EGU22-5746
https://doi.org/10.5194/egusphere-egu22-5746
EGU General Assembly 2022

Model Output Statistics (MOS) and Machine Learning applied to CAMS O_3 forecasts: trade-offs between continuous and categorical skill scores

Hervé Petetin[1], Dene Bowdalo[1], Pierre-Antoine Bretonnière[1], Marc Guevara[1], Oriol Jorba[1], Jan Mateu armengol[1], Margarida Samso Cabre[1], Kim Serradell[1], Albert Soret[1], and Carlos Pérez García-Pando[1,2]

[1]Barcelona Supercomputing Center, Barcelona, Spain (herve.petetin@bsc.es)
[2]ICREA, Catalan Institution for Research and Advanced Studies, Barcelona, Spain

Air quality (AQ) forecasting systems are usually built upon physics-based numerical models that are affected by a number of uncertainty sources. In order to reduce forecast errors, first and foremost the bias, they are often coupled with Model Output Statistics (MOS) modules. MOS methods are statistical techniques used to correct raw forecasts at surface monitoring station locations, where AQ observations are available. In this study, we investigate to what extent AQ forecasts can be improved using a variety of MOS methods, including persistence (PERS), moving average (MA), quantile mapping (QM), Kalman Filter (KF), analogs (AN), and gradient boosting machine (GBM). We apply our analysis to the Copernicus Atmospheric Monitoring Service (CAMS) regional ensemble median O_3 forecasts over the Iberian Peninsula during 2018–2019. A key aspect of our study is the evaluation, which is performed using a very comprehensive set of continuous and categorical metrics at various time scales (hourly to daily), along different lead times (1 to 4 days), and using different meteorological input data (forecast vs reanalyzed).

Our results show that O_3 forecasts can be substantially improved using such MOS corrections and that this improvement goes much beyond the correction of the systematic bias. Although it typically affects all lead times, some MOS methods appear more adversely impacted by the lead time. When considering MOS methods relying on meteorological information and comparing the results obtained with IFS forecasts and ERA5 reanalysis, the relative deterioration brought by the use of IFS is minor, which paves the way for their use in operational MOS applications. Importantly, our results also clearly show the trade-offs between continuous and categorical skills and their dependencies on the MOS method. The most sophisticated MOS methods better reproduce O_3 mixing ratios overall, with lowest errors and highest correlations. However, they are not necessarily the best in predicting the highest O_3 episodes, for which simpler MOS methods can give better results. Although the complex impact of MOS methods on the distribution and variability of raw forecasts can only be comprehended through an extended set of complementary statistical metrics, our study shows that optimally implementing MOS in AQ forecast systems crucially requires selecting the appropriate skill score to be optimized for the forecast application of interest.

Petetin, H., Bowdalo, D., Bretonnière, P.-A., Guevara, M., Jorba, O., Armengol, J. M., Samso Cabre, M., Serradell, K., Soret, A., and Pérez Garcia-Pando, C.: Model Output Statistics (MOS) applied to CAMS O_3 forecasts: trade-offs between continuous and categorical skill scores, Atmos. Chem. Phys. Discuss. [preprint], https://doi.org/10.5194/acp-2021-864, in review, 2021.