of sentences.

To referee #1,

First and foremost, we would like to thank the first reviewer for his effort to technically edit our manuscript. This really helps to improve the scientific quality of it. It must have cost the reviewer quite a bit of time and effort to comment and review our work. We heartily thanks the reviewer for this.

In the following, texts in black font are the original comments by the reviewer. Our response is in blue font.

2nd review of the manuscript by Tan et al. submitted to ACPD.

The authors have been able to clarify the manuscript and improve the section regarding lidar validation. However, there are still number of points that have to be improved before this manuscript can be published.

First of all, the link between the aerosol typing and the AOD prediction model is still missing. In section 3.5 the authors speculate on the effect of aerosol types on the performance of the prediction model but in my opinion the discussion is not convincing (see specific comments for more details). Especially, when the authors claim that the best model version is the “overall” version which is not season/aerosol type dependent.
Response: In the fact, we do not claim that "overall" version is the best version. Nevertheless it does give a satisfactory result in term of R^2 and RMSE. This "overall" model can be interpreted as an effective and representative model which can predict AOD in every period. Meanwhile, the "overall" model has taken into account the contribution of every dominant aerosol type in the period of study. We found that when the dominant aerosol type is BMA, then the AOD distribution range is wider. Therefore, a wider range of AOD distribution will lead to better performance of our AOD prediction model. All the issues relating to the AOD prediction model and aerosol typing, as raised by reviewer #1 in the “*specific comments*”, are highlighted in grey in this document. The responses to these **greyly-highlighted** can be found there.

Secondly, there are still too many confusing parts which need to be clarified. If the reviewers are not able to follow the story, then other scientist that read the paper quickly won't have a chance. It's hard to follow which data set is used where and how the different numbers relate. For example, the authors mention in the abstract that the R 2 value is 0.68 but in the conclusions they mention that R 2 values are < 0.72.
Response: We have gone through quite thoroughly the manuscript again ourselves, and have made many corrections/amendments. Many sections in the manuscript which are considered as confusing or insufficiently explained have been rectified or further elaborated. For example, we have clarified on the differences in the values of R^2 of 0.68 and 0.72 as mentioned above. If a reader follows the storyline in the revised Section 3.5 and Section 3.6, he shall be able to see the distinctions between the values of R^2 of 0.68 and 0.72. Basically the R^2 value of 0.68 is meant for the validation accuracy of the model,

while that of 0.72 is defined for the calibration accuracy of the model. We believe that after our effort, the revised version of the manuscript should now be more lucid, reader-friendly. All the issues of R^2 raised by reviewer #1 in the "*specific comments*" were highlighted in green in this document. The responses to these **greenly-highlighted** can be found there.

Thirdly, the lidar part has to be explained better. Now, a lot of information is missing (e.g. how the lidar profiles were calculated) and there are strange gaps in the data presented in figure 10. Furthermore, the figures should be presented in the same time format.
Response: As a matter of fact, the details of the LIDAR signal processing are described in our previous papers which were cited in P5, l17. To explain better the details involved, we have added the necessary information of how the LIDAR signal was processed in the latest version. The figures with 'strange gaps' were re-plotted in the revised manuscript, where we have added in explanation to the gaps in the captions of the related figures. We have also standardized the time format of those related figures. Further details of modification to these re-plotted figures can be found in our response to the last question, "*Figure 9: This figures could be left out because …*", in the "Specific comments" posed by reviewer #1. Note that since some selected figures from the original manuscript were removed as suggested by the reviewer, the labeling of figures in the revised manuscript has changed in comparison to the original one.

Finally, the motivation of the study is a bit fuzzy. In the abstract the authors mention atmospheric correction (for what do you need this in cloudy cases?) and in the introduction they mention air quality (which is redundant because you have the API data). There is also more meaningful motivation discussed in the introduction and these should be mentioned in the abstract and conclusions.
Response: As suggested, we have improved the Abstract in order to remove the fuzziness of our motivation, where we have stated explicitly the motivation of the present work. As a related correction, the word "atmospheric correction" as appeared in the Abstract, as well as in the Conclusion section, has been replaced by **"**for climatological studies and monitoring aerosol variation**"**.

As a response to the reviewer's comment "*which is redundant because you have the API data*", the word "air quality" appearing in the Introduction section has been replace by "aerosol".

I also think that there are too many figures and some of them do not bring much to the discussion, thus, they could be left out.
Response: We have removed the figures recommended to leave out.

Specific comments:
p1, l19: usually references are not given in the abstract.
Response: Changed.

p1, l26: the R 2 value is different than in the conclusions or the tables. What value this is?
Response: We have overhauled our manuscript that involves R^2 are carefully examined and restated if necessary, so that there is no more confusion or contradictory values of R^2 when they appear in the body text or in the tables.

p1, l30: "atmospheric correction": for what? If you are providing AOD values for cloudy cases, they are not much use for satellite products such as surface albedo because the measurements cannot be made in cloudy cases.
Response: The word "atmospheric correction" has been replaced by **"**for climatological studies and monitoring aerosol variation**"**.

p1, l33: "human anthropogenic": human is redundant
Response: As suggested, the word "human" is now removed in the revised manuscript.

p1, l35: "right circumstances": which are what?
Response: The "right circumstances" are referring to, for example, (*i*) cloud free condition; (*ii*) the data is not affected by high surface reflectance due to snow- and ice-cover. However, we decided to rephrase the sentence to avoid this ambiguity.

p1, l38: "small scale studies": give references
Response: Citations and references are added.

p2, l4: "sufficient measurements": measurements of what?
Response: We have rephrased the sentence and now it read **"**To better monitor and understand aerosol variation, sufficient measurements and a practical observation of aerosols are necessary…**"**.

p2, l15: combine → be mixed
Response: Changed.

p2, l17-18: You are using air quality index in your prediction. Isn't that a better metric for for air quality than AOD?
Response: Actually, air pollution index is a healthy index. It is announced by government agency to indicate the pollution level to the public. On the other hand, AOD is a scientific parameter commonly used in the research community, and has a wide range of scientific application including aerosol study. Therefore, AOD is better parameter as compared to the air pollution index.

To avoid ambiguity, we change the word "air quality" to "aerosol".

P2, l41: Mention here that you are using direct sun data.
Response: This statement has been added to the revised manuscript as suggested.

P3, l16: Use the same format for coordinates throughout the paper. (see p2, l41)
Response: We have standardized the format for coordinates throughout the paper as suggested.

p3, l34: give reference for Hysplit.
Response: References have been added as suggested.

P3, l40: AERONET → the AERONET web page
Response: Modified as suggested.

p4, l5: explain the calculation of API in more detail because this is an important parameter for the model and its link to PM is not straightforward. If I have understood correctly, the reported API value is the maximum of the individual indicies for different pollutants. This means that the conversion of API to PM10 could be overestimated at some occasions.
Response: You are right. The reported API value is the maximum of the individual indicies for different pollutants. However in our case we only consider API values which are mainly due to PM10. As a response to this issue, I have added in a few sentences to the text which reads

"Malaysia is mainly polluted by PM_{10} (DOE, 2010). Therefore only API that are predominantly due to PM_{10} is used in this study. API is computed using the technique developed by US-EPA. The Malaysian Department of Environment provides a standardized procedure on how to calculate API values (DOE, 1997). The conversion between API and PM_{10} has been shown in the guideline provided in DOE (1997)."

As for your reference, the conversion as provided by DOE (1997) is listed in the table below:

Concentration of PM_{10} (24 hours average)	API
Conc. < 50 μ g/m^3	API = conc.
50 < Conc. < 350	API = 50 {[conc. – 50] x 0.5}
350 < Conc. < 420	API = 200 {[conc. – 350] x 1.4286}
420 < Conc. < 500	API = 300 {[conc. – 420] x 1.25}
Conc. > 500 μ g/m^3	API = 400 + [conc. - 500]

P4, l22: calculation of %MRE: you call this parameter mean relative error but isn't this actually relative error of the mean? They sound the same but they are slightly different. Based on the presented equation the difference is calculated from the averaged AOD values. In my view, you should first calculate the difference between each individual comparison and then calculate the average from them to be more exact. I would suggest you calculate the %MRE like this

$\%MRE = \sum |(AOD_{p,i} - AOD_{m,i}) / AOD_{m,i} | / N * 100$,

where the subscript p refers to predicted, m to measured and i to individual measurements. N is the total number of measurements. I would also encourage to do the comparison with absolute values because the comparisons can result in positive and negative values which could improve the statistics by canceling out big differences. For example, if you have two measurements and the first measurement is overestimated by 0.6 and the second underestimated by the same amount that would give a %MRE of 0 with the current method and 0.6 with the method I'm proposing. If you think that absolute values are not suitable here, at least, calculate the average from the differences and not the difference from the averages.

Response: We have taken your advice to calculate the %MRE (which is also known as mean absolute percentage error, MAPE). However, there are some problems with this method. For example, it does not weigh between different magnitudes. In addition, combinations with very small or zero values (a series of small denominators) can cause large skew in the results. A better method to use is the generalization of MAPE, i.e., weighted mean absolute percentage error (wMAPE). The formula for wMAPE is given as

$$wMAPE = \left(\sum \left| \left((AOD_{p,i} - AOD_{m,i}) / AOD_{m,i} \right) \right| * AOD_{m,i} / \sum AOD_{m,i}\right) * 100$$

For details of these methods, see (Muhammad, 2010; Okkan, 2012).

Adopting your suggestion, we have included wMAPE in our revised manuscript as a way to quantify the accuracy of the prediction by our model.

P5, l12: “the measured counterpart”: measurement of what done with what? AERONET AOD?

Response: The ‘measured counterpart’ refers to AERONET AOD. This statement is made because we were referring to Fig. 8 in the original manuscript. However, we have removed Fig. 8 in the revised manuscript as suggested by reviewer #1. As a result this statement becomes irrelevant, hence is removed in the revised manuscript.

P5, l13: give reference for the lidar system.

Response: A reference is added as suggested. This reference contains the characteristic and other technical details of the specific lidar system used by our group in USM.

P5, l15-16: This sounds suspicious. Why didn't you just use all the available lidar data? That would give more reliable results.
Response: For the analysis as mentioned in P5, l15-16 (original manuscript), not all available LIDAR data can be used in this case. We have to use only LIDAR data that is coincide in time with the AERONET and predicted AOD data. The examination procedure is further elaborated in p.11 l6-8 in the original manuscript. However, as was advised by reviewer #2, we have removed and slightly rephrase the related sentences.

P5, l21: "values so obtained" → values obtained
Response: Modified as suggested.

p5, l27: is → are
Response Modified as suggested.

p5, l35: clarify the sentence. Why are Angstrom exponent and precipitable water related in this sentence?
Response: The sentence did not mean to relate Angstrom exponent to the precipitable water. We have broken the sentences into two to avoid such misunderstanding.

P6, l9-15: use the same nomenclature in the text and in the figure: October-November → post-monsoon
Response: We have modified our manuscript so that it now adopts a uniformised nomenclature when referring monsoonal seasons.

p6, l40: "frequency data" → data
Response: Modified as suggested.

p7, l13-15: What does turbidity have to do with this study? This sentence could be removed.
Response: This sentence is removed.

P7, l17: This information was already mentioned on rows 9 and 10. Remove the sentence.
Response: This sentence is removed.

P7, l30: This statement underlines my question: why are you doing aerosol typing in this way if it is so uncertain?
Response: At p7, l30, our intention was just to show that threshold criteria suggested by Pace et al. (2006) and Kaskaoutis et al. (2007) and Smirnov et al. (2002b, 2003) were not

suitable to be used in this study as their criteria resulted in higher percentage of indistinguishable aerosol type (MIXA). However, in the following discussion (in the text), we showed that the threshold criteria suggested by Toledano et al. (2007) is more suitable to be used in our study area, because the percentage of MIXA in the criteria by Toledano is much lesser that by others'.

P8, l16: "reach in" → reach
Response: Modified as suggested.

p9, l8-9: This sentence needs rewording.
Response: The sentence was modified according to the suggestion by reviewer #2.

P9, l14: If you are talking about the figure 1, mention it in the text.
Response: Done as suggested.

P9, l14-15: This sentence needs rewording.
Response: The sentence was modified according to the suggestion by reviewer#2.

P9, l16: Same model for all seasons or different models for different seasons.
Response: For the original manuscript, we did a calibration model for each monsoon season (northeast, pre-, southwest, and post-monsoon). Meanwhile, we have also done an overall calibration model (which combined all monsoonal seasons). The results show that the overall calibration model gives a reasonable R^2, RMSE and wMAPE. Therefore, we choose the overall calibration model as our standard reference model for AOD prediction. A relevant discussion on this issue has also been addressed in our response to the question "*P9, l30: How well does the overall version perform at different seasons?.......*".

P9, l24: "<" → values were smaller than
Response: Modified as suggested.

P9, l26: "vice versa", do you mean that predicted AOD increased atmospheric AOD? Reword.
Response: No. What we actually meant was accuracy of the AOD prediction can be enhanced if there are more aerosols in the air. If the amount of aerosols in the air is very little, the accuracy of the AOD prediction will drop. We have reworded the sentence accordingly.

P9, l30: How well does the overall version perform at different seasons? Is it better or worse than the seasonal models? You would assume that the seasonal models are better there if the seasons/aerosol types have some effect to the results.

Response: Yes, the seasonal models are indeed slightly better than the overall models because "the seasons/aerosol types have some effects to the results". Below we show **three tables for your reference.**

Table A (for calibration**)**

For each seasonal monsoon period, we used the measured AOD in data subset 1 to calibrate the coefficient $\{a_i\}$ using Eq. (2). The calibrated $\{a_i\}$ so obtained are then used in Eq. (2) to predict values of AOD, which are directly compared to the measured AOD values from data subset 1 for that particular seasonal monsoon period. We shall call this sets "seasonally-calibrated $\{a_i\}$". We then repeat the procedure described above but using all data subset 1 that includes all seasonal monsoons (the "overall"). We shall refer this set as "overall-calibrated $\{a_i\}$". This result in a total of five sets of $\{a_i\}$: one for each seasonal monsoon and the overall period. The values of R^2, RMSE and wMAPE in Table A are that obtained by comparing the predicted AOD and the measured AOD from data subset 1for each seasonal monsoon and the overall data.

Table B

The seasonally-calibrated $\{a_i\}$ are used to predict AOD values for each corresponding monsoonal period. These predicted values are directly compared to the measured AOD in data subset 2 for that particular monsoon period. The values of R^2, RMSE and wMAPE in Table B are that obtained by comparing the predicted AOD and the measured AOD from data subset 2 for each seasonal monsoon using the corresponding set of seasonally-calibrated $\{a_i\}$.

Table C

The overall-calibrated $\{a_i\}$ are used to predict AOD values. These predicted values are directly compared to the measured AOD in data subset 2 for each monsoon period as well as the overall data in data subset 2. The values of R^2, RMSE and wMAPE in Table C are that obtained by comparing the predicted AOD (using overall-calibrated $\{a_i\}$) and the measured AOD from data subset 2 for each seasonal monsoon and the "overall data".

Referring to Table B and C, despite the seasonal models give results that are slightly better than the overall model (by comparing the values of RMSE and wMAPE of the seasonal validation data against the overall validation data), the overall model can still predict a reasonable AOD. In addition, we can perform AOD prediction throughout a year by adopting only one overall model (instead of four seasonal models).

Table A

	Model calibration (data subset 1)			
Monsoon	**R^2**	**RMSE**	**wMAPE (%)**	**Number of data**
Northeast monsoon	0.41	0.11	0.40	129
Pre-Monsoon	0.64	0.11	0.37	65
Southwest monsoon	0.77	0.17	0.08	117

Post-monsoon	0.42	0.06	0.21	83
Overall	0.72	0.13	0.04	394

Table B

	Model validation (using seasonally-calibrated $\{a_i\}$)			
Monsoon	**R^2**	**RMSE**	**wMAPE (%)**	**Number of data**
Northeast monsoon	0.32	0.11	0.15	128
Pre-Monsoon	0.60	0.11	0.18	67
Southwest monsoon	0.75	0.17	0.10	118
Post-monsoon	0.29	0.07	0.24	82

Table C

	Model validation (using overall-calibrated $\{a_i\}$)			
Monsoon	**R^2**	**RMSE**	**wMAPE (%)**	**Number of data**
Northeast monsoon	0.30	0.12	0.17	128
Pre-Monsoon	0.32	0.15	0.32	67
Southwest monsoon	0.74	0.17	0.09	118
Post-monsoon	0.17	0.08	0.37	82
Overall	0.68	0.14	0.05	395

p9, 31: Why the R 2 values are different for northeast monsoon and pre-monsoon while the mean AODs, RMSE's and %MRE's are almost identical? I would argue that the better correlation is caused by a wider distribution of AOD values for the pre-monsoon and southwest monsoon and not the flow patterns.

Response: (We assume you are referring to p9, l22-30.) Yes indeed. The better correlation is caused by a wider distribution of AOD values. A relevant discussion on this issue has also been addressed in our response to the question "*P10, l1-11: I'm not convinced by this discussion. I think more important than the aerosol…*"

P9, l41: "broad region": isn't this range same for all seasons?

Response: If scrutinizing the seasonal curves pair by pair, the post-monsoon and northeast monsoon pair (purple and blue curves) appears to be broader and flatter whereas the other two seasons (red and green curves) are sharper and narrower. To be more quantitative, the slopes for the purple and blue pair begin to pick up at a relatively fast pace (compared to the red-green pair) at around Angstrom exponent of 1.1, dropping at around Angstrom exponent of 1.7 from peak values, maintaining a relatively flat profile between these two limits. Whereas the slopes for the red and green pair begin to pick up at a relatively gentle pace (compared to the purple-blue pair) at around Angstrom exponent of 1.1 (for red curve) and 1.3 (for green curve), dropping at around Angstrom exponent of 1.6 (for red curve) and 1.7 (for green curve) from peak values. The profile for the red-green pair is relatively narrower and sharper between their pick-up and dropping limits.

To clarify the ‘broad region’ issue raised, we have inserted our response above into the revised manuscript. The paragraph is inserted into the text between lines 31 and 42, page 9 in the original manuscript.

P10, l1-11: I'm not convinced by this discussion. I think more important than the aerosol types is the distribution of AOD values. The wider spread of AOD you have in the training of your model, the better performance it will have. This is in line with the fact that you got the best performance with the overall model which does not depend on aerosol types. How does that fit in with this discussion about aerosol types?

Response: We are talking about “dominant aerosol types” during a particular season. At the first place, we totally agree that distribution of AOD values is the key to have better performance (i.e., the wider the AOD distribution range, the better is the AOD prediction by our model). However, based on the observation of the AOD prediction, we found that the AOD distribution range is influenced by the dominant aerosol type in each monsoon season. To elaborate this issue more clearly, we have restructured and rephrased the content of p.9,l31 – l42 and p.10,l1 – l11 in the original manuscript by the following lines:

“High correlation was observed between the measured and predicted AOD for the pre-monsoon and southwest monsoon, in which similar air flow patterns occurred (Fig. 4b and c). On the other hand, the prediction accuracy of the AOD model in the post-monsoon and northeast monsoon seasons was moderate. The air flow patterns in Fig. 4a and d, which are associated with northeast and post-monsoons respectively, also show similarity. This observation is consistent with Figure 1b which displays the relative frequencies of occurrence of $Angstrom_{440–870}$. If scrutinizing the seasonal curves pair by pair, the post-monsoon and northeast monsoon pair (purple and blue curves) appears to be broader and flatter whereas the other two seasons (red and green curves) are sharper and narrower. To be more quantitative, the slopes for the purple and blue pair begin to pick up at a relatively fast pace (compared to the red-green pair) at around Angstrom exponent of 1.1, dropping at around Angstrom exponent of 1.7 from peak values, maintaining a relatively flat profile between these two limits. Whereas the slopes for the red and green pair begin to pick up at a relatively gentle pace (compared to the purple-blue pair) at around Angstrom exponent of 1.1 (for red curve) and 1.3 (for green curve), dropping at around Angstrom exponent of 1.6 (for red curve) and 1.7 (for green curve) from peak values. The profile for the red-green pair is relatively narrower and sharper between their pick-up and dropping limits. As a result, a clear correlation between aerosol optical properties in Fig. 1b and seasonal wind flow patterns in Fig. 4 is observed.

The broader and flatter curves in the post-monsoon and northeast monsoon indicate that coarser aerosols are more frequently loaded in the atmosphere. This observation is proved in Fig. 3, in which both monsoons are showing higher occurrence of MA and lesser BMA in Penang. In fact, we realized that AOD predominantly falls on smaller values if MA is the dominant aerosol type, because clean atmosphere is dominated by MA. When the atmosphere is dominated by UIA then the AOD values are larger than MA. Normally, when BMA is the dominant aerosol type, AOD values are large. In other words, if BMA is absent or small, AOD will have narrower range of distribution. As a result, only a moderate accuracy in the AOD prediction is obtained for the post-monsoon and northeast monsoon (refer to Table 3).

By comparing the types of dominant aerosol observed during each monsoon, we observe that the results, as obtained in Table 3, correlate well with the information from Fig. 3. Table 3 shows higher coefficients of determination of the proposed AOD prediction model, which can be associated with higher amounts of BMA during the pre-monsoon and southwest monsoon periods. Such observation implies that the aerosol types are possibly indirectly correlated with the AOD prediction model. This result was also noticed by Chen et al. (2013). However, the relationship between the predicted AOD and aerosol type as observed in our model is qualitative and preliminary. Further study is needed. In addition, as mentioned in Lee at al., (2012) and Gupta et al., (2013), the relationship between AOD and particulate matter at the surface depends also on extent of atmospheric mixing, relative humidity, chemical composition, aerosol size distribution, etc"

The overall model performs well because the training data set has taken in account the contribution of all aerosol types, leading to wider distribution of AOD values.

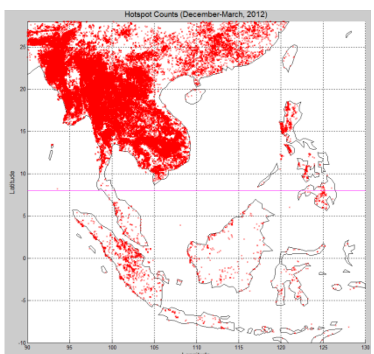
In addition, the wind flow patterns tell us the possible aerosol origins and the paths of aerosols being transported. **For your reference**, below we show the figures from satellite MODIS for fire detections from Terra and Aqua a(*i*), (*ii*) – d(*i*),(*ii*). These figures show seasonal fire plots in SEA for i) 2012 and ii) 2013. These figures show that fire activities occurred throughout the year, reaching the peak intensity during southwest monsoon period in Sumatra, Indonesia (location: southern Southeast Asia). BMA is significantly observed during this monsoon season, see Fig. 3 in the manuscript. As a result, we can infer that the transboundary aerosol (predominantly BMA) is originated from Sumatra. However, the occurrence of BMA is much reduced in the post-monsoon because of the reduction in open burning activities in Sumatra. Additionally, the transboundary aerosol from Sumatra is made less possible due to the wind flow pattern as shown in Fig. 4d in the manuscript.

On the other hand, northern SEA (Indochina) encompasses active biomass burning activities during northeast monsoon period, usually reaching maximum intensity in March or April (Lin et al., 2013). From the conceptual model, they found that biomass burning aerosols originating from Indochina are uplifted and transported in high- and low-level pathways by westerly and northeast monsoon flows, which lead the aerosols toward the southwest direction. We thus expect that these aerosols will continue to move toward our study site as wind circulation tends to blow toward the southwest direction (refer to the monthly mean streamline charts 1979–2010 of Lin et al. (2013)). These transported aerosols were finally mixed with the local aerosols.

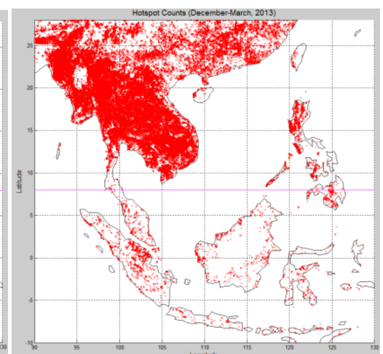
In Fig. 4c, 4b for the southwest monsoon and pre-monsoon, the wind flow patterns clearly show that it is highly possible that BMA could have originated locally and from Sumatra. Therefore, during both of these seasons the model has a better performance. On the other hand, we clearly observe that the northeast monsoon and post monsoon have similar flow patterns. Looking specifically at the northeast monsoon (Fig. 4a), we realize that BMA could have been originated from Indochina, based on the discussion of the results on Fig. 1b. Additionally, the studies by Lin et al., 2013, indirectly suggest that the BMA observed at Penang could also be originated from Indochina. Aerosols that have traversed through a geographically long distance will lose in quantity due to deposition and rainout. Therefore, the AOD distribution range is smaller during northeast monsoon. Following the chain of arguments presented above, a moderate performance of the AOD prediction model is obtained.

During post-monsoon, the wind flow pattern clearly shows that the wind is coming from the northeast direction. In this period, fire burning activities in Indochina is in-active [see Fig. d(i),d(ii) shown below], hence we do not expect BMA to come from Indochina. At the same time we also do not expect BMA from Sumatra (due to the wind flow pattern in Fig. 4(d) during this period). This observation is consistent with what has been observed in our dataset, in Fig. 3, for the post-monsoon period. As a result of reduced BMA in Penang in the post-monsoon period (due to the explanations presented above), AOD values are distributed in a smaller range, resulting in a weaker performance of the model.

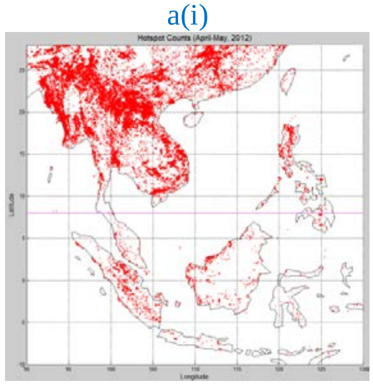
Apparently, BMA emission transportation from the fire fields are linked to the wind flow patterns. Following the discussion above, it can be inferred that the distribution of AOD values is indirectly related to the wind circulation (which could possibly transport aerosol across boundary) and the dominant aerosol types in the monsoon season.



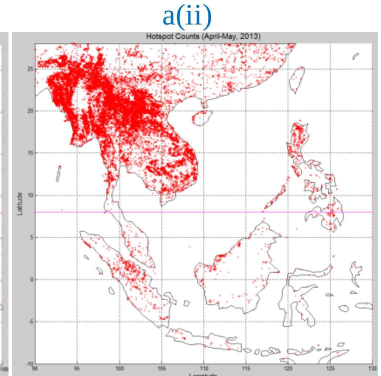
a(i)



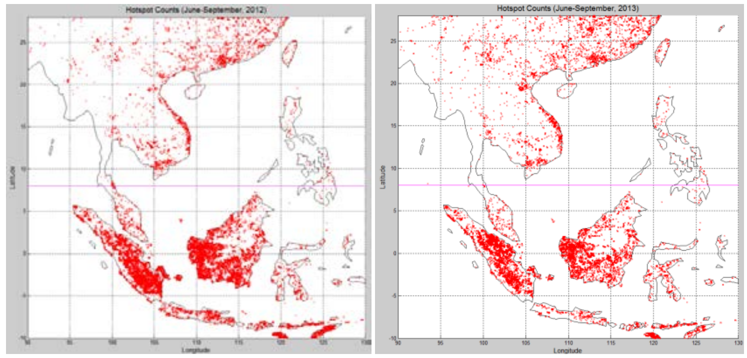
a(ii)



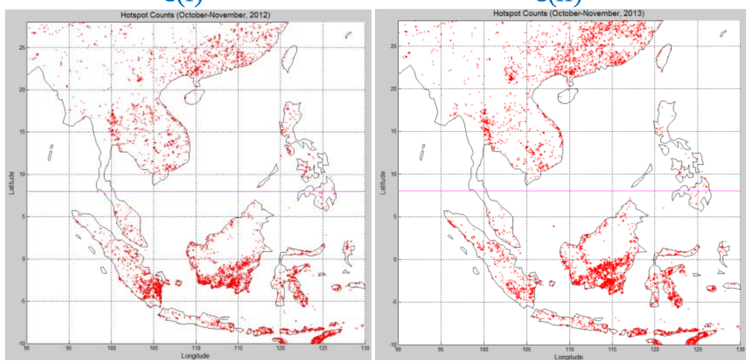
b(i)



b(ii)



c(i) c(ii)



d(i) d(ii)

Figure: Seasonal fire or hotspot counts map in SEA for i) 2012 and ii) 2013. (a) northeast monsoon, (b) pre-monsoon, (c) southwest monsoon, (d) post-monsoon. The northern (8˚N to 28˚N) and southern (10˚S to 8˚N) of SEA are separated by a pink line.

P10, l9: "air quality": do you mean PM10?

Response: Yes, indeed. To rectify the ambiguity, I have changed the word "air quality" to "particulate matter".

P10, l10: "environmental factors": chemical composition and aerosol size distribution are not environmental factors. They are aerosol properties that are embedded in the AOD values.

Response: The misleading word "environmental factors" were removed from the manuscript.

P10, l13-35: This is the most confusing part of the paper. It is not clear where all these values are coming from (what are the different versions of the predictions) and how do they relate to the previous values. For example, what is this R2 value of 0.68? How does it differ from 0.72 that was given in the previous section? Are these values compared to the same AERONET data as in the previous section or not? State clearly that which values are coming from which version of the model and what kind of AERONET data is used. I would suggest that you put all these values in a table so it will be easier to compare them.

Response: We have overhauled our manuscript that involves R^2 are carefully examined and restated if necessary, so that there is no more confusion or contradictory values of R^2 when they appear in the body text or in the tables. As suggested, we have put all these values in Table 3 "so it will be easier to compare them". In addition, the use of %MRE is replaced by wMAPE as a response to the question "*P4, l22: calculation of %MRE: you call this parameter mean relative error but isn't this actually relative error of the mean?...*".

P10, l26: Here you say that R2 was not improved but on line 23 you mentioned that R2 was enhanced. This is a really confusing part.
Response: We have tested the approach proposed by Lee et al. (2012) to remove potential outliers in overall data subset 1. We then obtained a new set of coefficients $\{a_i\}$ which have better R^2 and RMSE. However, when this set of coefficients are applied to predict AOD and compared to measured AOD for validation in data subset 2, the results remained the same as before filtering out those potential outliers.

As a measure to improve our presentation, maintain consistency and remove any potential ambiguity on R^2, we have rewritten a major portion of the manuscript in Section 3.6. Additionally, we have put all relevant values in a new table (we named it Table 4 in the revised manuscript, "so it will be easier to compare them".

P10, l42-43: The sentence about underprediction is not clear.
Response: We have rephrased the sentence. Now it reads
"By comparing the measured and predicted data in Fig. 5, it is found that only a few small time segments are significantly underpredicted, possibly due to the presence of aerosol residual layer beyond PBL."

P11, l4-5: Are you saying that overprediction in your estimates is caused by local fires?
Response: What I meant was: occurrence of significant overprediction might be caused by local fires near the measurement station.

P11, l11-14: This information is not needed and figure 6b could be left out. Instead, you should tell more about the lidar system. For example, how was overlap dealt with and what is the lowest usable altitude? Did you use single profiles in the analysis or did you use time-averages of the lidar data?
Response: Figure 6b is removed as suggested. The overlap height (R_o) is about 200 meters for our LIDAR system. There are so many details in the preprocessing of the LIDAR signal. We have actually cited our previous papers for the preprocessing procedures (as shown in the original version at P5, l17). We shall briefly illustrate the protocol here.

“First, background solar radiation in the LIDAR signal has to be deducted. Then the analog and photon signals in the LIDAR signal are combined to enhance the near and far field signals. The range corrected signal (RCS) is obtained by multiplying the glued signal with a range square. To increase the signal-to-noise ratio, every 10 data files (each file contains data taken for one minute) are averaged over to give a single 10-minute averaged data. Then, the spatial resolution is decimated by averaging over 10 bins (each bin is separated at a distance of 7.5m) spatially to obtain a 75m resolution profile. The RCS is then normalized by calibrating it against theoretical molecular backscatter according to the USSA976 standard atmosphere.

Often, the raw data are contaminated by presence of clouds during data-taking. Such contamination has to be removed so that the data is clean for the purpose of abstracting the values of AOD. LIDAR scattering ratio (defined as LIDAR signal divided by molecular backscatter signal) (Wang and Sassen, 2001; Lo et al., 2006) is used as a means to remove the cloud contaminated data. The referred LIDAR signal here is RCS, whereas molecular backscatter signal is referred to the attenuated molecular backscattering.
”

We have added the above description into the Methodology section in the revised manuscript. Meanwhile, the overlap height, R_o = 200 m, is also mentioned in the revised manuscript when Eq. (6) was discussed.

P11, l16-19: I still don't understand this sentence. Do you mean that the predicted AOD was overestimated by 0.04 at 10:00 and underestimated by 0.04 at 11:00? If so, you have to reword the sentence. What are the measured AODs at these times?
Response: Yes, indeed this was what we meant. We have reworded the sentence.

Now it reads:
“Aerosols had accumulated near the ground at 10:00 a.m. and our model indicated that the predicted AOD was overestimated by 0.039. By contrast, most aerosols at 11.00 a.m. were at a higher level; therefore the model predicted value was underestimated by 0.044.”

We could have added in a table containing numerical values of the measured and predicted AOD at different times to illustrate the underprediction and overpredicton cases. However, such numbers could be just omitted as these excessive details are slightly redundant, and should not undermine the development of our conclusions in this section.

P11, l38-39: I don't understand this sentence.
Response: What we originally meant was: We calculated the average value of the predicted AOD using 4493 data points, and we got a value of 0.31. The average value of the measured AOD from AERONET in the relevant periods was calculated to be 0.29. We wish to compare these two values (0.31 vs. 0.29) as a numerical evidence to justify the accuracy of our model. However, after a second thought we believe that such comparison is not really a valid criterion for checking the accuracy of our model in this case. Hence we removed these two lines from the manuscript, without any lost of

generality. To maintain the flow and continuity of the context, we replace the two lines by the following sentences:

"It is observed that the variation in the predicted AOD matches with that of the measured AOD from AERONET. Hence, as an application of the AOD predicting model, information missed out by sun photometer (i.e., AERONET) could be reasonably well reproduced. These 'retrieved' AOD can be used in other aerosol studies. For example, the diurnal variability of AOD can be significant, depending on location and dominant aerosol type (Arola et al., 2013). They observed that the measurement-based estimates of aerosol direct radiative forcing (also known as aerosol direct radiative effect) at regional or individual sites are substantially influenced by the diurnal variability of AOD. In (Pandithurai et al., 2007), they found that the diurnal AOD variation depends on meteorological factors such as relative humidity, winds, temperature and convection activities. Our model provides a helpful means to investigate the uniqueness of diurnal variability of AOD in different seasons of a specific region."

P11, l41: "examine into" → examine
Response: Modified as suggested.

P11, l41: Why were these windows selected? Why didn't you use all the available lidar data? It would be better to use all data available because then the validation is more comprehensive and more reliable.
Response: First, we wish to clarify to the referee that P11, l41 is not meant for validation but an application and visual comparison of the trend between the measurements (from AERONET) and prediction (from our model). In fact, the majority of our AOD measurement in Level 2.0 that could be used (i.e., the ones that coincide in time with the predicted AOD for comparison purpose) only contains very few data points per day. Therefore, we purposely select only days which contain more data points to see how well our predictions are against measurements. In this section we do not apply the LIDAR data (this statement is clearly stated in P11, l43 in the original manuscript). Anyway, the issue raised here is rendered irrelevant now as we have removed the text in p11, l41 – p12, l6 from the original manuscript. The removal is a response to the issues raised by the reviewer in *"p12, l3: You just show that the model gives values for each hour. This is quit … "* (see below).

P11, l43: "CIMEL sun photometer" → AERONET
Response: This issue is irrelevant now as we have removed the text in p11, l41 – p12, l6 from the original manuscript as a response to the issues raised by the reviewer in *"p12, l3: You just show that the model gives values for each hour. This is quit … "* (see below).

p12, l3: You just show that the model gives values for each hour. This is quite evident so I don't think you need a figure to show this. It would be more interesting to see a day when you have a lot of AERONET measurements. Then you could see how well the predictions match with the measurements.

Response: As a matter of fact, our AERONET data are very limited in number that coincides in time with the predicted AOD for comparison purpose. In a typical 9-hour measurement, the number of final usable data for comparison purpose (use to produce Fig. 8 in the original manuscript) is maximum 9 (but most of the time it is much less than that). Fig. 8 is meant for visual comparison of the variation (i.e., the trend) between the measurements (from AERONET) and prediction (from our model), so that we can conclude how well these two are matched. Fig. 8c is the best we can show as a case with the most data points (prediction in blue dotted line are compared against the measurements' red dotted line). Meanwhile, Fig. 8a and 8b (which contain even less number of data points) are only for supplementing purpose to Fig. 8c. These figures show that the prediction model is applicable because the predictions (as shown in blue dotted line) are quite well matched with the measurements. However, since the reviewer opined that such hourly data in Figure 8 is 'evident', we have removed Fig.8 and the related discussion associated with it. Specifically, the text in the original manuscript from p.11, l41 to p.12, l6 are removed.

P12, l7-16: This part could be left out because it doesn't bring much to the analysis. It's hard to compare backscatter profiles with AOD time series. Figure 10 is much more informative.
Response: This part is left out as suggested.

P12, l17 – p13, l12: This part should go to the method section. Define equations before referring to them (e.g. Eq. 5)
Response: We have moved the related portion to the methodology section as suggested.

p13, l13: Do you have to remote the lidar profile if it has a cloud? Couldn't you just mask the cloud and calculate the AOD?
Response: (We assume you meant 'remove' instead of 'remote' in your statement). In our procedure, we simply remove the data from LIDAR profile if it has a cloud by using aerosol scattering ratio . As for the second question, cloud masking indeed could increase the number of data available for AOD calculation. However we have not performed such procedure in our analysis as it could possibly lead to increased uncertainty of the resultant data.

See for example in (Wu et al.;Wang and Sassen, 2001;Lo et al., 2006)

In principle, we could either directly remove the cloud-affected data or do a cloud masking procedure. We choose the former way as it is easier to implement. Furthermore, despite not adopting the cloud masking procedure, the result obtained does not in any significant way affect our conclusion.

P13, l21: Explain the hypothesis also here.

Response: We have added in two more lines of elaboration related to this issue (at the last few lines in Section 3.7) so that the flow of explanation is now in better shape.

P14, l19-20: What do you mean by optical properties here?
Response: Each monsoon season is dominated by a different aerosol type. Therefore the optical properties of aerosol measured (e.g., Angstrom exponent, scattering and absorption of radiation, single-scattering albedo, refrective index, phase-function, symmetry factor of aerosols, etc) are different from one monsoon season to another. These are the original intention we meant when we said 'The variation in aerosol types for different monsoon seasons clearly yields distinct optical properties'. It is a general statement which is not specific to the present work of this paper.

Table 1: Month → Number of months
Response: Modified as suggested.

Table 3: Add the number of measurements and mean AOD for each season to the table.
Response: The number of measurement for each season is added to Table 3. However, the mean AOD will not be included since the mean AOD do not carry any extra information for this study.

Figure 4: I still think that this figure should be redone. The scales and regions are different which makes comparison between the plots difficult. If the scales do not provide any information why are they shown?
Response: Figure 4 has been modified. The scales, longitude and latitude of the figures are now uniformized. Meanwhile, as suggested by reviewer #2, their legends have been combined into one.

Figure 5 and 7: Still don't understand what is the difference between these two figures. Explain the difference more clearly in the text.
Response: Clarifications and explanations have been added into the captions in Fig. 5 and 7. Fig. 5 is meant for validation, while Fig. 7 is for application of the model. Upon reading these captions, one shall immediately understand the differences between these two figures.

Figure 6: Plot a) is confusing because you mention in the caption that there is a gap in the lidar data but it is not shown in the plot. Clarify the plot. As I already mentioned, b) is not necessary because you are showing backscatter profiles.
Response: Fig. 6a has been re-plotted. Fig. 6b is removed.

Figure 8: This figure could be left out because it shows an evident thing.
Response: The Figure 8 is left out as suggested.

Figure 9: This figure could be left out because Figure 10 is more informative. Maybe you could combine 9a with Figure 10. Why are there gaps in the time series shown in 9a? Now the plot is a bit misleading because the gaps are not shown.
Response: Fig. 9 is left out as suggested. We have combined 9a with Figure 10 where they are re-plotted. We have re-structured these figures to display the gaps. The revised manuscript refers them as Fig. 8a, b, and c. The captions have also been rephrased accordingly.

References

Arola, A., Eck, T. F., Huttunen, J., Lehtinen, K. E. J., Lindfors, A. V., Myhre, G., Smirnov, A., Tripathi, S. N., and Yu, H.: Influence of observed diurnal cycles of aerosol optical depth on aerosol direct radiative effect, Atmospheric Chemistry and Physics, 13, 7895-7901, 2013.

Chen, B. B., Sverdlik, L. G., Imashev, S. A., Solomon, P. A., Lantz, J., Schauer, J. J., Shafer, M. M., Artamonova, M. S., and Carmichael, G.: Empirical relationship between particulate matter and aerosol optical depth over Northern Tien-Shan, Central Asia, Air Quality, Atmosphere and Health, 6, 385-396, 2013.

Lo, C., Comstock, J. M., and Flynn, C.: An atmospheric radiation measurement value-added product to retrieve optically thin cloud visible optical depth using micropulse lidar, Rep. DOE/SC-ARM/TR, 77, 2006.

Muhammad, K. J.: The Effect of Optimization of Error Metrics, Master, Högskolan i Borås/Institutionen för Data- och affärsvetenskap (IDA), University of Borås, 1-43 pp., 2010.

Okkan, U.: Using wavelet transform to improve generalization capability of feed forward neural networks in monthly runoff prediction, Scientific Research and Essays, 7, 1690-1703, 10.5897/SRE12.110, 2012.

Pandithurai, G., Pinker, R. T., Devara, P. C. S., Takamura, T., and Dani, K. K.: Seasonal asymmetry in diurnal variation of aerosol optical characteristics over Pune, western India, Journal of Geophysical Research: Atmospheres, 112, 2007.

Wang, Z., and Sassen, K.: Cloud type and macrophysical property retrieval using multiple remote sensors, Journal of Applied Meteorology, 40, 1665-1682, 2001.

Wu, Y., Chaw, S., Gross, B., Zhao, Y., Moshary, F., and Ahmed, S.: Measurement of thin cloud optical properties using a combined mie-raman lidar.

To referee #2,

First of all, we would like to thank your effort to technically edit our manuscript. This really helps to improve the scientific quality of it.

Based on the suggestions of both referees #1 and #2, we have performed a lot of technical and grammatical corrections to refine the content of our manuscript. We have gone through quite thoroughly the manuscript and made many corrections/amendments. Now, the revised manuscript should be more lucid and reader-friendly. Additionally, we have modified our figure 4 according to the suggestion of referee #2. However, Figure 9 is left out as suggested by referee#1. There are several new references added into the manuscript, as were suggested by referee #2 and some by referee#1. The title is also modified as suggested by referee#2.

The content in the original manuscripts which was deemed unnecessary by referee #2 (and some by referee #1) was trimmed and appropriately improved. You may find the new version in the attachment.

For the sake of your convenience, we have highlighted the modified sections in different colors:

Yellow → Grammatical problems. Sentences are replaced by more suitable words suggested by reviewers.

Turquoise (light blue) → New discussion and amendments for clarification.

Grey → Additional citations and references.

Green → New addition information and restructuring of sentences.