The comments by the reviewer are in black and our replies to the reviewers are in **bold blue**.

RC1

The manuscript, “Aerosol characteristics and polarimetric signatures for a deep convective storm over north-western part of Europe – modeling and observations,” details the development and evaluation of an extended modeling system. This enables high-resolution modeling of aerosol-cloud interactions in the Terrestrial Systems Modeling Platform (TSMP), after extension with a chemical transport model and the polarimetric radar forward operator. The TSMP was evaluated at convection-permitting horizonal resolution against observations of a thunderstorm which took place in July 2015 over northwestern Germany. Overall, the extension of the model with the chemical transport model and polarimetric radar forward operator adds valuable capability to model predictions.

We are very thankful for the reviewer’s critique.

Beyond this, however, this reviewer found it difficult to pull out what the key conclusions of this study actually were, and what benefits the extended TSMP platform provided. The manuscript suffers from too much unnecessary detail, and sometimes repeated itself, and not enough necessary detail. Not only did the lack of clarity and explanation make it difficult to identify the key findings of the study, it also made it difficult to evaluate the claims and evidence that were more clearly presented.

In the revised manuscript, we have highlighted the key findings from the study: 1) Convective storm event in the model generates "aerosol tower" like features; 2) Model is able to capture the spatial pattern of precipitation and polarimetric features but with some biases; 3) Constraining the cloud droplet size distribution (CDSD) did not produce any drastic difference in the precipitation and CFADs of polarimetric variables but produced improvement in the simulated ZDR column-like features. Thus, running simulations with prognostic aerosol and use of forward operator can also additionally help to constrain the cloud droplet size distribution in the model.

The methods and results sections have been better explained, and extended details in the study have been moved to the supplementary to simplify the manuscript and better emphasize the main messages.

For example, none of the polarimetry aspects of the study were clearly explained; prior knowledge of both polarimetric radar and how they are interfaced to/emulated by models seems to be necessary. Even more specifically as an example, polarimetric radar measurements such as horizonal reflectivity, differential reflectivity, specific differential phase, and the cross-correlation coefficient are introduced and the TSMP’s performance evaluated for these parameters, but the manuscript does not explain to which meteorological or chemical parameters they relate, what these parameters tell us about the atmosphere and the model’s performance. This reviewer is familiar with regional and high-resolution modeling and aerosol-chemistry-convection interactions but not polarimetric radar, and assumes that this will be the case for at least some readers, so that much of the manuscript dealing with the polarimetric model implementation (even how it's interfaced with the model is unclear) and the related results are a bit shrouded in mystery. As a somewhat less critical example, the manuscript does not define what it considers to be northwestern Germany for the purpose of the study, and later on adds in mention of two additional German cities/radar sites, with insufficient location information for the deep convective event itself; a map may have been useful. Methods and results are not clearly explained – how are the column NO2 observations from satellite swaths and AERONET point observations compared to a model grid box, as another example. Why was this study undertaken in the first place is also left unclear (though such a modeling platform does have great potential).

We agree with the reviewer that lack of prior knowledge of polarimetric radar and interfacing of FO with the model might hinder some readers. While polarimetric radar observations are still comparably new, we think that a detailed explanation and further their sensitivities is beyond the scope of this manuscript. So, to address the reviewer's concern, we have added references to literature on radar polarimetry in general (e.g. Ryzhkov and Zrnic 2019; Kumjian et al. 2013; Trömel et al. 2021) and its use for microphysical fingerprinting in particular. We have also added a discussion about the specific polarimetric signature like the ZDR column that plays a role in our study. Further, we have added additional details on the forward operator (FO), including references to papers with in-detail descriptions of the FO and e.g. its coupling to the COSMO model (Zeng et al. 2016; Trömel et al. 2021; Shrestha et al. 2022a,b), to allow the reader to better understand this part without (further) (over)loading this manuscript. We have also now better explained the benefits of the extended TSMP platform.

The choice of study region has been better motivated in the revised manuscript. Fig. R1.1 shows the model domain with the extents of the two polarimetric radars: JuXPol and BoXPol. The dotted lines show the extent of the model domain used for evaluation purposes excluding the relaxation zone.

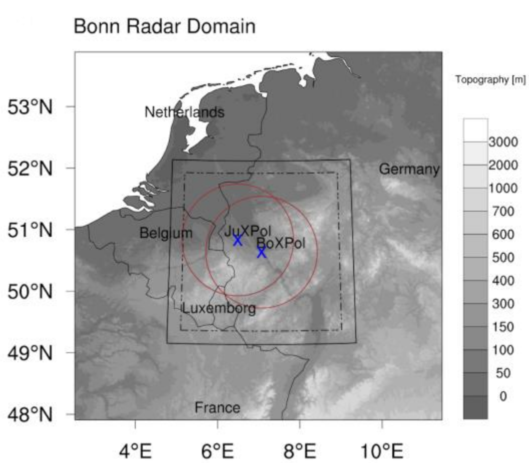

Figure R1.1: Bonn Radar Domain (included in the revised manuscript).

In the manuscript, column NO_2 observations from satellite swaths are compared with the model qualitatively (more quantitative comparison is discussed below) and AERONET point observations are compared to a model grid box nearest to the AERONET location.

Finally, the study was undertaken to improve our understanding of aerosol-cloud-precipitation interactions, by overcoming shortcomings in the representation of aerosol dynamics in the model and by evaluating the model in polarimetric radar space.

These examples are meant to be illustrative, as this problem is ubiquitous in the manuscript, and greatly impedes identification, interpretation, and evaluation of results.

We are thankful to the reviewer's constructive criticism. We have made an effort to bring overall clarity in the manuscript in terms of methods, results and discussion of findings.

Two Specific Comments:

Line 111: It is not really accurate to say that AERONET is considered to have a better accuracy than MODIS – one instrument is essentially a point-based observation and so is often likely more accurate for that specific location than a satellite, while the other observes over a larger swath and so loses some of the horizontal detail but can fill in the gaps between AERONET instruments. And both require a retrieval to turn the raw observations into useful measurements of aerosol properties, which introduces its own uncertainties. Which instrument is "better" for a study depends on what you're trying to compare or investigate. It's more accurate to say that MODIS and AERONET are complementary, and each has its advantages and disadvantages.

We agree with the reviewer that MODIS and AERONET are complementary to each other, and provide valuable data to compare spatial patterns and time-series individual locations, respectively. However, we maintain our statement that AERONET AOD observations are more accurate than MODIS. This statement has been made in many previous studies, see Giles et al. (2019) and references therein. It is much easier to retrieve AOD by looking from below the atmosphere to the sun (as done by AERONET) than observing reflected sun-light (as done by MODIS), where it is difficult to disentangle surface reflectance from aerosol scattering effects. AERONET has therefore always served as a ground-truth for MODIS observations. AERONET AOD observations have a 1-σ uncertainty of 0.02 (Giles et al., 2019), whereas MODIS has an uncertainty of 0.05$\pm$15 % (Levy et al., 2013), and data quality gradually decreases with increasing surface reflectance (Wei et al., 2018). Note also that by measuring sky radiance at multiple scattering angles in addition to direct-sun observations, AERONET can much more accurately determine further aerosol properties such as aerosol size distribution and phase-function than MODIS, where aerosol types and corresponding properties are prescribed in look-up tables as a function of time and space (Levy et al. 2007, Levy et al. 2013). We changed the sentence in revised manuscript to "These measurements have a better accuracy than MODIS (Giles et al. 2019) but are only available at a few locations."

Line 245: Comparison of satellite-based column observations to the NO2 columns computed from the model output is more than simply a first-order evaluation, but this again depends to an extent on what exactly you want to evaluate and be able to say with your study. Column quantities are quite useful for many applications, and again, the satellite-based column observations fill in gaps between ground-based or other types of in situ instruments and can, for example, provide information on chemical transport and transformation. The brief evaluation of the NO2 column comparisons already demonstrate a number of potential results and areas for model refinement that could be classified as more than first-order. Also, why were other trace gases not evaluated, such as O3 or HCHO which are available from satellite data products? And equally importantly, the satellite products come with specified uncertainties and data flags, so it is not enough to only acknowledge the uncertainty in the satellite estimates – the uncertainty in the comparison can therefore be quantified and would facilitate evaluation of the comparison.

The reviewer is right that this is more than just a first-order comparison since we are directly comparing model-simulated NO_2 VTCs with satellite retrieved VTCs over the same domain and at approximately the same time. What we tried to say is that this is a very limited evaluation since we are comparing the model with observations for a single day only. The product indeed comes with an estimated uncertainty for each VTC that could be used for the comparison between satellite and model. According to Boersma et al. (2011), Boersma et al. (2018) and Lamsal et al. (2021) typical uncertainties are of the order of 30% under clear-sky-conditions. This should also hold for the data used in our study, since we followed the recommenations in the Readme document available at https://disc.gsfc.nasa.gov/datasets/OMNO2_003/summary(i.e. filtering for data with a quality flag of zero and with a cloud radiance fraction < 0.5, as already mentioned in the manuscript). In order to compare simulated and observed NO2 VTCs more quantitatively, we have additionally created a scatter plot between simulated and observed VTCs (see Fig. R1.2). For this, we have interpolated the model output over the individual OMI pixels using an inverse distance squared algorithm.

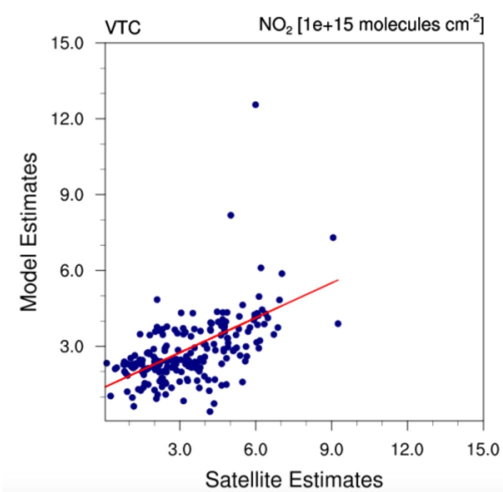

Figure R1.2: Scatter plot with regression line of NO2 VTCs between satellite and model estimates (r= 0.46). The model output was interpolated over the individual OMI pixels using an inverse distance squared algorithm.

We haven't included OMI HCHO and ozone for the following reasons: The HCHO product is extremely noisy. The retrieval uncertainty is 50-105%, with the lower end being valid only for highly polluted locations. HCHO products are therefore usually only presented as monthly, seasonal or yearly averages (e.g. De Smedt et al. 2021). An illustration of the extremely high noise in OMI data compared to MAX-DOAS measurements is shown in their Fig. 15. The same argument holds for SO_2. Comparing ozone would be quite interesting, but there is no official OMI tropospheric ozone product. Comparing the total column O_3 from OMI would not be meaningful as the column is strongly dominated by the stratosphere. Note also that due to the long lifetime of ozone, we would expect only very small gradients in the model domain.

The above discussion has been included in the revised manuscript.

References:

Boersma, K. F., Eskes, H. J., Dirksen, R. J., van der A, R. J., Veefkind, J. P., Stammes, P., Huijnen, V., Kleipool, Q. L., Sneep, M., Claas, J., Leitão, J., Richter, A., Zhou, Y., and Brunner, D.: An improved tropospheric NO2 column retrieval algorithm for the Ozone Monitoring Instrument, Atmos. Meas. Tech., 4, 1905–1928, https://doi.org/10.5194/amt-4-1905-2011, 2011

Boersma, K. F., Eskes, H. J., Richter, A., De Smedt, I., Lorente, A., Beirle, S., van Geffen, J. H. G. M., Zara, M., Peters, E., Van Roozendael, M., Wagner, T., Maasakkers, J. D., van der A, R. J., Nightingale, J., De Rudder, A., Irie, H., Pinardi, G., Lambert, J.-C., and Compernolle, S. C.: Improving algorithms and uncertainty estimates for satellite NO2 retrievals: results from the quality assurance for the essential climate variables (QA4ECV) project, Atmos. Meas. Tech., 11, 6651–6678, https://doi.org/10.5194/amt-11-6651-2018, 2018.

De Smedt, I., Pinardi, G., Vigouroux, C., Compernolle, S., Bais, A., Benavent, N., Boersma, F., Chan, K.-L., Donner, S., Eichmann, K.-U., Hedelt, P., Hendrick, F., Irie, H., Kumar, V., Lambert, J.-C., Langerock, B., Lerot, C., Liu, C., Loyola, D., Piters, A., Richter, A., Rivera Cárdenas, C., Romahn, F., Ryan, R. G., Sinha, V., Theys, N., Vlietinck, J., Wagner, T., Wang, T., Yu, H., and Van Roozendael, M.: Comparative assessment of TROPOMI and OMI formaldehyde observations and validation against MAX-DOAS network column measurements, Atmos. Chem. Phys., 21, 12561–12593, https://doi.org/10.5194/acp-21-12561-2021, 2021.

Kumjian, M. R. (2013). Principles and Applications of Dual-Polarization Weather Radar. Part I: Description of the Polarimetric Radar Variables. Journal of Operational Meteorology, 1.

Levy, R. C., Remer, L. A., and Dubovik, O. (2007), Global aerosol optical properties and application to Moderate Resolution Imaging Spectroradiometer aerosol retrieval over land, J. Geophys. Res., 112, D13210, doi:10.1029/2006JD007815.

Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F., and Hsu, N. C.: The Collection 6 MODIS aerosol products over land and ocean, Atmos. Meas. Tech., 6, 2989–3034, https://doi.org/10.5194/amt-6-2989-2013, 2013.

Ryzhkov, A. V., & Zrnic, D. S. (2019). Radar polarimetry for weather observations (Vol. 486). Cham, Switzerland: Springer International Publishing.
Shrestha, P., Trömel, S., Evaristo, R., & Simmer, C. (2022). Evaluation of modelled summertime convective storms using polarimetric radar observations. Atmospheric Chemistry and Physics, 22(11), 7593-7618.

Shrestha, P., Mendrok, J., Pejcic, V., Trömel, S., Blahak, U., & Carlin, J. T. (2022). Evaluation of the COSMO model (v5. 1) in polarimetric radar space–impact of uncertainties in model microphysics, retrievals and forward operators. Geoscientific Model Development, 15(1), 291-313.

Trömel, S., Simmer, C., Blahak, U., Blanke, A., Doktorowski, S., Ewald, F., ... & Quaas, J. (2021). Overview: Fusion of radar polarimetry and numerical atmospheric modelling towards an improved understanding of cloud and precipitation processes. Atmospheric Chemistry and Physics, 21(23), 17291-17314.

Wei, J., Sun, L., Peng, Y., Wang, L., Zhang, Z., Bilal, M., & Ma, Y. (2018). An improved high-spatial-resolution aerosol retrieval algorithm for MODIS images over land. Journal of Geophysical Research: Atmospheres, 123, 12,291– 12,307. https://doi.org/10.1029/2017JD027795

Zeng, Y., Blahak, U., & Jerger, D. (2016). An efficient modular volume-scanning radar forward operator for NWP models: description and coupling to the COSMO model. Quarterly Journal of the Royal Meteorological Society, 142(701), 3234-3256.