Response to comments from reviewer 2

Manuscript: A satellite-based analysis of semi-direct effects of biomass burning aerosols on fog and low cloud dissipation in the Namib Desert # EGUSPHERE-2024-1627

General comments

This manuscript assesses the impact of seasonally emitted biomass burning aerosol (BBA) on the fog and low cloud (FLC) dissipation in the Namib Desert, using long-term geostationary satellite observations, reanalysis data, and a statistical model tasked to disentangle the role of BBA from that of the co-varying synoptic-scale meteorological pattern. They find that the FLC dissipation time is significantly later on high BBA days, which is mainly attributed to the longwave radiative effect of the co-transported moisture and subsequent changes in the regional circulation pattern and atmospheric heating profiles. Although the ridge regression model they trained is able to reproduce the statistical mean difference in the FLC dissipation time between high and low BBA days, it fails to provide definitive conclusions about BBA effects, due to the underfitting issue.

The manuscript is well written and enjoyable to read. I find this work appears of sound methodology and is of great interest to the community, with particular implications for the climate and the hydrological cycle of the Namib Desert.

I do have a few minor points/comments on the paper that I would like the authors to consider and address, in order to improve the manuscript's clarity and the soundness of the conclusions.

First of all, we would like to thank the reviewer for the positive comments and for providing very helpful and constructive feedback. We thank the two reviewers in the acknowledgments of the updated version of the manuscript as follows "We are also grateful to the two anonymous reviewers for their careful and constructive feedback, which has helped improve the manuscript."
The comments and suggestions of reviewer 2 are incorporated in italics and addressed by the authors in blue.

Comments

Attribution to BBA semi-direct effect

To me, the evidence that the authors demonstrated for a dominant role of the changing meteorological patterns (circulation, heating profile, moisture LW effect) in delaying FLC dissipation time is convincing and robust, but there are two subtleties to this that I think the authors could address/discuss in the paper.

The definition of 'semi-direct' effect of BBA, one thing I think this paper wasn't very clear about is the definition of the 'semi-direct' effect (appears in the title) or BBA effect, one could argue it's the local effect of BBA on FLC, all else equal, or, one could define it as the net, integrated impact of the presence of BBA on FLC (accounting for large-scale circulation adjustments). For example, Diamond et al. (2022) discusses both the "large-scale" and "local" semi-direct effects of BBA on low clouds in the SE Atlantic.

It is true that additional precision regarding semi-direct effects definition was necessary in the manuscript.
- Added in the introduction: "Semi-direct effects in this study refer to the "large-scale" semi-direct effects, as defined in Diamond et al. (2022), involving atmospheric thermodynamic, stability, and circulation adjustments resulting from the absorption or scattering of solar radiation by aerosols."

Changes in meteorological conditions and coastal circulation that the authors focus on discussing in the paper, to me, is part of this "large-scale" BBA/moisture semi-direct effect. However, there could be potential contributions simply from the spatiotemporally correlated synoptic patterns and regional BBA transport, which has nothing to do with the moisture LW effect or the BBA absorption, one such example is the mid-latitude intrusion events (more frequent in Sept.) that can constrain the smoke plume closer to the continent (e.g., Zhang & Zuidema, 2021, ACP). I wonder if the method used in this study can indicate/control such correlations?

By comparing the synoptic differences on typical (i.e., averaged) high and low BBA days, the method may not capture relatively uncommon events, such as the intrusion events you mentioned. These events can significantly impact BBA transport and distribution on specific days, but they will not be clearly represented in the averaged synoptic situations used in this study. This an additional limitation that should be discussed:
- Added at the end of the Section 3.2: "However, by comparing situations averaged over hundreds of days, out-of-the-ordinary events, such as mid-latitude intrusion events (Zhang and Zuidema, 2021), which can significantly impact BBA transport and distribution on specific days, are not well captured in this study."

Clarity on the methodology

It wasn't clear to me whether only June-October are used in the analyses or the whole year was used? (L82 says 15 years of data are used while L344 says only months June- October are used)

Only June–October is selected from the 15 years of available data. We modified the sentences for better clarity.
- Modified in Section 2: "As the study focuses on the interactions between FLCs and BBA, all analyses are conducted during the BBA season from June to October (De Graaf et al., 2014) over a 15-year period (2004–2015)"
- Modified in the conclusion: "During the BBA season (June–October) of the investigated 15-year period (2004–2018)"

L83 'some of the analyses' and L128 'For some analyses,' are not clear, please clarify and be specific

Modified in Section 2: "...**all analyses** are conducted during the BBA season from June to October..."
Modified in Section 2.3: "**If not explicitly indicated,** 8 UTC fields are selected..." the 8 UTC fields are used every time for the ERA5 data except for Fig. 6 a) & b) where every morning time step is used.

L103, please briefly summarize how does this work, such that one doesn't need to go to Pauli et al. (2022) to grasp the idea of this method (one can of course read it if more details are sought).

We added information about the method.
- Modified in Section 2.1: "This method, as described by Pauli et al. (2022), employs logistic regression to predict the probability of a data point belonging to one of two binary classes: FLC or no FLC. By defining the transition between classes when this probability exceeds 50%, it becomes possible to determine the time of the transition from one class to the other and, as a result, determine the formation and dissipation times of each individual FLC event."

What are the predictors used in the ridge regression? It says broadly "the predictors are the spatial fields of ERA5 meteorology and BCAOD from CAMS" I assume all the variables mentioned in Section 2.3 are included in the training?

The predictors are now explicitly listed: "The predictors are the spatial fields of MSLP, Z650, T2M, SST, Q650, Q975 and EIS (as defined in Section 2.3 from ERA5 and BCAOD from CAMS."

What exactly are the thresholds for BCAOD (25th and 75th percentile)? (a histogram as a supplementary figure would be nice)

For the classification into the low and high BBA conditions, before using a percentile approach we tried to define common AOD thresholds across the two regions. However as the distributions are quite different between the Angolan Namib (AN) and the Central Namib (CN), it was not possible to define reliable thresholds with approximately the same number of members in both groups and both regions.
- We added a figure in Appendix A (Figure 1 in this document) in Appendix A showing the distribution of CAMS BCAOD.
- Added in Section 2.2: "In the CN, the first quartile is characterized by a high density, followed by an almost exponential decrease into and throughout the third quartile. In contrast, the AN exhibits a more linear increase in density within the first quartile, followed by a gradual decline as BCAOD increases (see Appendix A). These differences are attributed to the distance from emission sources."

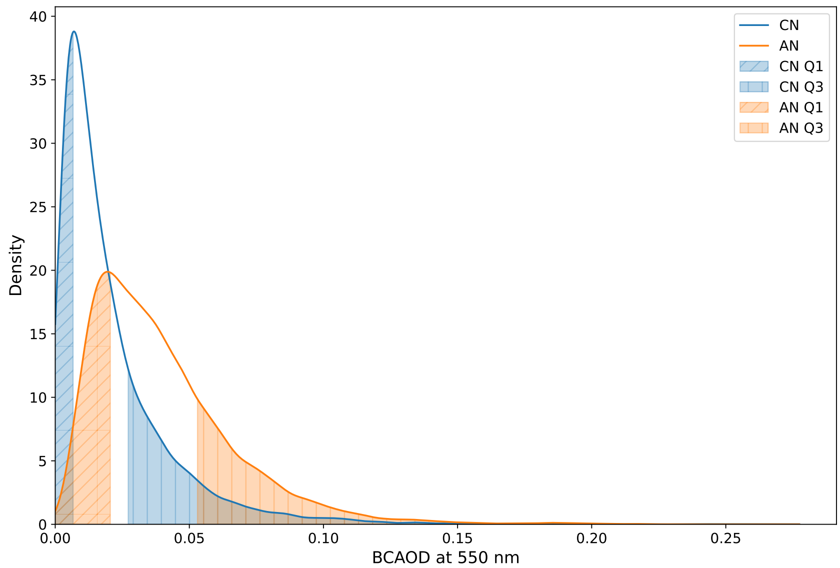

Figure 1: BBA season mean climatology (2004–2018) of the BCAOD probability density distribution in the AN and CN, along with their respective 1st and 3rd quartiles.

How good is this column-integrated BCAOD reanalysis, since there are aircraft measurements from the field campaigns in the region, I wonder if this product can be validated against observations?

Although aircraft measurements are proven to be very useful, they may not be the most appropriate for evaluating the columnar AOD used in this study.
Gueymard and Yang (2020) conducted an extensive validation of CAMS using AERONET observations. The authors of this study write in their conclusion: "From the global analysis presented here, it can be concluded that, for many applications, AOD data from reanalyses such as MERRA-2 or CAMS offer significant advantages over more customary satellite observation".

I wonder if the authors have considered using the above-cloud AOD (ACAOD, by Kerry Meyer at NASA) product to indicate BBA loading, given it's more observationally-based

Thank you for the suggestion, this is indeed an excellent product. However, we prefer to use reanalysis data instead of satellite products or field campaigns (mentioned in the previous comment) for several reasons: there are no spatiotemporal data breaks, cloud interference is significantly reduced, it offers hourly temporal resolution, and it is convenient to combine with ERA5 data as predictors in our statistical model, requiring only an interpolation of ERA5 onto the CAMS grid.

L120, assuming 15 years of Jun-Oct days are used, a 25th percentile will yield 562 days for each 'high BBA' and 'low BBA' group, is there additional screening involved? or missing data?

Yes there is additional screening involved: because of the CAMS resolution, there are very few pixels per region (CN and AN). We chose a rather aggressive approach by discarding every day where at least one pixel has missing data to have a consistent number of CAMS pixels per region. This leads to approximately 300 days per quartile as written in Section 2.2.
- Added in Section 2.2 : "They are referred to as 'high BBA days' and 'low BBA days,' **each containing around 300 days after discarding those where at least one pixel of the region is outside the defined thresholds.**"

L194, why only 200 days? instead of 300 days (L120)

From the 300 days we get from CAMS, we discard again around 100 days with missing data in the FLS dataset (for example during clear sky days) and we end up with around 200 days per group in Fig.3 as written in Section 3.1.
- Added in Section 3.1: "Then, the data is separated into two groups of high and low BBA loading with around 200 days in each group **after discarding days without FLC events.**
- Figure 3 modified to show the number of datapoints per bar.

In Section 2.4, CALIPSO is introduced, but I don't see how it is used. I couldn't find anywhere in Section 2 or 3 where CALIPSO data is used/discussed. L279 states the lack of skill of the statistical model could be due to the lack of vertical information of BBA (it confuses me as I thought you used CALIPSO to get the vertical dimension).

We used CALIPSO to obtain the latitudinal and longitudinal cross-sections in Figure 2. This data is utilized to characterize the cloud and aerosol layers, as well as their relative positions. Unfortunately, it is not feasible to implement this data as

a predictor in our statistical model due to several differences between the satellite data and the reanalysis data, including missing data, differences in temporal and spatial resolution or cloud interference.
- Added in Section 2.4: "**To characterize the cloud and aerosol layers, as well as their relative positions, data** from the active-sensor platform of the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) is **used and presented in Figure 2.**"

Statistical model

In general, I wonder what's the rationale to stick to this ridge regression linear model (given the low R2 values), over other non-linear models, such as CNN or random forest??

A CNN or random forest may seem like a good idea, but unfortunately, neither will work for this study. For CNNs, we simply do not have enough data (around 200 days per group) to achieve proper model performance. In the case of random forests, the problem lies in the high number of correlated predictors, which provide redundant information, reducing the effectiveness of the random feature selection and increasing the risk of overfitting. We use ridge regression, which works well with our relatively small amount of data and addresses the issue of correlated predictors through regularization. Additionally, it allows us to interpret the model's predictions using the coefficient map, providing a clear understanding of the model, something that is harder to achieve with more complex models like CNNs.

What's the ensemble spread in terms of member skill and the prediction of the mean delay in dissipation time?.

The ensemble spread in terms of member skill is 0.04 in the AN and 0.03 in the CN
The ensemble spread in terms of the prediction of the mean delay in dissipation time is 7 minutes in the AN and 9 in the CN.
- Modified in Section 3.3: "The ensemble average skill in the AN is $R^2 = 0.34$ with a standard deviation of 0.04, and in the CN, it is $R^2 = 0.30$ with a standard deviation of 0.03."

Could you show the scatter plot between truths and predictions for this model?

We added the plot to the Appendix C (Figure 2 in this document). The spread around the y = x line is quite large, with a tendency for over-predicting in both AN and CN. Nevertheless, the highest density of points can be found around the y = x line. These results are expected given the R^2 values of the models.
- Added in Section 3.3: "The plot of actual versus predicted dissipation times (see Appendix C) exhibits a relatively large spread around the line of perfect agreement, with a tendency for over-prediction in both AN and CN. Nevertheless, the highest density of points broadly follows the diagonal line."

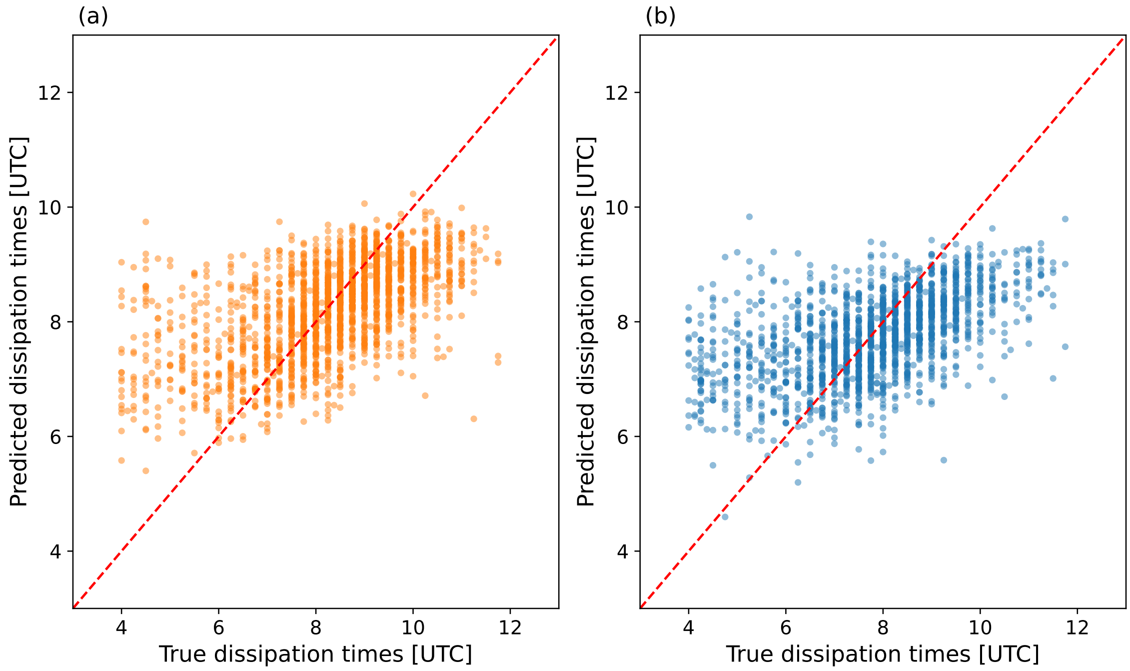

Figure 2: Ensemble mean predicted vs. true values for dissipation times [UTC] for the AN (a) and CN (b) models. The red dashed line represents the y = x line, indicating where predicted values equal true values.

As mentioned in L279, have the authors tried other predictors to try to improve the skill of the model?

Other meteorological fields were tested such as Boundary Layer Height, U and V components of winds at 650 hPa and at 10 meters or relative humidity. This didn't lead to a higher skill but increased the overfitting. The predictors are already highly correlated and therefore we argue that the model has indirect access to information not explicitly present as predictors (for example coastal circulation with MSLP). So we tried to select as few predictors as possible to reduce the risk overfitting but still have a complete picture of the meteorological situation. Nevertheless, some important information for the prediction of FLC dissipation is probably missing. One important aspect in the lifetime of FLCs is the temporal evolution of the meteorological fields, but here we only have a fixed picture at 08:00 UTC. Therefore, adding some information about the temporal variations of these fields before the dissipation could also increase the skill. Additionally, the predictand (dissipation times) is probably too complex to be accurately modeled by a linear model, but despite its limitations, this model still improves our understanding of the system.

Editorial

L103, what is 'logistic regression'? a typo?

Logistic regression is a statistical method used for binary classification tasks, where the goal is to predict the probability that a given input belongs to one of two categories.

Pauli et al. (2022) applied the logistic regression on the data binary data obtained by Andersen and Cermak (2018) to get formation and dissipation times of FLC events.
- Modified in Section 2.1: "This method, as described by Pauli et al. (2022), employs logistic regression to predict the probability of a data point belonging to one of two binary classes: FLC or no FLC. By defining the transition between classes when this probability exceeds 50%, it becomes possible to determine the time of the transition from one class to the other and, as a result, determine the formation and dissipation times of each individual FLC event."

L133, do we know which direction?

ERA5 simulations are steered towards a static baseline of climatological means. This baseline does not change dynamically with atmospheric conditions or real-time emissions, such as biomass burning events for example.
- Changed in Section 2.3: "Therefore, it is important to keep in mind that this prescribed forcing constrains the model's radiative environment by limiting the effects of aerosol variability."

Figure 2, cross and star labels on the map is reversed?

Yes thank you for pointing this error. This is now fixed.

Figure 3, caption, please define IQR at first use.

Caption modified in Figure 3 and 7 to: "... the whiskers extend to the highest (lowest) data points still within 1.5 times the interquartile range from the third (first) quartile.

L230-235, what's the reason for such a strong land-sea contrast (reversed in the sign) in terms of T2M difference between high and low BBA days?

We interpreted this as a heat low which is created by the long-wave absorption in the free troposphere caused by water vapor and BBA aerosols. Northerly advection of warm air masses is also likety to contribute to this heat low.
- Added in Section 3.2: "As such, the observed negative pressure anomalies are a clear sign of a heat low anomaly, **which is defined as an area of low atmospheric pressure caused by intense surface heating.** This phenomenon seems to be mainly driven by the greenhouse warming of the moist free-tropospheric air masses."

L349, as mentioned in my first comment, BBA effects can include these large-scale circulation adjustments, depending how one defines it.

Yes so not only there is the issue of disentangling water vapor effects from BBA effects but also the circulation adjustments included in the large-cale semi-direct BBA effects.
- Modified in the conclusion: "However, the grouping of high and low BBA days has been shown to lead to meteorological sampling biases, complicating the separation of meteorological effects from the **"large-scale"** semi-direct BBA effects."

L366, perhaps, given the primary role of BBA - modulating the large-scale synoptic pattern/circulation, the use of this regression model is not suited?

Whether it is due to missing information in the predictors, non-linearity of the system, or a predictand that is too complex, the model exhibits underfitting issues. That said, it still provides valuable insights by quantifying the processes leading to delays in the dissipation times of FLC in the region. That's why we argue that, despite its limitations, the model is well-suited for this study and improves our understanding of the system.

L375, one possible way of doing this could be the "synoptic matching (or locking)" method used in Quaas et al. (2021, ERL), where, given a location and time, they search in their climatological database for a day with the synoptic pattern that matches the current one the best, but with different aerosol states.

This is a nice idea; however, as we can see in this study, the synoptic patterns in the region between high and low BBA days are so different that they will likely lead to an extremely small data sample. Therefore, I don't think this approach would work here. However, expanding the study area or changing the definitions of high and low BBA days could definitely result in an interesting study.

L375-377, given that CALIPSO is used (?) already in this study and that the key issue is the covariation between meteorology and BBA, I struggle to see how can EarthCARE offer better ways to disentangle aerosol effects.

You are right, EarthCARE will improve our understanding of aerosl-cloud interactions but it will not directly address the covariation between meteorology and BBA.
-Deleted in the conclusion: "and potentially help disentangle aerosol effects"

References

Andersen, H. and Cermak, J.: First fully diurnal fog and low cloud satellite detection reveals life cycle in the Namib, Atmospheric Measurement Techniques, 11, 5461–5470, https://doi.org/10.5194/amt-11-5461-2018, 2018.

Gueymard, C. A. and Yang, D.: Worldwide validation of CAMS and MERRA-2 reanalysis aerosol optical depth products using 15 years of AERONET observations, Atmospheric Environment, 225, 117 216, https://doi.org/10.1016/j.atmosenv.2019.117216, 2020.

Pauli, E., Cermak, J., and Andersen, H.: A satellite-based climatology of fog and low stratus formation and dissipation times in central Europe, Quarterly Journal of the Royal Meteorological Society, 148, 1439–1454, https://doi.org/10.1002/qj.4272, 2022.