This paper presents an evaluation of a forecast of mineral dust in support to the CADDIWA experiment over the western part of Africa. Also, an innovative way to produce an ensemble forecast is presented by using lead time of a single model instead of using several models. The coupled regional model WRF-CHIMERE is deployed in forecast mode during the summer 2021. I think this publication is suitable for publication after the following corrections.

Major comments

- My feeling is that the new methodology (using an ensemble of forecast lead) proposed in this paper could be applied to the general purpose of weather forecast. Since the WRF-CHIMERE also deliver data, I strongly suggest to apply the method to the main meteorological data available in this zone at synoptic stations like airports as shown on this map (white spots):

- Also, instead of showing a difference between two lead times in appendix, I would discuss the variability of forecast lead in terms of standard deviation. A standard deviation is appropriate to describe a variability.

Minor comments

The paper deserves to be carefully checked, there are many typo errors. Among them:

I would suggest to reformulate the title like “Variability of mineral dust forecast during the 2021 CADDIWA experiment”

I have reformulated the abstract like this:

“As an operational support to the CADDIWA field campaign, the coupled regional model WRF-CHIMERE is deployed in forecast mode during the summer 2021. The simulation domain covers the West Africa and the East Atlantic and allows the modeling of dust emissions and their transport to the Atlantic. On this route, we find Cape Verde which was used as a base for measurements during the CADDIWA campaign. Meteorological variables and mineral dust concentrations are forecasted on a horizontal grid with a 30 km resolution and from the surface to 200 hPa. Each day, the simulation starts the day before (D-1) and up to five days ahead (D+4). For each day, we thus have six different calculations, with expectantly a better precision the closer we get to the analysis (lead D-1). In this study, a quantification of the forecast variability of wind, temperature, precipitations and mineral dust concentrations according to the modelled lead is presented. It has been shown that the forecast quality

does not decrease with time and the high variability observed on some days for some variables (wind, temperature) does not explain the behavior of other dependent and downwind variables (mineral dust concentrations). A new method is also tested to create an ensemble without perturbing input data, but considering six forecast leads available for each date as members of an ensemble forecast. It has been shown that this new forecast based on this ensemble, is able to give better results for two AERONET stations on the four available for Aerosol Optical Depth. This could open the door to further testing with more complex operational systems."

Line 16: this sentence is strange "Leading the African continent to arrive above the Atlantic sea, they can generate tropical storms", please reformulate

Line 34: analyzed

Line 43: Change "And for those institutes that don't have the computer resources to do ensemble simulations" to "And for institutes which do not have sufficient computing resources to perform classical ensemble simulations"

Line 58: again five days?

Lin 76: I would say "The goal is not to perform a comparison of model with observations"

Line 135: diagnoses

Line 157: I am a bit surprised by the definition this normalized RMSE, can the author justify this formula?

Line 174: I would write : "desert- and a maritime-infuenced…"

L191-193: the last sentence needs to be reformulated

L201: what "model realizations" means?

L209: I would say "Having only si**x** members, once ranked from the lowest to the highest, this value is in fact the mean average of the 3th and 4th members"

L211: Through the paper harmonize the dates, 1 September and 31 September here, but check everywhere. In any case 31th is not correct.

L213 : realization?

L224: laos?

L254: Capo verde or Cape verde, check in the whole paper and harmonize

I suggest this conclusion:

"In this study, the first goal was to examine the variability of the forecast. This forecast was daily performed for six days, during the period August to October 2021 and as support for the CADDIWA field campaign. For meteorological variables (2m temperature, 10m wind speed, total precipitation rate) and surface concentrations of mineral dust, the day to day variability was quantified. Comparisons at two sites were performed, Bodélé (desert area and important source of dust) and Capo Verde (where the measurements of DACCIWA were coordinated). It has been shown that the wind speed is highly variable for day to day forecast while the temperature is stable over land but more variable over sea and shores (Capo Verde being a group of little Islands). The less stable parameter is

the precipitation at one location when the model may forecast an event one day and not at all the day after (and vice-versa).

Fits, one goal of the study was to examine if a large forecast variability at one site (such as Bodélé) may have a visible impact at a downwind remote site (such as Capo Verde). No evidence of a transport of variability (or a transport of stability) was found during the forecast. The large variability of wind speed, precipitation and temperature induce a large variability of the surface concentration of mineral dust. Between forecast leads, large differences were found both for the correlation and the bias. Considering the model configuration used for this study where no direct and indirect effects of aerosols on meteorology and only mineral dust as natural emissions was taken into account, this variability could be underestimated.

A next study could be to replay this forecast with a model version including all anthropogenic and natural emissions in the CHIMERE model with an exhaustive evaluation with the measurements of the experiment to come.

Second, a new way combining forecast leads was tested to improve the predictions. Considering that several forecast leads may be considered as the members of an ensemble, they are combined from (D-1) to (D+4) for all coinciding dates computing the mean and median values. These new "forecast leads" are compared, with all others members, to the Aerosol Optical Depth measurements of AERONET using correlation, nRMSE and bias statistics. It is noticeable that the forecast is not impaired when increasing the lead time. But it is also noticeable that out of four sites, the best scores for two sites are with the ensemble for the period of the CADDIWA campaign. It is not the case for an extended analyzed period, highlighting that the scores are close from one lead to another. The ensemble methodology provides the best scores when the AOD values are the most important and the most variables in time.

This result opens perspectives for forecasting in general. It would be interesting to test this hypothesis on operational systems: if the combination of the previous forecasts allows to improve the initial conditions of a new forecast, it would allow to perform less ensemble simulations for the same day and thus to reduce considerably the computing cost.”