



Supplement of

DustNet (v1): skilful neural network predictions of dust aerosols over the Saharan desert

Trish E. Nowak et al.

Correspondence to: Trish E. Nowak (trish.research@nym.hush.com) and Stefan Siebert (s.siebert@exeter.ac.uk)

The copyright of individual parts of the supplement might differ from the article licence.

S1 Supplements for Methods section

S1.1 AOD imputation

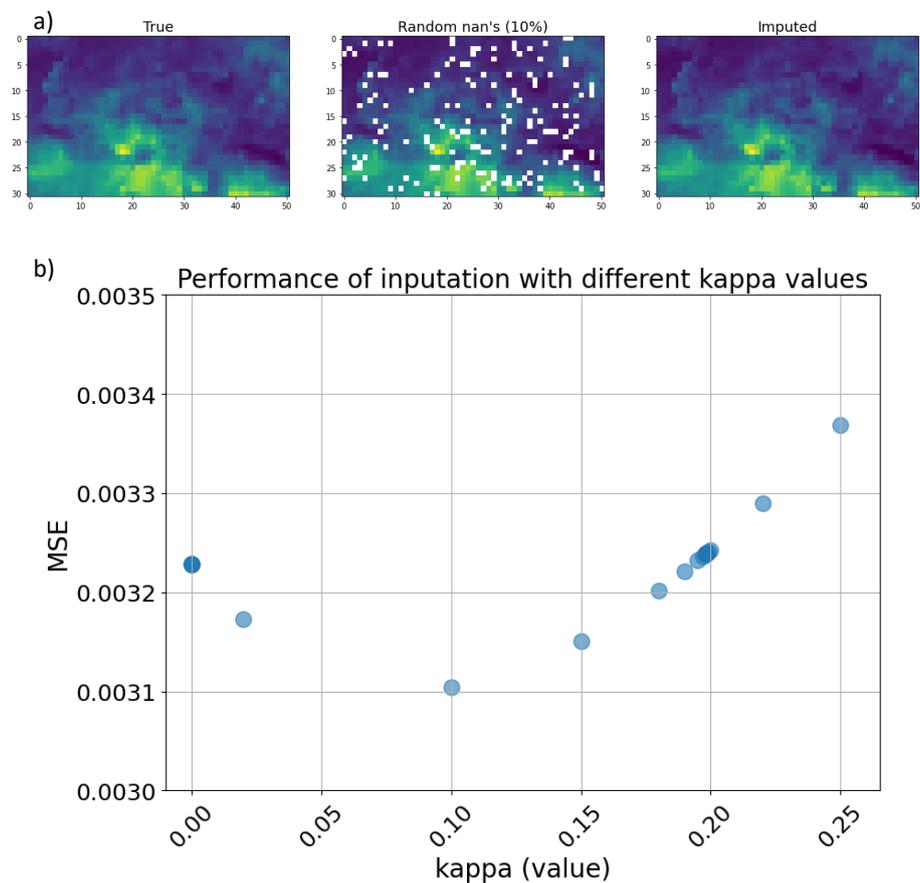


Figure S1. Performance validation of imputation code. Three visual inspections of imputation in **a)** show the "true" values on the left, randomly assigned 10% of missing values in the middle, and imputed with Lattice Kriging method values in the right panel. In **b)** the results of MSEs between "true" and imputed values are displayed for different kappa values (Kriging hyperparameter) used during imputation. We tested 24 different kappa values ranging between 0.000001 and 0.1998. The MSE appears to be insensitive to changes in kappa values producing only marginal improvements in the overall MSE.

S1.2 DustNet - training *versus* validation

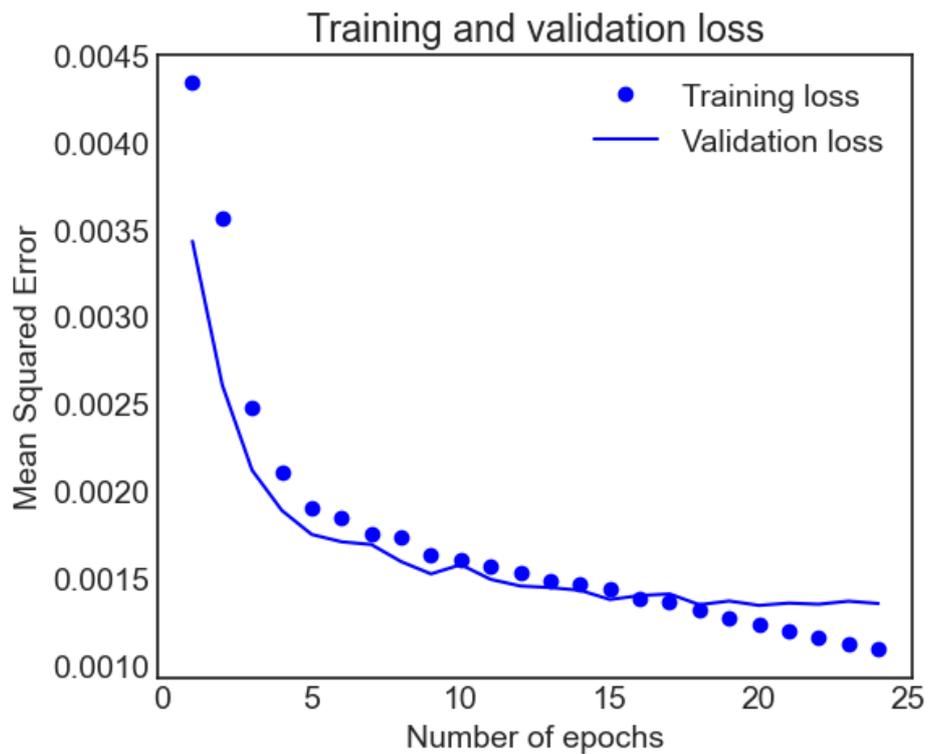


Figure S2. Training and validation loss for the optimal model - DustNet. The model's architecture ensures Early Stopping is performed following the 4th iteration without any improvement in validation loss. Here, stopping occurred after 24 epochs and the model with the lowest ratio of training to validation loss was saved and used for predictions.

S1.3 Daily AOD from MODIS *versus* DustNet predictions - visual inspection

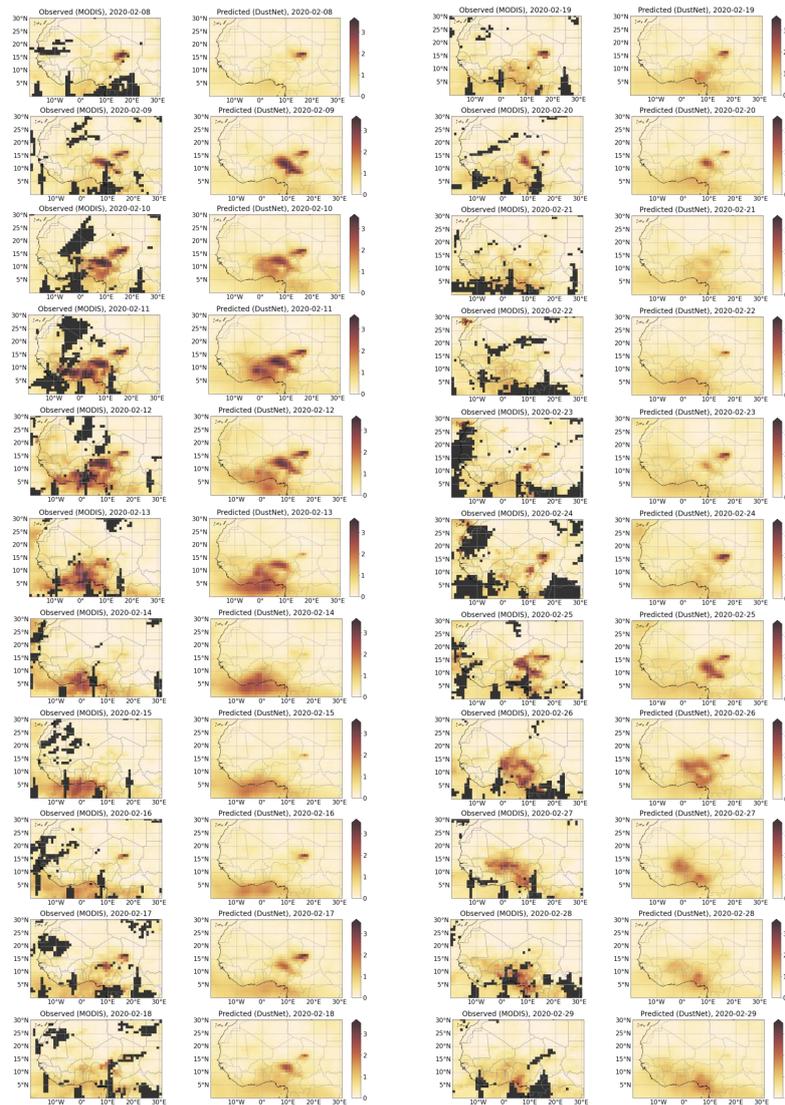


Figure S3. Visual comparison of daily AOD values as observed by MODIS (mean of Aqua and Terra) (**left panel** in both columns) and corresponding DustNet predictions (**right panel**) for selected continuous 3 weeks (22 days), from 8th - 29th February 2020. The dark grey colour in the MODIS maps represents missing values. Despite an initial assumption of heavy reliance on the past 5 days of AOD during training, DustNet presents a skillful ability to predict the next time-step (24-hr) which visibly differs from the last 5 days. This is evident on 13th -14th Feb and 21st Feb, where the AOD values start to decrease despite an increasing past trend. Similarly, prediction of an increasing AOD from 22nd - 26th Feb was captured, despite the previous 5 days of decreasing AOD. The south-western direction of aerosol transport during boreal winter is also skillfully captured (10th - 14th Feb), as is the position of the Bodélé Depression during dust generation episodes (22nd - 26th Feb), but without overly relying on this location as a constant dust source (27th - 29th Feb).

5 S2 Supplementary figures for Results

S2.1 Density curves of daily mean AOD predictions

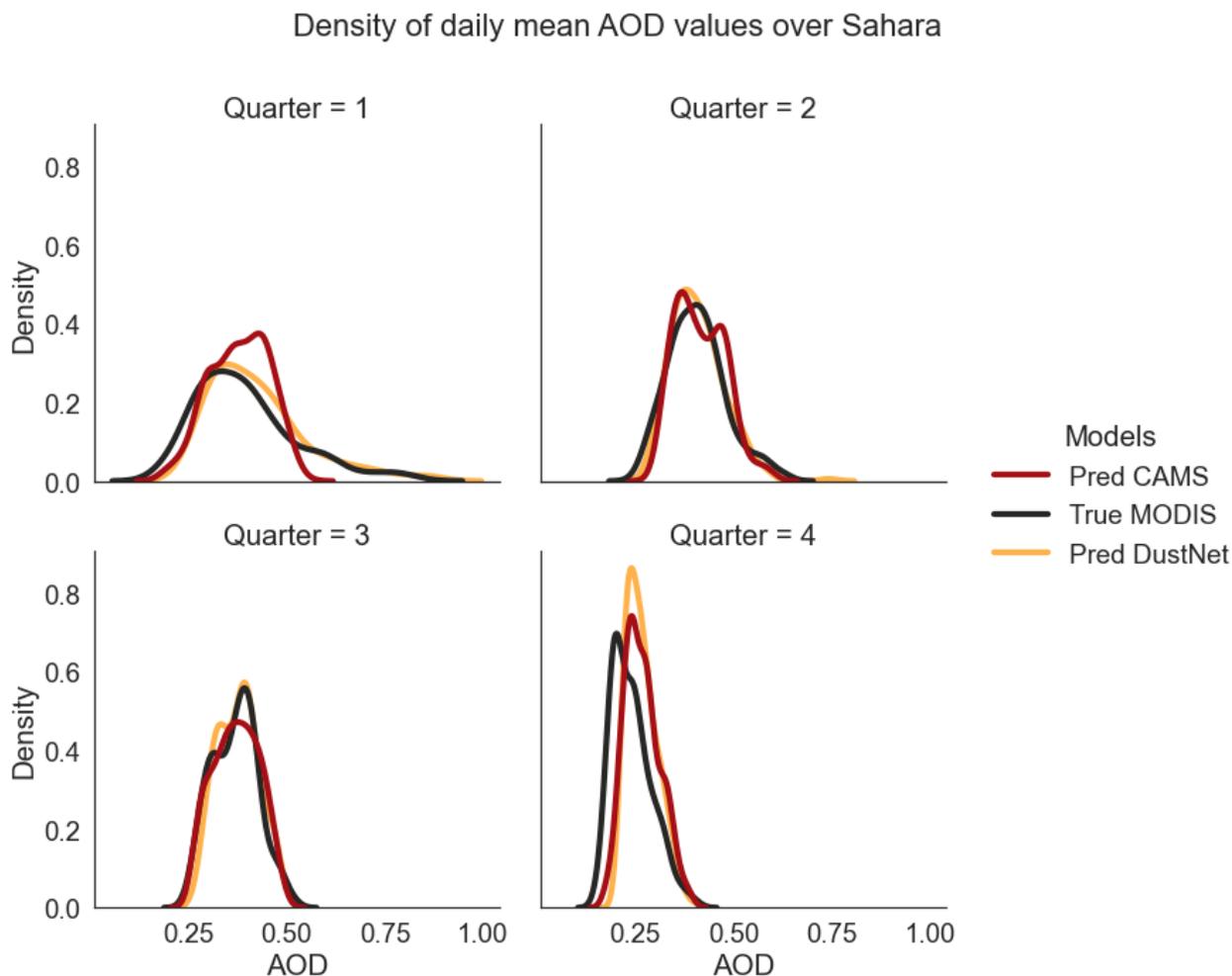


Figure S4. Seasonal mean distribution of daily predicted and true AOD values. The data was averaged over the study region for the testing period of 2020-2022, and shows CAMS forecasts (red) DustNet predictions (yellow) and ground-true MODIS (black). The long tail, indicative of higher AOD values, is clearly missing in the CAMS distribution (red) for Quarter 1: January - March, while the lower AOD values are overestimated. The opposite is true for Quarter 4: October - December, where lower AOD values tend to be underestimated by both CAMS and DustNet in comparison to MODIS. Both models forecast fairly well during Quarter 2 and 3, although DustNet captures the bimodal distribution of AOD in Quarter 3 more skillfully than CAMS.

S2.2 Feature importance

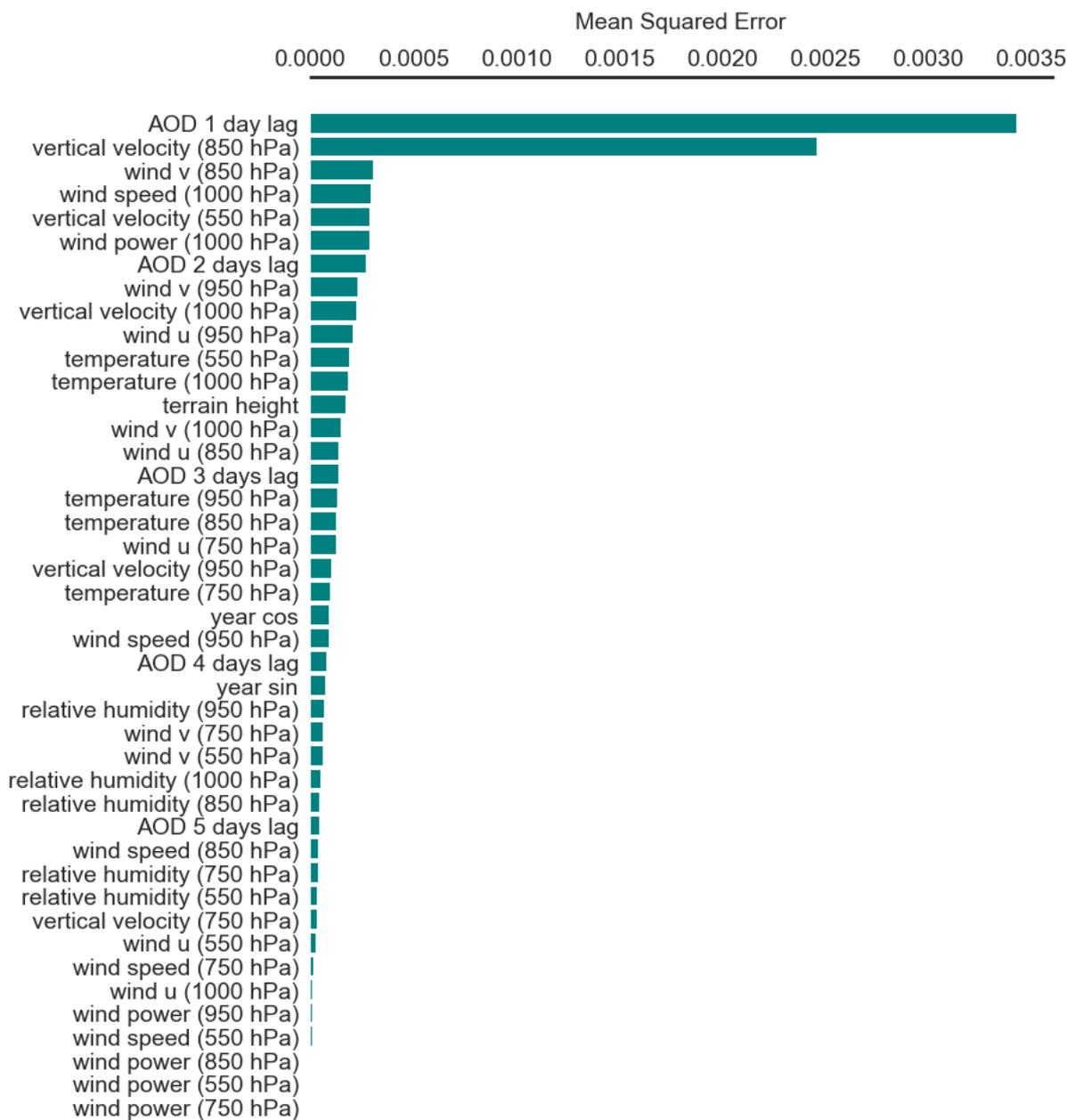


Figure S5. Feature importance analysis for all features included in the DustNet model based on Mean Squared Error (MSE). Features with the highest MSE indicate the highest relevance to the model predictions, while features with the lowest MSE are deemed as nearly irrelevant.