Reply to RC1

We thank the referee for the very helpful comments. We have revised the manuscript accordingly. In the following please find our replies to the individual comments.

> This manuscript from Klingmüller and Lelieveld presents a dust scheme that leverages techniques from machine learning (ML) to help represent dust emitted from arid and semiarid regions. The model setup is well described and the comparison with observations is sufficiently clear to appreciate regions where ML improves upon the classical representation of dust emission that is currently used in the EMAC model used by this group. The paper is sufficiently well written to warrant publication but I would like to see minor questions answered that can improve on it and help the reader follow the choices that have been made in this work.
> A first thing that should be explained since it is can be seen as both a wise or a curious choice is why the authors choose to compare the model results to the dust aerosol optical depth at 10 mm instead of the more classical comparison at a wavelength of 550nm.

We have added the motivation "In the infrared, coarse particles scatter radiation more efficiently than submicron particles. Therefore, observations at a wavelength of 10 μm are most sensitive to mineral dust and sea salt and are little affected by other aerosols, resulting in a particularly reliable DAOD retrieval."

> Lines 73-75: please explain to the reader why you need to normalize surface friction velocity, soil moisture, snow depth, KAI, geopotential and clay fraction, is it inherent to the way the ML technique is used? It is hard for the reader to guess.

We have added the explanation "This normalisation maps all input variables and associated gradients into a similar and relatively small range to support the model parameter optimisation routine during the training process."

> Line 109: you imply that the dust emissions are injected in the first 4 lowest layers of the model, if that is the case, please give the rational of why you chose to do that instead of injecting dust in the lowest model layer and what are the averaged heights of these first 4 layers.

The dust emissions are only injected into the lowest of the 4 layers, which is now more clearly stated.

> Line 123: the text seems to indicate that aerosol wet scavenging is a function of total precipitation in a model gridbox. Physically, this is not the choice since aerosols are scavenged as a function of the amount of precipitation formed in the aerosol layer and is also dependent on the rate of precipitation coming from above. Please explain better the choice made here.

Considering only the leading order effect that more precipitation means more scavenging is one of the simplifications to reduce the complexity of the model. The motivation is that, due to the transport between emissions and scavenging, primarily the main statistical properties of the latter influence the emissions submodel during training. In retrospect, this seems justified by the good performance of the trained emissions in the EMAC model with its very detailed wet scavenging parametrisation. We have added: "This approximation considers the correlation between precipitation and dust removal as the dominant statistical relationship with the strongest influence on dust source parameters during training. More detailed parametrisations could improve future versions of the model".

> Line 143: this is the first time you mention f_loss which is defined below in line 155. You should at least introduce what f_loss represents before this line.

We now avoid mentioning f_{loss} in this line.

> Lines 177-178: "The temporal correlation coefficients of the observed and predicted hourly DAOD values within each grid cell are typically greater than 0.5 over the regions affected by desert dust (Fig. 3)." You give the impression to the reader that the correlation coefficients are always above 0.5, as you describe later on in the paper it might be the case for spring and summer and it is not the case for the 2 other seasons.

The temporal correlation discussed in Lines 177f considers all seasons and is typically greater than 0.5 in the dusty regions (Fig. 3). Later on, in lines 183ff, we discuss the spatial correlation, which is typically greater than 0.5 during the dusty seasons. We have added inline equations to make this more clear.

> Caption of Figure 1: Please spell out that MAE stands for mean absolute error and RMSE stands for random means square error.

We now define all 4 numbers shown in the plot in the caption.

> Color chart in Figure 3: the use of red and saturated red make it difficult to appreciate the differences between regions that have a correlation coefficient of 0.5 compare to 0.7 or even 0.9. Please take a color scale that allows to appreciate this differences more accurately.

We have increased the contrast.

> Line 198: To appreciate an annual mean emission of 4.3 Gt/yr it would be informative to give the fraction of this emission that are particles below a diameter of 1mm since they will influence much more the shortwave and Kok et al., (2017) have established a constraint on this fraction as well as on the total emission. Please indicate what is the cutoff of the dust size distribution in your model for comparison with other models. It would be of interest to know how much you emit for the larges regions emitting dust (see paper by Kok et al 2021)

We have added details about the accumulation and coarse mode parameters which clarify that the accumulation mode (which receives 5.3 % of the dust emissions) has a count median diameter $< 1.4\ \mu$m and that there is no strict upper limit for the particle size in the coarse mode (but no additional mode for super coarse particles). We now also present the fractional emissions from the regions considered by Kok et al. 2021 in the new Table 1.

> Line 219: When you explain how EMAC DAOD is obtained when dust and seasalt are present, you should indicate the assumptions made for the density of dust and of seasalt to allow other researchers to make a comparable evaluation.

We have added the densities.

> Color bar of Figure 7: you should extend this color bar as the AOD scales of 0.0, 0.1,... 0.8 are too close one another to be legible.

We have extended the colour bar.

> You could have pushed further the comparison with observations by comparing yearly mean dust deposition over the globe, this is done et Checa-Garcia et al., (2021) for instance.

We have added a comparison with deposition observations in the new Fig. 12.

> Thank you for this interesting contribution.
> References:
> Kok et al., 2017, Smaller desert dust cooling effect estimated from analysis of dust size and abundance, Nat com., doi: 10.1038/ngeo2912
> Kok, J. F., Adebiyi, A. A., Albani, et al., 2021: Contribution of the world's main dust source regions to the global cycle of desert dust, Atmos. Chem. Phys., 21, 8169–8193, https://doi.org/10.5194/acp-21-8169-2021.
> Checa-Garcia, R., Balkanski, Y., Albani, S., et al.: Evaluation of natural aerosols in CRESCENDO Earth system models (ESMs): mineral dust, Atmos. Chem. Phys., 21, 10295–10335, https://doi.org/10.5194/acp-21-10295-2021, 2021.