

RESEARCH ARTICLE

Impact of Saharan dust outbreaks on short-range weather forecast errors in Europe

Kilian Hermes^{1,2}  | Julian Quinting¹  | Christian M. Grams¹  |
Corinna Hoose¹  | Gholam Ali Hoshyaripour¹ 

¹Institute of Meteorology and Climate Research, Department Troposphere Research, Karlsruhe Institute of Technology, Karlsruhe, Germany

²Institute for Climate and Atmospheric Science, University of Leeds, Leeds, UK

Correspondence

Kilian Hermes, Institute for Climate and Atmospheric Science, University of Leeds, Leeds, UK.

Email: kilian.hermes@mail.de

Funding information

H2020 European Research Council, Grant/Award Number: 714062; Helmholtz Association, Grant/Award Number: SPREADOUT/VH-NG-1243

Abstract

Mineral dust, the most abundant atmospheric aerosol by mass, interacts with radiation directly and alters cloud properties indirectly. Many operational numerical weather prediction models account for aerosol direct effects by using climatological mean concentrations and neglect indirect effects. This simplification may lead to shortcomings in model forecasts during outbreaks of Saharan dust towards Europe, when climatological mean dust concentrations deviate strongly from actual concentrations. This study investigates errors in model analyses and short-range forecasts during such events. We investigate a pronounced dust event in March 2021 using the pre-operational ICOSahedral Nonhydrostatic weather and climate model with Aerosols and Reactive Trace gases (ICON-ART) with prognostic calculation of dust and the operational European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS) model, which deploys a dust climatology. We compare model analysis and forecast with measurements from satellite and in situ instruments. We find that inclusion of prognostic aerosol and direct radiative effects from dust improves forecasts of surface radiation during clear-sky conditions. However, dust-induced cirrus clouds are strongly underestimated, highlighting the importance of representing indirect effects adequately. These findings are corroborated by systematic quantification of forecast errors against satellite measurements. For this we construct an event catalogue with 49 dust days over Central Europe between January 2018 and March 2022. We classify model cells by simulated and observed cloudiness and simulated dustiness in the total atmospheric column. We find significant overestimations of brightness temperature for cases with dust compared with cases without dust. For surface shortwave radiation, we find median overestimations of 6.2% during cloudy conditions with dust optical depth greater than 0.1, however these are not significant compared with cloudy conditions without dust. Our findings show that the pre-operational ICON-ART and the operational IFS model still do not reproduce cloudiness adequately during events with Saharan dust over Central Europe. Missing implementations of prognostic dust, particularly of indirect effects on cloud formation, lead to significant underestimations of cloudiness and potentially overestimations of surface radiation.

KEYWORDS

aerosol–cloud interaction, aerosol–radiation interaction, mineral dust, numerical weather prediction

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *Quarterly Journal of the Royal Meteorological Society* published by John Wiley & Sons Ltd on behalf of the Royal Meteorological Society.

1 | INTRODUCTION

Dust is the most common natural aerosol in Earth's atmosphere by mass (Textor et al., 2006) and interacts with various components of the Earth system (Knippertz, 2014; Kok et al., 2023). Like other aerosol species, mineral dust interacts directly with radiation, by scattering and absorbing solar radiation and absorbing and emitting terrestrial radiation (Liao & Seinfeld, 1998). As a consequence of this direct effect, the absorption of radiation by aerosol alters the stability of the atmosphere. This is known as the semidirect effect (Hansen et al., 1997). Mineral dust particles also interact with cloud microphysical processes by acting as cloud condensation nuclei (CCN) or ice nucleating particles (INP; Karydis et al., 2011). This alters cloud properties such as cloud brightness, cloud-top height, or cloud lifetime, it can lead to generally increased cloudiness, and is known as the "indirect effect" (Lohmann & Feichter, 2005). The Sahara contains Earth's most productive dust sources (Ginoux et al., 2012; Prospero et al., 2002). Significant amounts of Saharan dust are transported northwards (Shao et al., 2011), and outbreaks with strongly elevated dust concentrations frequently affect the Mediterranean and Europe (e.g., d'Almeida, 1986; Moulin et al., 1998; Mona et al., 2006; Merdji et al., 2023). While high wind-speed events are most efficient for causing dust emissions (Cowie et al., 2015), the synoptic processes generating these winds differ spatially and seasonally. Particularly during spring, North African cyclones are efficient for causing dust-emission-generating peak wind speeds over North Africa (Bou Karam et al., 2010; Fiedler et al., 2014). A center for cyclogenesis is in the lee of the Atlas mountains (Flaounas et al., 2022). Under the influence of upper-level troughs and positive differential vorticity advection, surface depressions form and can intensify quickly into cyclones (Bou Karam et al., 2010; Knippertz & Fink, 2006). A similar synoptic situation with a pronounced trough spanning into Northwest Africa, which generates southwesterly flow over the western Mediterranean, is typical for Saharan dust events affecting Europe (Barkan et al., 2005). Interactions with the polar jet can additionally trigger the formation of North African cyclones and enable dust transport far poleward (Francis et al., 2018, 2019). Flaounas et al. (2015) showed that North African cyclones contribute to up to 70% of extreme dust events over the Mediterranean. On a synoptic scale, Fluck and Raveh-Rubin (2023a); Fluck and Raveh-Rubin (2023b) showed that the dry intrusion air stream associated with deep equatorward penetrating high-latitude air masses mobilises dust effectively due to extreme surface winds when reaching the lower levels. Once mobilised, the dust might become embedded in the warm sector of a North African cyclone. Indeed, Francis et al. (2022)

recently showed that atmospheric rivers (AR) can be efficient drivers for Saharan dust transport to Europe, with 78% of ARs associated with severe dust episodes over Europe. More recently, Merdji et al. (2023) showed that dust transport and dust fractional contribution to the total aerosol load is particularly enhanced during summer and spring, and that during these seasons transport tends to happen in higher atmospheric layers than during winter.

Today, models that are currently used for operational numerical weather prediction (NWP) make simplifications with respect to dust to circumvent computational costs. On the one hand, dust emission, transport, and concentrations are not calculated prognostically. Instead, operational NWP models rely on climatological mean values (e.g., ECMWF, 2021, Reinert et al., 2021) for implementing dust effects. On the other hand, even in the models with prognostic calculation of the dust life cycle, the semidirect and indirect effects of dust are only partially implemented (e.g., Rémy et al., 2022).

Previous case studies have assessed whether the inclusion of dust–cloud interaction in models leads to improvements of weather forecasts during dust events in Europe. Rieger et al. (2017) investigated the effect of dust on forecasts of photovoltaic (PV) power generation and found an improvement in radiation forecasts by including prognostic calculations of dust and its effects, where the direct radiative effect from dust dominates the improvements. Weger et al. (2018) found that cirrus cloud cover and ice content, in particular, increase after including prognostic calculation of dust and its effects. Recently, Seifert et al. (2023) suggested a novel subgrid parametrisation which improves the representation of dust-induced cloud decks and thus of the indirect effect of dust.

During spring 2021, several outbreaks of Saharan dust with dust transport towards Europe occurred, leading to elevated dust concentrations (exceeding the mean dust optical depth (DOD) from 2018–2021 tenfold) over large parts of Western and Central Europe for several days. For one of these outbreaks, Magnusson et al. (2021) showed a misrepresentation in cirrus cloud cover, surface radiation, and 2-m temperature over the Iberian Peninsula in the forecast from the operational NWP model of the European Centre for Medium-Range Weather Forecasts (ECMWF), and improvements in forecasts with prognostic calculation of dust and the inclusion of dust radiative effects.

This study follows up on these findings and addresses the following research questions.

- What is the cause for the observed model errors during the dust events in spring 2021? Is dust related to the model errors?

- How large are the typical errors during dust events over Central Europe in an operational model in comparison with satellite observations/retrievals?

Section 2 gives an overview of the data products that are used for this study and summarises model implementations of mineral dust and retrieval methods of satellite products. Section 3.1 presents a case study, which analyses the capability of two current models to reproduce the conditions during a selected dust event in early March 2021 and aims to trace back model errors to their cause. Section 3.2 picks up on the findings from the case study and seeks to generalise the typical synoptic situation during dust outbreaks towards Central Europe and to quantify model forecast errors in brightness temperature (BT) as a proxy for high clouds and surface incoming shortwave radiation (SIS) climatologically over many cases. Finally, Section 4 summarises our results and draws conclusions.

2 | MODEL DESCRIPTIONS AND MEASUREMENT DATA

2.1 | Model data

For the purpose of this study, we use data from two model families: firstly from the ICOSahedral Nonhydrostatic weather and climate model (ICON), which is currently deployed operationally by the Deutscher Wetterdienst (the German Weather Service, DWD), and secondly from the Integrated Forecasting System (IFS) developed by the ECMWF. The following gives a brief summary of model characteristics and implementations of aerosol effects.

2.1.1 | ICON and ICON-ART

ICON is a nonhydrostatic atmospheric model and was developed with the aim of providing a global model for both weather and climate predictions. It is based on an icosahedral grid and allows local grid refinement, so-called nesting (Zängl et al., 2015). The ICON model has been used for operational weather forecasting by the DWD on a global scale since January 2015 and on a regional scale since 2016 (Reinert et al., 2021).

Aerosols and Reactive Trace gases (ART) is a sub-module of ICON, which enables the simulation of the life cycle of aerosols, trace gases, and their interactions in the atmosphere. In ICON-ART, aerosol processes are simulated online, including emission and removal processes. Aerosol particles are represented by lognormal modes, where the median diameter of the distribution is a diagnostic variable (Rieger et al., 2015).

ICON-ART offers a flexible and interoperable framework that can be configured to include tropospheric and stratospheric chemistry (Schröter et al., 2018), aerosol dynamics (Muser et al., 2020), aerosol–radiation (Hoshyaripour et al., 2019; Rieger et al., 2017), and aerosol–cloud interactions (Seifert et al., 2023). At the time of this study, only aerosol–radiation interaction were implemented in the pre-operational forecast system. This forecast runs on the experimental system at DWD, mainly used for research and development purposes. Aerosol–cloud interactions, that is, indirect effects, were not yet implemented into the system used here. Recently, Seifert et al. (2023) suggested a subgrid parametrisation for dusty cirrus, but the full aerosol–cloud interaction remains the topic of ongoing developments.

All data from ICON-ART that we use in this study were calculated on a R2B06 global domain with R2B07 nest over north Africa and Europe with 60 vertical levels, which translates to an effective horizontal grid spacing of 19.7 km. We retrieve DOD and SIS on the native icosahedral grid. The model was initialised with DWD reanalysis at 0000 UTC and we only consider forecast lead times of less than 24 hr to minimise departure from model analysis.

2.1.2 | IFS and IFS-AER

The ECMWF IFS uses prognostic variables for temperature, humidity, and cloud properties as well as monthly-mean climatologies of trace gases and aerosols as input for radiative calculations. The IFS further provides a postprocessing system to simulate cloudy brightness temperatures expected from satellites to allow a like-to-like comparison with the observed satellite imagery. This system takes a large number of atmospheric model profiles into account for input into the radiative transfer model, including temperature, specific humidity, ozone mass mixing ratio, cloud cover, specific cloud liquid, ice, rain, and snow water content, as well as surface parameters including skin temperature, 10-m wind speed, 2-m temperature and dewpoint temperature, volumetric soil water layer, and convective available potential energy (Lupu & Wilhelmsson, 2016). The aerosol climatologies used in the IFS are derived from analysis and short-range forecast data within the Copernicus Atmosphere Monitoring Service (CAMS: Bozzo et al., 2017). The IFS (model cycle 47R3) provides two forecasts from unperturbed initial conditions: a forecast at the regular horizontal grid spacing of 18 km and a forecast with a decreased horizontal grid spacing of 9 km (*hres*). Additionally a medium-range ensemble forecast (*ens*) is provided, containing 50 model runs from perturbed initial conditions, which are generated by

adding small-amplitude perturbations, and stochastically perturbed model physics during the model integration. All simulations are performed with 137 vertical levels (for *eps*, 91 vertical levels before May 11, 2021), spanning from surface pressure to 0.01 hPa at the top level. Model initialisation is performed at 0000 and 1200 UTC (ECMWF, 2021). For the purpose of this study we use data for dust events from several years. For each event we retrieve data from the most recent operational version (model cycles CY43R3–CY47R3). We retrieve total cloud cover (TCC), SIS, and simulated BT from the *hres* run, and remap data onto a regular latitude–longitude grid with 0.25° grid spacing. For analysis we only consider forecast lead times of less than 12 hr, to minimise departure from model analysis.

The Integrated Forecasting System aerosol scheme (IFS-AER) is developed within the CAMS framework and provides the operational IFS model with extensions for simulating tropospheric aerosols, chemically interactive gases, and greenhouse gases. Dust emissions in IFS-AER are computed dynamically using prognostic variables from the meteorological model, while emission threshold surface wind speeds (threshold friction velocity after July 9, 2019) are derived from prognostic variables and a climatology. A bulk-bin aerosol scheme is used for modelling aerosol size distributions. For mineral dust there are three bins within the limits 0.03, 0.55, 0.9, 20 μm , which represent the fine, coarse and super-coarse modes. Since IFS cycle 45r1 (June 26, 2018), operational analyses and forecasts have been performed with interactive aerosols as input for the radiation scheme to compute the direct radiative effect of aerosols. Currently, there is no representation of aerosol–cloud interactions in IFS-AER, however an implementation is planned for the future. The horizontal model resolution is 40 km with 60 (137 after July 9, 2019) vertical levels, spanning from surface pressure to 0.01 hPa at the top level. Forecasts are initialised at 0000, 0600, 1200, and 1800 UTC (Rémy et al., 2019; Rémy et al., 2022). Hereafter we refer to IFS-AER as the CAMS model. For the purpose of this study, we use data from the operational near-real-time forecast (based on IFS model cycles 43R3–47R3). We retrieve DOD remapped to 0.25° grid spacing at the time of model initialisation (analysis).

2.1.3 | ERA5

ERA5 is the most recent global reanalysis product from ECMWF and covers the period from 1950 to the present. It is calculated with the ECMWF IFS, model cycle 41r2, and accounts for aerosol by using climatological mean values. The horizontal grid spacing is 31 km with 137 vertical levels, spanning from surface pressure to 0.01 hPa at the top

level (Hersbach et al., 2020). We use ERA5 data in the case study for assessing the synoptic situation.

2.2 | Observational data

For the evaluation of forecast models we use data derived from instruments on the geostationary *Meteosat Second Generation (MSG)* satellites operated by the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT). All products are from measurements with the Spinning Enhanced Visible and Infrared Instrument (SEVIRI), the main instrument based on the *MSG* satellites. SEVIRI employs 12 spectral channels, spanning from the visible range to the infrared (0.4–13.4 μm). It provides a sampling resolution of 1–3 km at nadir, and is continuously scanning with a 15-min repeat cycle (Schmetz et al., 2002).

2.2.1 | *Meteosat* cloud mask

We use the cloud mask from the EUMETSAT Satellite Application Facility on Climate Monitoring (CM SAF), contained in the product group CLAAS ed. 3.0 (Meirink et al., 2022). Products derived from passive satellite image data require information about the type of scene contained within a pixel. Therefore, cloud masking is an essential step in the processing of *MSG* images. The cloud-masking algorithm relies on measured radiances from seven SEVIRI channels in the visible and infrared and a climatology of cloudiness. It is trained and validated with reference data from the Cloud–Aerosol Lidar with Orthogonal Polarization (CALIOP) on the *Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO)* satellite. The combination of measured radiances and a climatology allows for the estimation of a probability that a pixel is cloudy, given the measured radiance. For the purpose of this study we only use the binary cloud mask, which classifies a pixel as cloudy if the cloud probability exceeds a threshold of 50% and clear-sky otherwise (Karls-son et al., 2020). The *MSG* CM SAF cloud mask covers a circular area with a maximum extent from 81.3° S– 81.3° N and from 81.25° W– 81.25° E on a pixel basis. It is saved on the native SEVIRI grid in satellite projection. For the purpose of this study, we use retrievals from the latest product version 004 on their native grid.

2.2.2 | Surface incoming shortwave radiation

We use SIS from EUMETSAT CM SAF, derived via the Surface Solar Radiation Data Set–Heliosat Edition 2 (SARAH-2) methods (Pfeifroth et al., 2018b). The

algorithm uses the MAGIC SOL method, which is a combination of the Heliosat method (Beyer et al., 1996) with the SPECMAGIC clear-sky model (Pfeifroth et al., 2018). For the calculation of solar irradiance during cloudy conditions, a retrieved cloud albedo and a clear-sky model are used. For the calculation of solar irradiance during clear-sky conditions, a look-up table approach is used which takes into account the effects from aerosol. For aerosols including dust, a modified monthly mean climatology derived from the ECMWF Monitoring Atmospheric Composition and Climate (MACC) reanalysis is used. For water vapour, the retrieval algorithm uses the vertically integrated value from the daily ECMWF *hres* analysis at 1200 UTC. For ozone, monthly mean values of the vertically integrated ozone column from ERA-Interim reanalysis are used. Uncertainties during clear-sky conditions for deviations of 0.1 from the monthly mean aerosol optical depth (AOD) relative to a background AOD of 0.2 are estimated to be about $10 \text{ W}\cdot\text{m}^{-2}$ ($20 \text{ W}\cdot\text{m}^{-2}$) for a solar zenith angle of 60° (0°). For cloudy conditions, these uncertainties are reduced with increasing cloud optical depth. Uncertainties due to deviations in the integrated water vapour of about 1–2 mm result in uncertainties of about $3 \text{ W}\cdot\text{m}^{-2}$, with decreasing uncertainties towards higher water-vapour levels (Pfeifroth & Trentmann, 2018).

Comparison of this SIS product with ground-based observations across Europe has shown high accuracy, with mean absolute errors of $6.9 \text{ W}\cdot\text{m}^{-2}$ relative to measurements. The product has further been shown to be capable of reproducing observed spatio-temporal SIS trends and distributions (Pfeifroth et al., 2018a). Even though the SARA H-2 methods deploy an aerosol climatology, we expect to capture changes in the aerosol indirect effect (Mueller et al., 2015) through measured variations in cloud brightness. For the purpose of this study, we use instantaneous SIS derived via the SARA H-2.1 method, product version 410. It is derived from measurements with SEVIRI on MSG and covers a circular area with a maximum extent from 65° S – 65° N and from 65° W – 65° E with a resolution of 0.5° . This product is available with 30-min time spacing and the same version the of retrieval algorithm from February 20, 2018 onward (Pfeifroth et al., 2018b), and extends beyond the end date of our analysis (March 2021).

As the direct radiative effect of aerosol is implemented with a prescribed climatology, we expect the satellite product to overestimate SIS for cases with DOD above the climatological mean.

2.2.3 | GridSat cloud-top products

The GridSat product combines data from most international meteorological satellites in geostationary orbit and

provides an intersatellite calibrated product, covering most of the globe between 70° S and 70° N from 1980 until the present. It is continuously extended to include most recent dates. The spatial resolution is 0.07° , which is equivalent to about 8 km at the Equator. The temporal resolution is 3 hr. GridSat provides data in three channels (infrared, water vapour, visible: Knapp et al., 2011).

For the purpose of this study, we use the National Oceanic and Atmospheric Administration (NOAA) Fundamental Climate Data Records (FCDR) of brightness temperature near 11 microns. For the examined area of Europe, this is derived mainly from *Meteosat-11* measurements.

2.2.4 | Radiosonde and station data

The Swiss meteorological service MeteoSwiss performs radiosoundings at the station Payerne every day at 0000 UTC and 1200 UTC. We use data from the launch on March 3, 2021 at 1200 UTC.

The DWD deploys a network of stations with instruments measuring SIS. We use measurements from two selected stations, Köln/Bonn and Hohenpeißenberg, which deployed a pyranometer and scanning pyr heliometer/pyranometer (ScaPP), respectively, during March 2021. Instruments are regularly calibrated, and perform automatic dark current corrections, keeping systematic measurement errors minimal. Relative errors for 1-min measurements reach up to 2.5% for the pyranometer, and the ScaPP is expected to perform similarly to the 20% uncertainty reported for hourly values, according to manufacturer information and internal validation by DWD Meteorological Observatory Lindenberg–Regional WMO radiation centre, respectively. We use the climate product of 10-min mean values with 10-min time spacing and automatic quality control and correction (DWD Climate Data Center, 2022).

2.3 | Data processing and averaging

For validation of forecast data against satellite observations, we perform several operations for spatial and temporal data alignment. We perform all calculations on the resolution of the respective model grid, hence we remap satellite data to the model grid. For TCC, BT, and SIS, this is the regular grid used in the IFS *hres* run. For comparison of SIS with ICON(-ART) values, this is the icosahedral grid used in the operational ICON model. Radiation data from ICON and IFS models are given as three-hourly mean and accumulated values, respectively. SIS satellite data from SEVIRI on MSG are provided as instantaneous data every

30 min, which we average over 3-hr periods to gain a measure comparable with the model values. In the following we refer to data from both variables, the instantaneous BT at 1200 UTC and the 3-hr mean SIS from 1200–1500 UTC, as the 1200 UTC values.

2.4 | Event selection

Currently operational weather forecast models such as ICON or IFS do not include prognostic aerosols and their direct, semidirect, and indirect effects. To explore whether there is a consistent, statistically robust error during dusty conditions, and in order to quantify the typical model errors during such conditions, we perform a statistical analysis. Inconsistent model versions, satellite instruments, or retrieval algorithms can limit data consistency. Due to changing retrieval algorithms of satellite products and data availability, we limit our analysis to the period from January 2018–March 2022.

We create an event catalogue of multiple events with elevated dust concentrations over the focus area of the case study. This area covers Austria, Belgium, continental France (without Corsica), Germany, Liechtenstein, Luxembourg, Netherlands, Switzerland. With this we limit analysis to approximately the area in which the case study shows large model errors, as well as to land areas to avoid skewed results from potential variations in the quality of satellite retrievals over land or water. Other than previous studies using relative deviations of on aerosol index from mean values to define dusty days (Barkan et al., 2005), we use absolute thresholds for DOD for differentiating from background aerosol and in order to ensure the selection of events with high absolute dust loadings. We classify cells as dusty if the simulated DOD exceeds a background aerosol threshold value of 0.1. This threshold is consistent with the background AOD threshold used by Holben et al. (2001) and Kambezidis and Kaskaoutis (2008), and of similar magnitude to the dust aerosol climatology used in the ECMWF IFS model for the study area (Bozzo et al., 2017). Additionally we compute the median DOD over the area, in order to select events where a large number of cells show elevated DOD. Based on these computed area values, we manually tune thresholds so that they result in the selection of the dust events over Central Europe during the beginning and end of February 2021 and beginning of March 2021: If more than 25% of cells are classified as dusty and the median DOD is greater than 0.075, we select the date as a dust day. For both criteria, we use DOD data from the near-real-time CAMS forecast at initialisation time, which shows reasonable skill compared with measurements (Flentje et al., 2021; Gueymard & Yang, 2020). Finally we manually analyse the selected

days for other prominent aerosol species over the selected areas, to ensure mineral dust is the dominant aerosol.

2.5 | Case classification

For the quantification of model errors during dust events, reference data are required. At the time of this study, the ECMWF operational analysis does not reproduce the cirrus cloud cover during dusty conditions reliably, hence it cannot be used as reference. In contrast, satellite retrievals for BT and SIS are available as operational products with consistent retrieval algorithms. We hence validate the model forecast directly against satellite products in order to provide a quantitative assessment of model errors. To investigate model errors in the presence or absence of dust, we apply a cell-based approach. For this we classify model cells into different cases: cells affected by the hypothesised source of model error (with dust) and those without the hypothesised source of model error (without dust). Furthermore, we differentiate between cloudy cells and clear-sky cells in order to quantify direct or indirect effects.

Therefore, we classify all model cells by two criteria: firstly by cloudiness, which we derive from the cloud mask and via the output variable TCC (total cloud cover) from the IFS model. In the case of the IFS model, we chose thresholds of TCC <25% and >75% for the classifications clear-sky and cloudy, respectively. We only take values from cells where cloudiness from satellite and model agree. This lowers the risk of the double-penalty problem concerning the cloud cover during the evaluation. Secondly, we classify all model cells by the presence of dust, which we determine via the total DOD from the CAMS model. If the cell value is greater than or equal to the threshold value we classify the cell as dusty; if it is less than the threshold we classify it as not dusty. We use a DOD threshold of 0.1 to differentiate from background aerosol. Additionally, we test the variation of DOD threshold values from 0.01–1 in order to investigate the sensitivity of model errors to increasing dust concentrations. We apply both criteria to each model cell and event, in order to obtain different classes for enabling differentiation between direct effects of dust during clear-sky conditions and indirect effects during cloudy conditions. This results in four specific cases: (1) clear-sky with dust, (2) cloudy with dust, (3) clear-sky without dust, (4) cloudy without dust. This is summarised in Figure 1.

In a subsequent step we calculate the model error for each cell for both BT and SIS, as follows:

$$BT_{\text{difference}} = BT_{\text{model}} - BT_{\text{satellite}},$$

$$SIS_{\text{ratio}} = \frac{SIS_{\text{model}}}{SIS_{\text{satellite}}}.$$

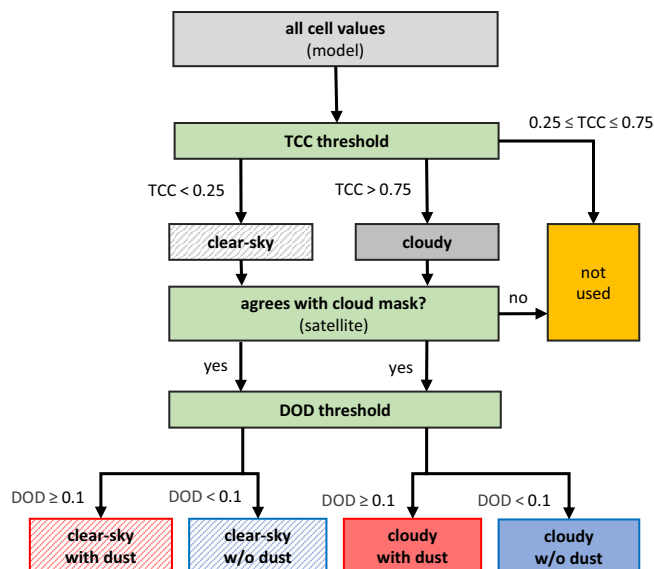


FIGURE 1 Classification of four different cases depending on simulated total cloud cover (TCC), satellite cloud mask, and simulated dust optical depth (DOD). The criteria are applied to each grid-cell value from all events in the event catalogue. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

To account for temporal autocorrelations between selected dust days, we summarise consecutive dust days into the same event. Next we compute the area-weighted median model error relative to satellite data from all cells within the respective case and per dust event, in order to account for spatial correlations. We then use the distribution of median model errors per event and case for statistical testing and error quantification, and for intercomparison between the four cases.

3 | RESULTS

We first present the results of a case study for a particularly pronounced event with Saharan dust over Europe. We use the pre-operational ICON-ART model to investigate how a model with prognostic aerosol reproduces observed conditions. We compare this with the ICON model without ART and the ECMWF IFS model, which both deploy aerosol climatologies. The IFS model in particular is widely used for operational forecasting, and simulations of dust concentrations are available from its coupled version IFS-AER. In a second step, we make use of the available IFS and IFS-AER analyses and short-range forecasts and perform a statistical analysis for a large set of dust events to investigate whether model errors are systematic and quantify deviations from observations.

3.1 | Dust event case study

3.1.1 | Synoptic situation

During February and March 2021, several events occurred during which high loads of Saharan dust were transported towards Western and Central Europe. In a recent publication, Francis et al. (2022) showed that two dust events in February 2021 coincided with ARs. In the following we focus on a later episode in early March 2021 and base our synoptic analysis on data from CAMS (DOD) and ERA5 (all other variables).

At the end of February and beginning of March 2021, large parts of Europe were under the influence of a pronounced anticyclone with its high-pressure centre over Germany (not shown). This system formed an omega-like situation over Central Europe and caused large-scale subsidence with stable clear-sky conditions. West of the anticyclone, a trough spans from a low-pressure region over the Bay of Biscay via the western Iberian peninsula into Morocco. These conditions can favour the formation of North African cyclones (e.g., Knippertz & Fink, 2006). At the same time, the SEVIRI desert dust red-green-blue (RGB) composite (Lensky & Rosenfeld, 2008) shows pink colours in the south of Algeria (not shown), a known dust source region during spring (Ginoux et al., 2012). Elevated dust appears in pink colours in this RGB product. Pink colour patterns, presumably dust, are advected northwards and reach the Mediterranean coast on the morning of March 1. Multiple stations in the north of Algeria report dust on March 1, 2021. Additionally, stations in Er-Rachidia (Morocco) record gusts of up to $59 \text{ km}\cdot\text{hr}^{-1}$ and stations in Tozeuz/Nefta (Algeria) gusts of up to $65 \text{ km}\cdot\text{hr}^{-1}$, a favourable condition for further dust emission. The southerly flow over the western Mediterranean advects subtropical air masses and Saharan dust northward. On March 2, the high-pressure system over Central Europe starts to weaken. Over the western Mediterranean, southerly flow persists and advects Saharan dust further, and a high-level cloud shield forms. On March 3, Western and Northern Europe are under increasing cyclonic influence (Figure 2a). A cut-off cyclone is evident over North Africa. Ahead of an approaching short-wave trough over Western Europe, satellite BT from March 3 at 1200 UTC shows a high-level cloud extending from the Iberian peninsula over central France towards Central Europe (Figure 2b). Enhanced southwesterly to westerly winds in the mid and upper troposphere favour the advection of Saharan dust in these layers. Compared with the dust outbreaks during February 2021 (Francis et al., 2022), the transport of integrated water vapour (IVT) is much lower at the event examined here (Figure 2a). This is

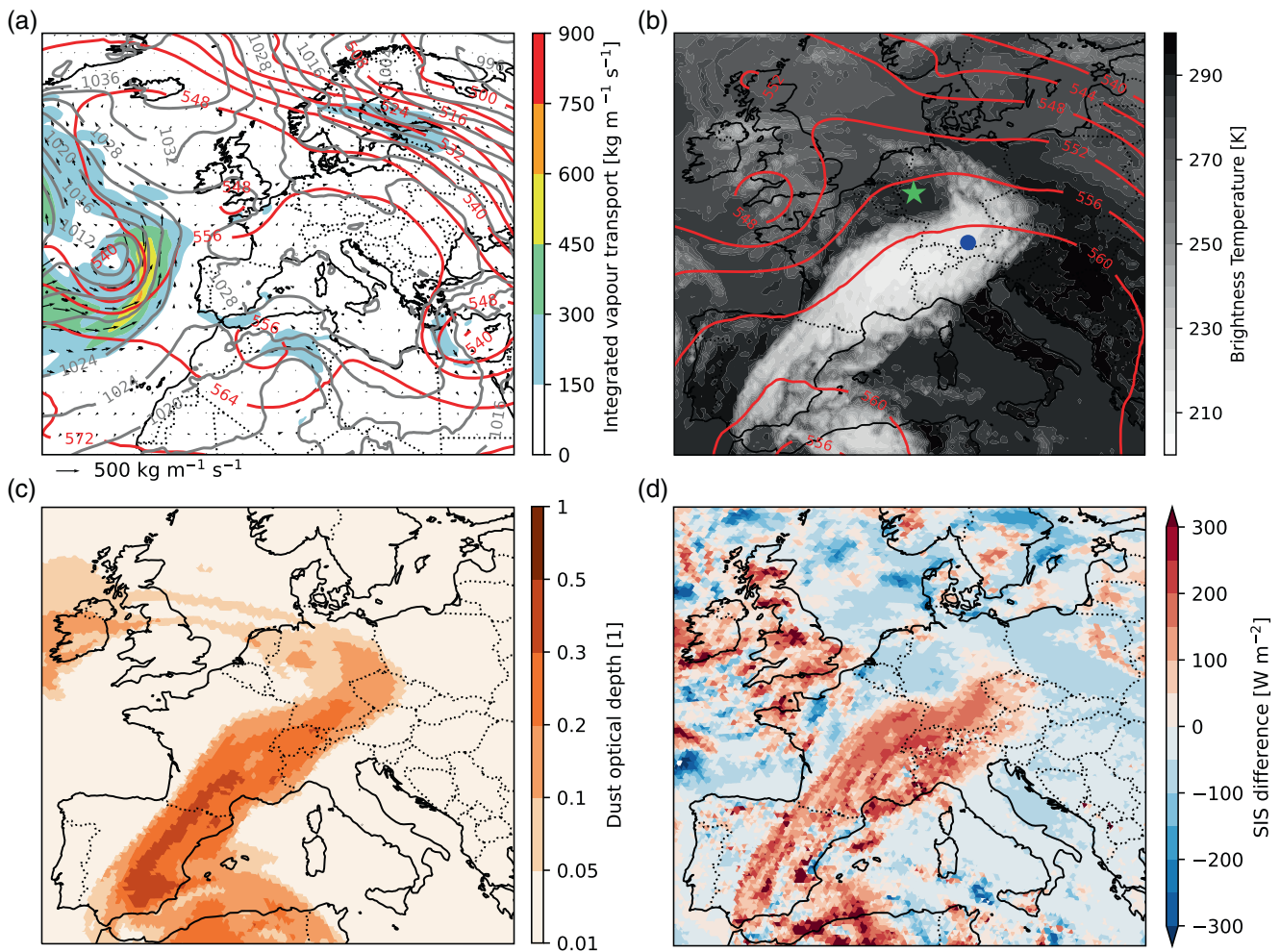


FIGURE 2 Synoptic situation, modelled dust, and model errors over Europe on March 3, 2021, 12 UTC. (a) Integrated vapour transport (IVT, computed in analogy to Francis et al., 2022), 500-hPa geopotential height (contours in geopotential decameters (gpdm) with a 8-gpdm contour interval), and mean sea-level pressure (contours in hPa with a 4-hPa contour interval) over North Africa and Europe. All data from ERA5. (b) Brightness temperature (shading in K) and 500-hPa geopotential height (contours in gpdm with a 4-gpdm contour interval). The DWD station Köln/Bonn is marked with a star and Hohenpeißenberg is marked with a filled circle. Brightness-temperature data are retrieved from Gridsat satellite composites with 0.07° resolution. Geopotential height from ERA5. (c) Dust forecast from ICON-ART. Colour shading shows DOD above 5000 m. Data are retrieved from a pre-operational +12 hr forecast with a grid spacing of 20 km. (d) Absolute difference in surface shortwave radiation between ICON-ART model and MSG satellite, mean from 1200–1500 UTC. Model data are retrieved from a pre-operational +15 hr forecast with a grid spacing of 20 km and from CM SAF operational products using the SARA-H-2 methods, respectively. [Colour figure can be viewed at wileyonlinelibrary.com]

corroborated by the AR database (Guan, 2022) developed by Guan and Waliser (2015) and based on Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) reanalysis data, which does not identify ARs in the western Mediterranean during the period examined. We evaluate the capability of current models to reproduce the observed conditions on March 3, 2021 in a case study. For this, we analyse firstly a simulation with the pre-operational ICON-ART model with prognostic dust and secondly the operational IFS with a dust climatology.

3.1.2 | Pre-operational forecast with prognostic dust: ICON-ART

At the time of this study, ICON-ART is used for pre-operational dust forecast by DWD. We use this forecast for evaluating the extent of the dust plume. On March 3, 2021, 1200 UTC, the forecast shows the dust plume extending into altitudes above 5000 m and spanning from North Africa and the Iberian peninsula via France to Central Europe (Figure 2c). This implies overlap of the dust plume with the observed high-level cloud (Figure 2b) and

suggests that dust might play a role in the formation of this cloud. Further examination indicates the model simulating vertical structures of dust (Appendix Figures A1,A2) and moisture (Appendix Figure A4) with reasonable skill.

Cloud cover and cloud properties play an important role for the radiative budget in the atmosphere (Boucher et al., 2013). Consequentially, model errors in cloudiness manifest as errors in radiative quantities. For assessing this, we further compare SIS from ICON-ART with SIS derived from MSG. Figure 2d shows the difference in SIS between the ICON-ART model and the satellite retrieval. The model shows slightly lower values than the satellite for clear-sky regions, with differences of up to $20 \text{ W}\cdot\text{m}^{-2}$ over Italy and the central Mediterranean and up to $65 \text{ W}\cdot\text{m}^{-2}$ over Poland. Taking into account that the SIS satellite retrieval shows a positive bias during conditions with elevated aerosol concentrations over the climatology (see Section 2.2.2), we consider this a reasonable agreement. In areas of low or broken clouds, varying patterns of disagreement between model and satellite can be seen, which can mostly be related to temporal or spatial shifts in cloudiness between model and satellite. In the area of high DOD and high-level clouds, namely the area between the Iberian peninsula and Central Europe, model and satellite show strong and spatially consistent deviations in SIS. Absolute differences exceed $100 \text{ W}\cdot\text{m}^{-2}$ in a wide swath, and exceed $300 \text{ W}\cdot\text{m}^{-2}$ at individual locations. This translates into relative model errors in SIS of up to 50%. Spatial consistency of the deviations with the high-level cloud layer and dust plume suggests the model has problems reproducing the observed high cloudiness.

ICON-ART reproduces cloud cover in lower layers and regions with moderate DOD. However, also in ICON-ART the cirrus cloud shield in the area of elevated DOD is absent (not shown). Recalling that our ICON-ART version does not include aerosol–cloud interaction, we take this as a strong indication that dust played an important role for the formation of the high-level cloud and suggest it being a dusty cirrus cloud (Seifert et al., 2023).

For evaluating the plausibility of SIS values, we compare SIS data from model and satellite with ground-based measurements. We further include model data from ICON with different aerosol implementations, namely a setup with aerosol climatology (ICON) and the pre-operational setup with prognostic aerosol (ICON-ART), for examining whether the improved representation of direct aerosol effects improves forecast of SIS for the examined March 3 event. We focus on two selected stations: Köln/Bonn, which is located in the west of Germany and experiences clear-sky conditions on March 3, and Hohenpeißenberg in the south of Germany, which is located directly under the cloud shield on March 3 (for locations see Figure 2b). We evaluate SIS for both stations

and the different cloudiness and dust conditions from March 2–4 in Figure 3. On March 2, Köln/Bonn and Hohenpeißenberg are both located under clear-sky conditions with low DOD. SIS from station measurements and satellite retrieval shows good agreement. Both ICON and ICON-ART show very similar values with and without aerosol effects, slightly underestimating SIS compared with ground- and satellite-based retrievals at both stations.

On March 3, ICON-ART shows increasingly dusty conditions for both stations, while the satellite image shows Köln/Bonn under clear-sky conditions and Hohenpeißenberg under the dense cirrus cloud shield. SIS model data from ICON shows similar values to those on March 2, indicating that ICON does not reproduce the reduced SIS during dusty conditions. In contrast, ICON-ART clearly shows a reduction in SIS by $50 \text{ W}\cdot\text{m}^{-2}$ for Köln/Bonn and by $75 \text{ W}\cdot\text{m}^{-2}$ for Hohenpeißenberg on March 3 at 1200 UTC compared with March 2 and compared with ICON. For Köln/Bonn, the ICON-ART value matches the

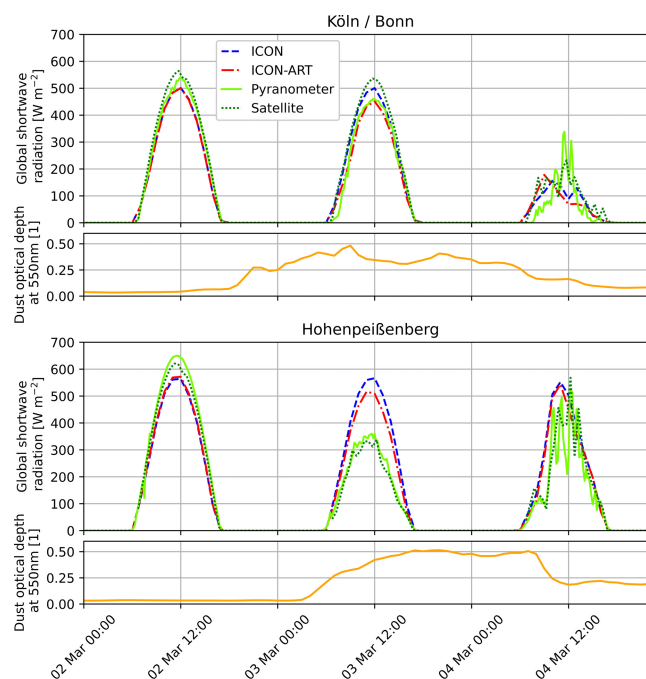


FIGURE 3 Intercomparison of SIS from model forecasts, satellite, and station measurements and dust forecast from March 1–3, 2021 at Köln-Bonn and Hohenpeißenberg. ICON and ICON-ART values are retrieved with one-hour spacing from operational and pre-operational forecasts, respectively, with daily re-initialisation at 0000 UTC. Station data with 10-min spacing are from station-based measurements from DWD. Satellite data with 30-min spacing are retrieved from CM SAF operational products using the SARA-H-2 methods. The dust forecast is from ICON-ART. For Köln-Bonn the simulated station in ICON is located about 17.5 km SE from the DWD station. [Colour figure can be viewed at wileyonlinelibrary.com]

station measurement very well. Satellite SIS shows slightly reduced values compared with the previous day but also overestimates radiation compared with the pyranometer measurement. For Hohenpeißenberg, values from both model setups are still much higher than the station measurement and satellite retrieval. This suggests that the implementation of prognostic aerosol leads to an improvement of the direct radiative effect in ICON-ART, especially under clear-sky conditions. However, the current level of model complexity is not sufficient for reproducing SIS adequately during dusty conditions with clouds. On March 4, both stations experience changing conditions of cloudiness and reduced DOD compared with March 3. SIS from satellite retrievals agrees well with the station measurements. Models show no clear difference between ICON and ICON-ART but smoothing due to reduced time resolution in comparison with measurements. For Köln/Bonn the models roughly reproduce daily SIS, whereas for Hohenpeißenberg, where DOD remains higher, both models still overestimate SIS.

From intercomparison of SIS values from station measurements, satellite, and ICON with and without prognostic dust, several characteristic features can be summarised. Firstly, satellite retrievals generally match well with ground-based measurements. Discrepancies during dusty clear-sky conditions might be related to the use of an aerosol climatology in the satellite retrieval. Secondly, the model forecast during dusty clear-sky conditions improves with the inclusion of prognostic aerosol accounting for direct radiative effects from dust. Thirdly, the forecast during dusty cloudy conditions is not captured well by either model, which suggests that the implementation of prognostic aerosol with direct radiative effects is not sufficient for the model to capture cloudiness and subsequently SIS during these conditions.

3.1.3 | Operational model without prognostic dust: ECMWF IFS

Since the pre-operational ICON-ART model shows problems with reproducing high clouds during the dust event studied here, we examine the ECMWF IFS model to evaluate how the dust event on March 3, 2021 is reproduced by a model that implements dust via prescribed climatologies and is currently used for operational weather forecasting. To keep discrepancy from the model analysis low, we base our analysis on IFS products at initialisation time (BT) or from the first timestep when temporally accumulated variables are available (SIS, +3 hr forecast). We only use data from the IFS *hresrun*.

We firstly analyse the simulated BT from IFS in Figure 4, a direct model output that can be compared with the GridSat BT (derived from *MSG*) in Figure 2b. Comparing cloud patterns over North Africa, the Bay of Biscay, or the Baltic, there is a good agreement in cloud structures between model and satellite on March 3. Focusing on the region with the observed high-level cloud, model and satellite differ. The discrepancies between model and satellite are most pronounced in the area of high DOD according to the ICON-ART dust forecast and reach values up to 75 K, suggesting the high cloud layer is highly underestimated or missing in the model. By analogy with the previous section, we analyse SIS from the IFS model relative to data derived from the *Meteosat* satellites via SARAH-2 methods (Figure 4). Additionally we analyse output from model runs with increased forecast lead times (not shown). For data in clear-sky regions, there is good agreement between model and satellite. In areas of low or broken clouds, there are varying patterns of discrepancy between model and satellite, which are mostly related to temporal or spatial shifts of clouds between model and satellite. These

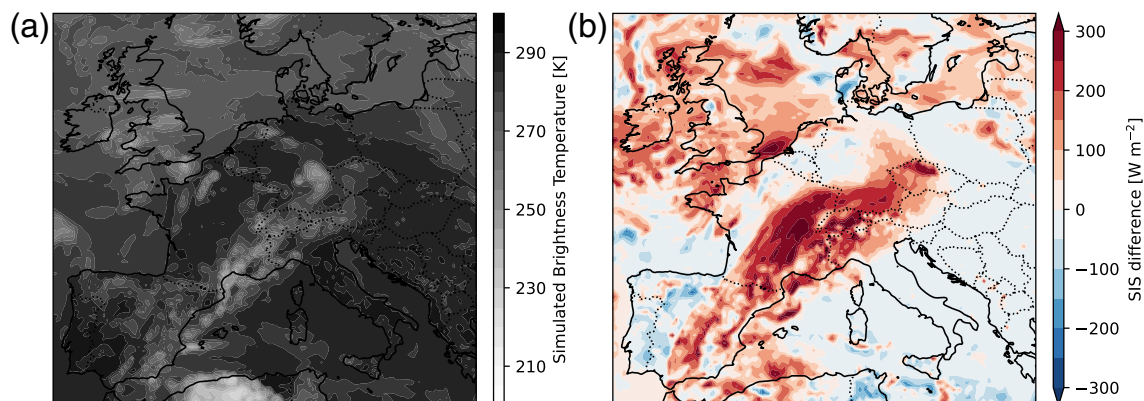


FIGURE 4 Model simulation with ECMWF IFS for March 3, 2021, 1200 UTC. Data are retrieved from the *hres* run with a grid spacing of 9 km. (a) Simulated brightness temperature. Data are retrieved from a forecast at initialisation time. (b) Absolute difference in surface shortwave radiation between ECMWF IFS model and *MSG* satellite, mean from 1200–1500 UTC. Data are retrieved from a +3 hr forecast and from CM SAF operational products using the SARAH-2 methods, respectively. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

differences increase with increasing forecast lead times, resulting in the noisy pattern. In the area of the dusty cirrus cloud, the model does not show strong reductions in SIS. In contrast, SIS derived from satellite shows largely reduced values in this area. Absolute differences between model and satellite exceed $100 \text{ W}\cdot\text{m}^{-2}$ in a spatially consistent area between the Iberian peninsula and Central Europe and exceed $300 \text{ W}\cdot\text{m}^{-2}$ at individual locations. This translates into relative model errors of up to 50% in the area of the (missing) cirrus cloud, and agrees well with the error pattern observed for ICON-ART. This prominent error pattern under the cirrus cloud is consistent between the model analysis and forecasts with increased lead times.

3.1.4 | Case study summary

During the beginning of March 2021, Saharan dust was advected towards Central Europe. Other than during dust events in February 2021, this did not coincide with an AR. Satellite data show a high and spatially consistent high-level cloud over Western and Central Europe on March 3, which overlaps with the simulated dust plume at altitudes above 5000 m. The prominent high-level cloud on March 3 is consistently underestimated or missing in a pre-operational model with prognostic calculation of aerosols (ICON-ART) as well as in operational models using aerosol climatologies (ICON, ECMWF IFS). The pre-operational ICON-ART model with prognostic aerosol and direct effects reproduces the general weather conditions reasonably on March 3, but does not properly reproduce the high-level cloud in the area of the dust plume. The missing dusty cirrus cloud leads to model errors in SIS of more than $100 \text{ W}\cdot\text{m}^{-2}$ in a wide swath, and peaks above $300 \text{ W}\cdot\text{m}^{-2}$ (about 17% and 50% of total SIS). Intercomparison with ICON shows improvement of the direct effect on SIS with the inclusion of prognostic dust. This, however, still underestimates indirect effects during cloudy conditions. As the model shows nearly perfect spatial overlap of the dust plume with model errors in cloudiness, mineral dust is efficient, as INP and indirect effects are not yet implemented into ICON-ART: we conclude that dust is very likely the cause of model errors for the event examined here. Similar errors in clouds and SIS are recorded for the operational IFS model. Errors exist in both model analysis and forecast with various lead times, suggesting a missing implementation of relevant processes in the model leading to this error.

Comparison of SIS satellite and station data shows good agreement during continuously cloudy conditions with dust, where the optical depth from clouds dominates extinction in the atmosphere, as well as during clear-sky

conditions without dust. The SIS retrieval from satellite measurements, however, does not account for increased direct radiative effects during elevated dust concentrations. This becomes particularly obvious during clear-sky conditions with dust when compared with station measurements. SIS from satellite retrievals should therefore be handled with caution when used as a reference. ICON shows improvement of the SIS forecast under dusty clear-sky conditions with the inclusion of the ART module. ART adds prognostic aerosol, but only includes the direct radiative effect of dust. Under the cloudy conditions with dust studied here, the direct aerosol effect is not sufficient for reproducing SIS from measurements. Models do not reproduce the extensive high-level cloud in the area of the dust plume. As the model errors in cloudiness align spatially with dust above 5000 m, and comparison with station data shows improvements with the inclusion of direct effects, we conclude that dust effects are likely the source that causes these errors. Recently, Seifert et al. (2023) suggested a novel subgrid parametrisation for dusty cirrus clouds and showed that only with this parametrisation is ICON-ART able to simulate the formation of these clouds. We conclude that, in particular, the missing implementation of indirect aerosol effects is likely the cause of the models not reproducing the high-level cloud in regions of high dust concentrations.

3.2 | Statistical analysis

3.2.1 | Event catalogue

As shown in the previous section, omitting prognostic dust and dust–radiation–cloud interactions in operational weather forecast models can lead to large errors in radiation and cloudiness forecasts during Saharan dust outbreaks. To verify if such errors occur systematically during dust events, we investigate the dust events in recent years. To extract the “dusty days”, we apply the selection criteria explained in Section 2.4 for the period from January 2018–March 2022 for each day at 0000 and 1200 UTC. This yields the selection of 49 individual days and translates into an average of 11.5 dust days over Central Europe per year. Clustering of consecutive dust days into events results in 24 dust events. Manual screening confirms mineral dust as the dominant aerosol for all events. All dates selected for the event catalogue are summarised in Table 1. Most dust days occur during spring (MAM, 24 days), followed by summer (JJA, 12 days). During autumn (SON, 4 days) and winter (DJF, 8 days), dust events over Central Europe are less frequent. This seasonality agrees with other studies (e.g., Moulin et al., 1998; Israelevich et al., 2012; Merdji et al., 2023), which that show Saharan dust transport to

TABLE 1 Selected dates for the dust event catalogue from application of the selection criteria to the area of Central Europe.

Selected dust days				
2018-04-07	2019-04-22	2020-01-23	2021-02-06	2022-03-15
2018-04-08*	2019-04-23*	2020-05-13	2021-02-07*	2022-03-16*
2018-04-09*	2019-04-24*	2020-11-07	2021-02-22	2022-03-17*
2018-04-23	2019-06-10		2021-02-23*	2022-03-18*
2018-05-26	2019-06-11*		2021-02-24*	2022-03-29
2018-05-27*	2019-06-14		2021-03-02	2022-03-30*
2018-05-28*	2019-06-15*		2021-03-03*	
2018-05-31	2019-06-25		2021-03-04*	
2018-06-11	2019-06-26*		2021-04-01	
2018-08-07	2019-06-27*		2021-04-02*	
	2019-07-24		2021-06-18	
	2019-07-25*		2021-06-19*	
	2019-10-13		2021-06-20*	
	2019-10-14*			
	2019-10-23			
	2019-12-17			
	2019-12-18*			

Note: Consecutive dust days, which were summarised into the same event as the previous date, are marked with an asterisk.

the Mediterranean region and Europe is most active during spring and summer.

3.2.2 | Synoptic situation

For assessing the mean synoptic situation during days with Saharan dust over Europe, we compute composites from all dust days in the event catalogue. Figure 5 shows the composite of 500-hPa geopotential and total DOD from CAMS. This shows a pronounced trough over the Iberian peninsula, which extends over the Atlas mountains and Algerian desert regions, while the Central Mediterranean is under anticyclonic influence. The synoptic composite resembles the mean synoptic conditions observed for dust transportation days to Central Europe by Barkan et al. (2005). The pronounced trough spanning into the Atlas mountains is characteristic for conditions that favour the formation of North African cyclones, which are efficient for dust transport to the Mediterranean region (Flaounas et al., 2015). With the tendency for elevated wind speeds and quasigeostrophic forcing ahead of the trough axis, dust emission and lifting into higher atmospheric layers can be assumed. For dust days in our catalogue, dust mainly reaches Europe with the general flow via the Western Mediterranean.

3.2.3 | Statistical results

We analyse BT difference and SIS ratio for the four different case classes (Section 2.5) and the operational forecasts of ECMWF. We only use data from the IFS *hres* run. Figure 6 shows the median of these relative errors over all dust events in the evaluated Central Europe area (see definition in Section 2.5 and Figure 5), but for different DOD thresholds classifying with/without dust. Note that a DOD threshold of 0.1 is used for initial event detection. Further note that we only use cells where model and satellite agree on cloudiness (15.2% of cells do not agree and were discarded for the analysis), and that area-median errors were computed for each case class per event prior to computing the median over all events. Generally, for higher DOD thresholds model errors increase. For very high threshold values, only a few or no dust events contain cells that exceed the threshold, hence the number of events with dusty cases is low or not available (clear-sky with dust). This results in abrupt variations in values for cases with dust with increasing DOD threshold. For both BT and SIS, the case “cloudy with dust” stands out and shows a strong increase of the model error with increasing DOD threshold. For the other cases, the increase of model errors with increasing DOD threshold is small. Thresholds for TCC are chosen as outlined in Section 2.5 (<25% and >75%

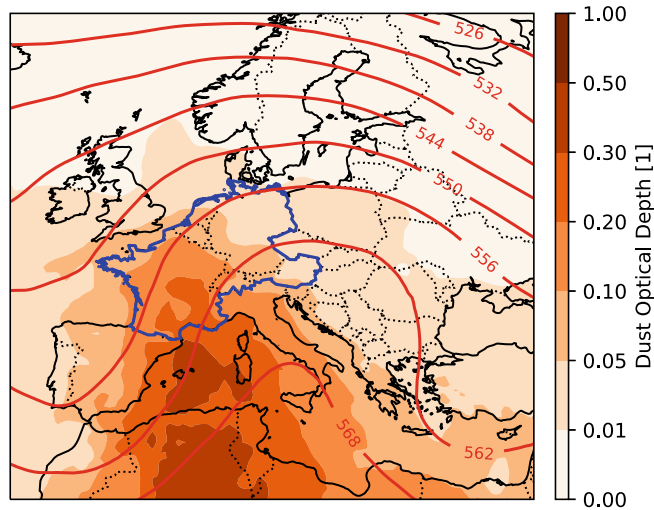


FIGURE 5 Composite of the large-scale synoptic situation from all dust days in the event catalogue, each day at 1200 UTC. We show the 500-hPa geopotential field (labeled contours in geopotential decameters (gpdm) with a 6-gpdm contour interval), total DOD (colour shading), and the area of evaluation in Central Europe (unlabeled contour). Fields are retrieved from the CAMS near-real-time forecast at initialisation time. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

for classification as clear-sky and cloudy, respectively). Stricter thresholds do not show a significant difference for clear-sky conditions and BT, but increase the SIS errors during cloudy conditions with dust (not shown), while reducing the total sample size. Also independent of the DOD threshold, cloudy grid cells with dust feature a higher median error in BT and SIS compared with their counterpart without dust. This strongly points to dust playing an important role in causing large errors in cloud parameters, and subsequently surface radiation.

For further statistical analysis we chose the threshold DOD of 0.1 at individual grid points. This value is exceeded during each dust day, so that we include values from each dust event. Table 2 summarises model errors relative to satellite measurements for the four cases at this threshold. Comparing model errors between the four cases, prominent characteristics stand out.

Firstly, the cloudy cases show a wide distribution of model errors (Figure 6, solid boxes), compared with a narrow distribution for clear-sky cases (Figure 6, dashed boxes). As this persists in both cases with dust (red) and cases without dust (blue), we relate the broad error distribution in the cloudy cases to non-dust-related variations in cloudiness and cloud properties between model and satellite. The narrow distribution of clear-sky errors suggests consistency of model values relative to satellite retrievals.

Secondly, the case cloudy with dust shows the largest median model error for both BT (7.0 K) and SIS (6.2%).

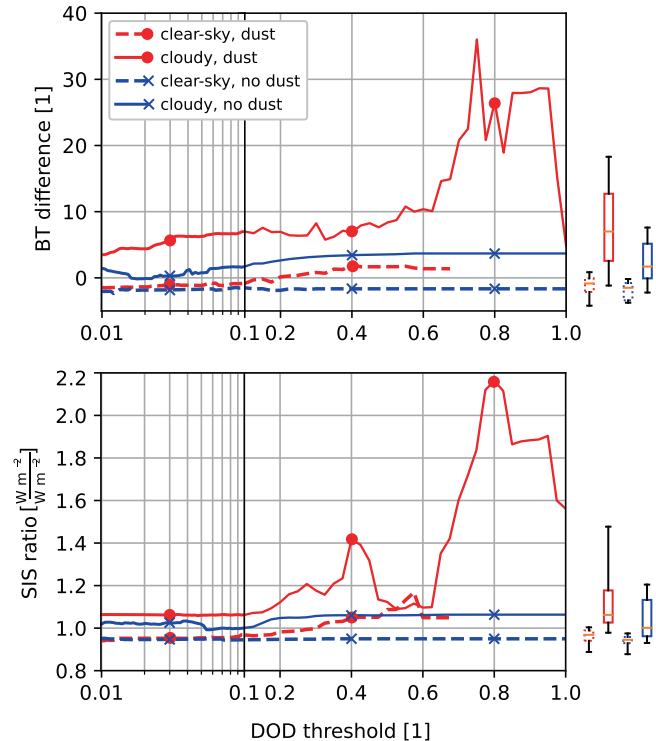


FIGURE 6 Sensitivity of BT difference (model – satellite) and SIS ratio (model/satellite) to the DOD threshold value for classification of cells with/without dust. Cases with dust in red with square markers, cases without dust in blue with cross markers. Dashed lines indicate clear-sky conditions, solid lines indicate cloudy conditions. Line plots show the median value, box plots show quantiles (0.10, 0.25, 0.5, 0.75, 0.90) for a DOD threshold of 0.1. Data from all events in the event catalogue at 1200 UTC each day are shown, alongside model data from the ECMWF IFS (*hres*) at the time of initialisation at 1200 UTC (BT) or the first timestep when temporally accumulated variables are available (SIS, +3 hr forecast), and satellite data for BT and SIS from GridSat and MSG (CM SAF), respectively. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

Median model errors are smaller for the other cases and of similar magnitude between cases. Setting median errors for cases with dust in relation to cases without dust, information about the effect of dust can be drawn. We test the significance of the differences in the medians between the case with dust and the corresponding case without dust via bootstrapping the difference of medians. We draw samples from the distributions of area medians from all events for both cases, with a sample length equal to the sample size of the distribution. Note that all distributions contain values for all 24 events, and therefore are of equal length. We calculate the difference in sample medians for each bootstrap replicate. We use a total number of 1000 bootstrap replicates for calculating the bootstrapping distribution. For all tests, we select a 95% confidence interval around the bootstrapping distribution median and call the differences significant where this interval does not extend over zero.

TABLE 2 Deviations of the IFS model values for BT and SIS from satellite retrievals for a DOD threshold of 0.1 for all events from the catalogue.

Case		Deviation from satellite			Sample size
		Median	Mean	Std	Rel. / abs.
Clear-sky with dust	BT	-0.8 K	-1.5 K	2.8 K	6.4%/4374
	SIS	-3.2%	-2.5%*	9.8%	
Cloudy with dust	BT	7.0 K*	8.54 K*	9.1 K	29.5%/20,147
	SIS	6.2%	14.1%*	19.4%	
Clear-sky w/o dust	BT	-1.5 K	-2.0 K	2.2 K	19.8%/13,481
	SIS	-5.5%	-6.1%*	3.7%	
Cloudy w/o dust	BT	1.7 K*	2.4 K*	4.1 K	44.3%/30,250
	SIS	0.1%	5.9%*	14.8%	

Note: BT difference in absolute values, SIS as deviation from the ratio 100% (SIS ratio-1). Statistics are computed from the distributions of area-median errors per case and event. Sample size on cell level over all dust days before collecting in events and area-averaging per event. Cells where model and satellite cloudiness do not agree are excluded. Cases with significant difference between conditions with and without dust at the 95% confidence level are marked with an asterisk. Abbreviations: BT, brightness temperature; SIS, surface incoming shortwave radiation.

For clear-sky cases, the differences in BT median errors between the cases with dust and without dust reach 0.7 K. For SIS, the differences reach 2.3%. Both are not significant. For cloudy cases, the differences in BT median errors between the cases with dust and without dust reach about 5.3 K and are significant. For SIS the differences reach 5.1% and are not significant. We apply the same procedure for testing the differences in the mean errors between the case with dust and the corresponding case without. This confirms significance for BT under cloudy conditions with dust. Additionally, the mean errors in SIS under clear-sky and cloudy conditions with dust are significantly different from the corresponding cases without dust. Median differences do not agree with this, therefore we conclude no robust significance for differences in SIS. These findings suggest that the indirect effect plays the dominant role for causing forecast errors in properties such as cloud-top height or cloud optical depth during dust events. While SIS errors during cloudy conditions are positive, a significant effect from dust cannot be confirmed when comparing the cases with/without dust. This is partly consistent with the findings in Section 3.1, which show that the indirect effect dominates the absolute amplitude of model errors in a model without prognostic aerosol when compared with station measurements. The clear indications for the direct aerosol effect causing forecast errors during clear-sky conditions when compared with station measurements cannot be confirmed by our multi-event analysis. However, it must be noted that the SIS satellite product likely shows a positive bias during clear-sky conditions with highly elevated aerosol optical depth, due to the deployment of a prescribed aerosol climatology in the retrieval method (Pfeifroth & Trentmann, 2018, see Section 2.2.2).

Thirdly, errors for clear-sky cases are negative. This is consistent between cases with and without dust for BT and SIS, and might be attributed to a general clear-sky bias in model or satellite data.

Our analysis only assesses cases where model and satellite agree on cloudiness, and excludes cases where model and satellite do not agree whether a cell is “clear-sky” or “cloudy”. Since Section 3.1 has shown that dust can lead to conditions where models do not reproduce observed clouds, we investigate whether the IFS model reproduces total cloud fraction compared with the satellite cloud mask. We compute cloud fraction per event directly from the TCC model variable, and the cloud fraction in the satellite cloud mask from the fraction of clear-sky pixels from all pixels. For each event we then compute the mean over cells or pixels above/below the DOD threshold. Figure 7 shows the mean cloud fraction over all events for increasing DOD thresholds. Model and satellite cloud fraction agree well for DOD<0.1, with maximum deviations of model values from the satellite cloud mask of up to 2.3%. With increasing DOD, cloud fraction from the satellite cloud mask increases. This increase is underestimated by the model, which shows up to 19.8% lower cloud fraction than the satellite for a DOD greater than or equal to 0.7. This matches with the findings in Section 3.1, which show that models without prognostic dust might not reproduce cloudiness during conditions with dust, again pointing to the indirect effect from dust playing an important role for causing model errors during such conditions. We conclude that the increase in cloud fraction with increasing DOD is likely related to indirect effects from dust. We hypothesise that this effect might additionally be enhanced by increased moisture transport during AR events, which have been shown to coincide frequently with dust episodes

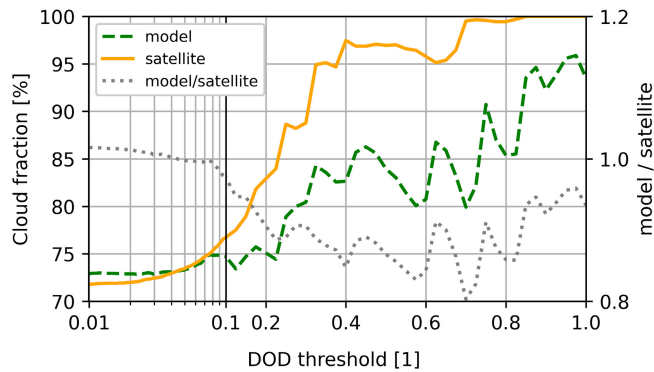


FIGURE 7 Cloud fraction from model, satellite, and ratio of model to satellite cloud fraction for different DOD threshold values for classification of cells with/without dust. Data from all events in the event catalogue at 1200 UTC each day. Model data from the ECMWF IFS (*hres*) at the time of initialisation at 1200 UTC (BT). Satellite data from the MSG CM SAF cloud mask. [Colour figure can be viewed at wileyonlinelibrary.com]

over Europe (Francis et al., 2022). However, the underestimation of cloud fraction in the model can have various sources beyond the availability of CCNs, INPs, or water vapour, and a bias in the cloud-mask classification algorithms cannot be ruled out. A systematic assessment of this model error is beyond the scope of this study but might be investigated in future work. Concluding, the analysis of 49 dust days over Central Europe shows strong sensitivity of model errors to dusty and cloudy conditions. Comparing absolute model errors with satellite ones, and relative errors of the dust case with those of the no dust case, we conclude that the misrepresentation of cloud properties plays a dominant role for causing model errors during dust events. We further find that the model underestimates increased cloud fraction during dust events. With Saharan dust transport towards Europe occurring multiple times per year, our statistically robust quantification of model errors emphasises the need for the inclusion of the dust indirect effect into NWP models in order to improve the representation of cloud properties during such events.

4 | CONCLUSION

Current operational weather forecasting models, such as the ICON model at DWD and IFS model at ECMWF, deploy prescribed aerosol climatologies for implementing dust effects on clouds and radiation. In this study we show that during Saharan dust outbreaks, when concentrations of mineral dust are greatly above climatological means, this approach leads to a significant underestimation of cloudiness and an overestimation of global radiation.

We first examined a dust event on March 3, 2021 and showed that both IFS and the ICON model are incapable

of reproducing the observed cirrus clouds co-located with the dust plume. Even coupled with the aerosol module ART, the pre-operational ICON-ART does not reproduce observed cloudiness, but improves the direct aerosol effect in clear-sky regions. Consequentially, deviations in SIS relative to satellite retrievals reach up to 50% of total SIS for both IFS and ICON-ART in cloudy conditions. Nevertheless, both models reproduce the larger synoptic situation with skill, and for ICON-ART the structure of the dust plume matches satellite retrievals. Thus, the implementation of direct effects improves the SIS forecast during dusty clear-sky conditions in this case but is not sufficient to reproduce measured values during dusty and cloudy conditions. As the deviations from measurement data exist for both model analysis and forecast, and as models reproduce the general weather situation in most areas without dust with skill, we conclude that the lacking implementation of dust indirect effects is the most likely cause for this model error.

We further show via statistical analysis of all dust events over Central Europe from 2018 to spring 2022 that this underestimation does not occur only during a particular dust event but is statistically robust throughout all 49 dust days. Our statistical analysis shows that the model deviations from satellite retrievals occur most prominently under dusty conditions with clouds. Absolute errors are highest for cloudy conditions with dust, with median absolute errors in BT of 7.0 K and median absolute errors in SIS of 6.2% for dust cases with DOD equal to or greater than 0.1. Differences between the cases with and without dust are highest for cloudy conditions, with differences in median errors in BT of 5.3 K, which are significant at a 95% confidence level. This indicates a systematic effect of dust on cloud properties such as cloud depth, cloud height, and cloud optical thickness. For SIS, differences in median errors reach 6.1% but are not significant between the cases with/without dust. For clear-sky cases, absolute errors in BT and SIS are much smaller than for cloudy conditions and are not significant. However, as the satellite retrieval of SIS is known to be particularly biased under clear-sky conditions with aerosol concentrations greatly above the climatological mean, quantification of clear-sky SIS against satellite data should only be done with caution. Our case study shows that a bias in the satellite retrievals of SIS during clear-sky conditions with dust could obscure model errors during such situations. We hence conclude that we cannot prove nor rule out a significant error from direct dust effects. We suggest testing the usage of near-real-time model fields of aerosol for future versions of SIS retrievals, and a comparison with retrievals deploying an aerosol climatology. We further show that cloud fraction increases with increasing DOD, a behaviour that the IFS model does not reproduce sufficiently. We suggest future research

to assess the cause of the underestimated cloud fraction during dust events in detail, in order to disentangle the effect of dust on cloud formation from other processes: for example, from potentially increased moisture transport when dust days over Europe coincide with AR events (Francis et al., 2022).

Our study shows that operational weather forecasting models currently still lack the skill to reproduce cloudiness during dust events adequately. Analysis of dust events from several years shows that these underestimations occur frequently. With 49 extremely dusty days occurring over the area examined during the study period, this highlights the importance of implementing prognostic dust in numerical weather prediction models. Despite pre-operational models simulating direct effects of mineral dust, which improves the forecast during clear-sky conditions with elevated concentrations of dust, indirect effects must be considered in order to account for dust serving as INP and altering cirrus cloud formation. A first promising attempt in this direction has recently been proposed by Seifert et al. (2023), who show improved dusty cirrus representation, including the case studied here, with a novel subgrid parametrisation. The inclusion of the indirect aerosol effect is a topic of ongoing development in the ICON-ART model system and can contribute to an improved forecast of cloudiness and, closely connected, surface radiation.

AUTHOR CONTRIBUTIONS

Kilian Hermes: data curation; formal analysis; investigation; methodology; software; validation; visualization; writing – original draft; writing – review and editing. **Julian Quinting:** data curation; resources; supervision; writing – review and editing. **Christian M. Grams:** conceptualization; funding acquisition; supervision; writing – review and editing. **Corinna Hoose:** funding acquisition; supervision; writing – review and editing. **Gholam Ali Hoshyaripour:** conceptualization; funding acquisition; project administration; resources; supervision; writing – review and editing.

ACKNOWLEDGEMENTS

The contribution of JQ and CMG is funded by the Helmholtz Association as part of the Young Investigator Group “Sub-seasonal Predictability: Understanding the Role of Diabatic Outflow” (SPREADOUT, grant VH-NG-1243). CH acknowledges funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 714062 (ERC Starting Grant “C2Phase”). We acknowledge ECMWF and Deutscher Wetterdienst (DWD) for granting access to operational forecast data and ERA5 reanalysis data. We

further acknowledge DWD’s Climate Data Center (CDC) for providing station measurements. We acknowledge the Copernicus Atmosphere Monitoring Service for providing atmospheric composition data. We acknowledge NOAA’s National Climatic Data Center for providing the GridSat Climate Data Record (CDR) radiation data. This CDR was originally developed by Ken Knapp and colleagues for NOAA’s CDR Program. We acknowledge EUMETSAT Satellite Application Facility on Climate Monitoring (CM SAF) for providing the cloud mask and SARA-2 radiation data. We thank Florian Filipitsch and Ralf Becker from DWD for providing additional information on measurements of solar radiation by the station network operated by DWD. We further thank the University of Wyoming and MeteoSwiss for providing radiosonde data. We finally thank three anonymous reviewers for their constructive comments, which helped to improve this article. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors declare no potential conflict of interest.

DATA AVAILABILITY STATEMENT

Data are available from the following sources.

ICON (operational) and ICON-ART (pre-operational) forecast: <https://www.doi.org/10.5445/IR/1000157325>.

ECMWF IFS forecast (ens, hres): <https://www.ecmwf.int/en/forecasts/dataset/operational-archive> (last access: 2023-09-20).

CAMS global atmospheric composition forecasts, ERA5 geopotential fields: <https://cds.climate.copernicus.eu/cdsapp> (last access: 2023-09-20).

NOAA Climate Data Record of Gridded Satellite Data from ISCCP B1 (GridSat-B1) 11 micron brightness temperature: <https://www.ncei.noaa.gov/products/gridded-geostationary-brightness-temperature> (last access: 2023-09-20).

EUMETSAT CM SAF Cloud Mask (CLAAS-3.0) and SIS (SARA-2.1, instantaneous): <https://wui.cmsaf.eu/> (last access: 2023-09-20).

SIS station measurements: <https://cdc.dwd.de/portal/> (last access: 2023-09-20).

We provide code for reproducing key analysis and visualising non-trivial model output.

Statistical analysis: <https://doi.org/10.5281/zenodo.10204763>.

Visualisation of ICON data on native grid: https://github.com/alihoshy/art_pytools.

The data that support the findings of this study are openly available in KITopen at <https://www.doi.org/10.5445/IR/1000157325>. Further data are available from ECMWF, Copernicus, EUMETSAT, NOAA, and DWD as referenced in the article. Restrictions apply to the

availability of these data, which were used under their specific licenses for this study.

ORCID

Kilian Hermes  <https://orcid.org/0000-0002-8374-5897>

Julian Quinting  <https://orcid.org/0000-0002-8409-2541>

Christian M. Grams  <https://orcid.org/0000-0003-3466-9389>

Corinna Hoose  <https://orcid.org/0000-0003-2827-5789>

Gholam Ali Hoshyaripour  <https://orcid.org/0000-0001-7770-7838>

REFERENCES

- Barkan, J., Alpert, P., Kutiel, H. & Kishcha, P. (2005) Synoptics of dust transportation days from Africa toward Italy and central Europe. *Journal of Geophysical Research: Atmospheres*, 110, D07208.
- Beyer, H.G., Costanzo, C. & Heinemann, D. (1996) Modifications of the heliosat procedure for irradiance estimates from satellite images. *Solar Energy*, 56, 207–212.
- Bou Karam, D., Flamant, C., Cuesta, J., Pelon, J. & Williams, E. (2010) Dust emission and transport associated with a Saharan depression: February 2007 case. *Journal of Geophysical Research: Atmospheres*, 115, D00H27.
- Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P. et al. (2013) Clouds and Aerosols. In: Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V. & Midgley, P.M. (Eds.) *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press.
- Bozzo, A., Remy, S., Benedetti, A., Flemming, J., Bechtold, P., Rodwell, M.J. et al. (2017) *Implementation of a CAMS-based aerosol climatology in the IFS*. Reading, United Kingdom: European Centre for Medium-Range Weather Forecasts.
- Cowie, S.M., Marsham, J.H. & Knippertz, P. (2015) The importance of rare, high-wind events for dust uplift in northern Africa. *Geophysical Research Letters*, 42, 8208–8215.
- d'Almeida, G.A. (1986) A model for saharan dust transport. *Journal of Applied Meteorology and Climatology*, 25, 903–916.
- DWD Climate Data Center. (2022) Ten-minute sum of station observations of global radiation in j/cmsup2; for Germany, version v21.3. Technical Report, DWD. https://cdc.dwd.de/sdi/pid/OBS_DEU_PT10M_RAD-G/DESCRIPTION_OBS_DEU_PT10M_RAD-G_en.pdf
- ECMWF. (2021) *IFS documentation CY47r3 - part IV: physical processes*. Technical report. Reading, United Kingdom: European Centre for Medium-Range Weather Forecasts.
- Fiedler, S., Schepanski, K., Knippertz, P., Heinold, B. & Tegen, I. (2014) How important are atmospheric depressions and mobile cyclones for emitting mineral dust aerosol in North Africa? *Atmospheric Chemistry and Physics*, 14, 8983–9000.
- Flaounas, E., Davolio, S., Raveh-Rubin, S., Pantillon, F., Miglietta, M.M., Gaertner, M.A. et al. (2022) Mediterranean cyclones: current knowledge and open questions on dynamics, prediction, climatology and impacts. *Weather and Climate Dynamics*, 3, 173–208.
- Flaounas, E., Kotroni, V., Lagouvardos, K., Kazadzis, S., Gkikas, A. & Hatzianastassiou, N. (2015) Cyclone contribution to dust transport over the Mediterranean region. *Atmospheric Science Letters*, 16, 473–478.
- Flentje, H., Mattis, I., Kipling, Z., Rémy, S. & Thomas, W. (2021) Evaluation of ECMWF IFS-AER (CAMS) operational forecasts during cycle 41r1–46r1 with calibrated ceilometer profiles over Germany. *Geoscientific Model Development*, 14, 1721–1751.
- Fluck, E. & Raveh-Rubin, S. (2023a) A 16-year climatology of the link between dry air intrusions and large-scale dust storms in North Africa. *Atmospheric Research*, 292, 106844.
- Fluck, E. & Raveh-Rubin, S. (2023b) Dry air intrusions link Rossby wave breaking to large-scale dust storms in Northwest Africa: four extreme cases. *Atmospheric Research*, 286, 106663.
- Francis, D., Eayrs, C., Chaboureaud, J.-P., Mote, T. & Holland, D.M. (2018) Polar jet associated circulation triggered a Saharan cyclone and derived the poleward transport of the African dust generated by the cyclone. *Journal of Geophysical Research: Atmospheres*, 123, 11899–11917.
- Francis, D., Eayrs, C., Chaboureaud, J.-P., Mote, T. & Holland, D.M. (2019) A meandering polar jet caused the development of a Saharan cyclone and the transport of dust toward Greenland. *Advances in science and research*, 16, 49–56.
- Francis, D., Fonseca, R., Nelli, N., Bozkurt, D., Picard, G. & Guan, B. (2022) Atmospheric rivers drive exceptional Saharan dust transport towards Europe. *Atmospheric Research*, 266, 105959.
- Ginoux, P., Prospero, J.M., Gill, T.E., Hsu, N.C. & Zhao, M. (2012) Global-scale attribution of anthropogenic and natural dust sources and their emission rates based on MODIS deep blue aerosol products. *Reviews of Geophysics*, 50, RG3005.
- Guan, B. & Waliser, D.E. (2015) Detection of atmospheric rivers: evaluation and application of an algorithm for global studies. *Journal of Geophysical Research: Atmospheres*, 120, 12514–12535.
- Guan, B. (2022) [Data] global atmospheric rivers database, version 3.
- Gueymard, C.A. & Yang, D. (2020) Worldwide validation of CAMS and MERRA-2 reanalysis aerosol optical depth products using 15 years of AERONET observations. *Atmospheric Environment*, 225, 117216.
- Hansen, J., Sato, M. & Ruedy, R. (1997) Radiative forcing and climate response. *Journal of Geophysical Research: Atmospheres*, 102, 6831–6864.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J. et al. (2020) The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146, 1999–2049.
- Holben, B.N., Tanré, D., Smirnov, A., Eck, T.F., Slutsker, I., Abuhasan, N. et al. (2001) An emerging ground-based aerosol climatology: aerosol optical depth from AERONET. *Journal of Geophysical Research: Atmospheres*, 106, 12067–12097.
- Hoshyaripour, G., Bachmann, V., Förstner, J., Steiner, A., Vogel, H., Wagner, F. et al. (2019) Effects of particle nonsphericity on dust optical properties in a forecast system: implications for model-observation comparison. *Journal of Geophysical Research: Atmospheres*, 124, 7164–7178.
- Israelevich, P., Ganor, E., Alpert, P., Kishcha, P. & Stupp, A. (2012) Predominant transport paths of Saharan dust over the Mediterranean Sea to Europe. *Journal of Geophysical Research: Atmospheres*, 117, D02205.
- Kambezidis, H.D. & Kaskaoutis, D.G. (2008) Aerosol climatology over four AERONET sites: an overview. *Atmospheric Environment*, 42, 1892–1906.
- Karlsson, K.-G., Johansson, E., Håkansson, N., Sedlar, J. & Eliasson, S. (2020) Probabilistic cloud masking for the generation of cm saf

- cloud climate data records from avhrr and sevir sensors. *Remote Sensing*, 12, 713.
- Karydis, V., Kumar, P., Barahona, D., Sokolik, I. & Nenes, A. (2011) On the effect of dust particles on global cloud condensation nuclei and cloud droplet number. *Journal of Geophysical Research: Atmospheres*, 116, D23204.
- Knapp, K.R., Ansari, S., Bain, C.L., Bourassa, M.A., Dickinson, M.J., Funk, C. et al. (2011) Globally gridded satellite observations for climate studies. *Bulletin of the American Meteorological Society*, 92, 893–907.
- Knippertz, P. (2014) Meteorological aspects of dust storms. In: Knippertz, P. & Stuut, J.-B.W. (Eds.) *Mineral dust: a key player in the earth system*. Dordrecht: Springer Netherlands, pp. 121–147.
- Knippertz, P. & Fink, A.H. (2006) Synoptic and dynamic aspects of an extreme springtime Saharan dust outbreak. *Quarterly Journal of the Royal Meteorological Society*, 132, 1153–1177.
- Kok, J.F., Storelvmo, T., Karydis, V.A., Adebisi, A.A., Mahowald, N.M., Evan, A.T. et al. (2023) Mineral dust aerosol impacts on global climate and climate change. *Nature Reviews Earth & Environment*, 4, 71–86.
- Lensky, I.M. & Rosenfeld, D. (2008) Clouds-aerosols-precipitation satellite analysis tool (CAPSAT). *Atmospheric Chemistry and Physics*, 8, 6739–6753.
- Liao, H. & Seinfeld, J. (1998) Radiative forcing by mineral dust aerosols: sensitivity to key variables. *Journal of Geophysical Research: Atmospheres*, 103, 31637–31645.
- Lohmann, U. & Feichter, J. (2005) Global indirect aerosol effects: a review. *Atmospheric Chemistry and Physics*, 5, 715–737.
- Lupu, C. & Wilhelmsson, T. (2016) A guide to simulated satellite images in the IFS.
- Magnusson, I.T., Parrington, M., Flemming, J. & Forbes, R. (2021) *Saharan dust events in the spring of 2021*, Vol. 168., Reading, United Kingdom: European Centre for Medium-Range Weather Forecasts.
- Meirink, J.F., Karlsson, K.-G., Solodovnik, I., Hüser, I., Benas, N., Johansson, E. et al. (2022) CLAAS-3: CM SAF CLOUD property dAtAset using SEVIRI-Edition 3.
- Merdji, A.B., Lu, C., Xu, X. & Mhawish, A. (2023) Long-term three-dimensional distribution and transport of Saharan dust: observation from CALIPSO, MODIS, and reanalysis data. *Atmospheric Research*, 286, 106658.
- Mona, L., Amodeo, A., Pandolfi, M. & Pappalardo, G. (2006) Saharan dust intrusions in the Mediterranean area: three years of Raman lidar measurements. *Journal of Geophysical Research: Atmospheres*, 111, D16203.
- Moulin, C., Lambert, C., Dayan, U., Masson, V., Ramonet, M., Bousquet, P. et al. (1998) Satellite climatology of African dust transport in the Mediterranean atmosphere. *Journal of Geophysical Research: Atmospheres*, 103, 13137–13144.
- Mueller, R., Pfeifroth, U. & Traeger-Chatterjee, C. (2015) Towards optimal aerosol information for the retrieval of solar surface radiation using Heliosat. *Atmosphere*, 6, 863–878.
- Muser, L.O., Hoshyaripour, G.A., Bruckert, J., Horváth, Á., Malinina, E., Wallis, S. et al. (2020) Particle aging and aerosol-radiation interaction affect volcanic plume dispersion: evidence from the raikoke 2019 eruption. *Atmospheric Chemistry and Physics*, 20, 15015–15036.
- Pfeifroth, U., Sanchez-Lorenzo, A., Manara, V., Trentmann, J. & Hollmann, R. (2018a) Trends and variability of surface solar radiation in Europe based on surface- and satellite-based data records. *Journal of Geophysical Research: Atmospheres*, 123, 1735–1754.
- Pfeifroth, U. & Trentmann, J. (2018) *Algorithm theoretical baseline document, ICDR SEVIRI radiation based on SARA-2 methods*. Technical memorandum. Reading, United Kingdom: European Centre for Medium-Range Weather Forecasts. https://www.cmsaf.eu/SharedDocs/Literatur/document/2018/saf_cm_dwd_icdr_sev_rad_atbd_1_3_pdf.html
- Pfeifroth, U., Trentmann, J., Hollmann, R., Selbach, N., Werscheck, M. & Meirink, J.F. (2018b) ICDR SEVIRI radiation-based on SARA-2 methods. Satellite Application Facility on Climate Monitoring. https://wui.cmsaf.eu/safira/action/viewDoiDetails?acronym=SARAH_V002_01
- Prospero, J.M., Ginoux, P., Torres, O., Nicholson, S.E. & Gill, T.E. (2002) Environmental characterization of global sources of atmospheric soil dust identified with the Nimbus 7 Total ozone mapping spectrometer (toms) absorbing aerosol product. *Reviews of Geophysics*, 40, 1–31.
- Reinert, D., Prill, F., Frank, H., Denhard, M., Baldauf, M., Schraff, C. et al. (2021) *DWD database reference for the global and regional ICON and ICON-EPS forecasting system, v2.1.7*. Technical report. Offenbach, Germany: Deutscher Wetterdienst.
- Rémy, S., Kipling, Z., Flemming, J., Boucher, O., Nabat, P., Michou, M. et al. (2019) Description and evaluation of the tropospheric aerosol scheme in the European Centre for Medium-Range Weather Forecasts (ECMWF) integrated forecasting system (IFS-AER, cycle 45r1). *Geoscientific Model Development*, 12, 4627–4659.
- Rémy, S., Kipling, Z., Huijnen, V., Flemming, J., Nabat, P., Michou, M. et al. (2022) Description and evaluation of the tropospheric aerosol scheme in the integrated forecasting system (ifs-aer, cycle 47r1) of ECMWF. *Geoscientific Model Development*, 15, 4881–4912.
- Rieger, D., Bangert, M., Bischoff-Gauss, I., Förstner, J., Lundgren, K., Reinert, D. et al. (2015) ICON-ART 1.0—a new online-coupled model system from the global to regional scale. *Geoscientific Model Development*, 8, 1659–1676.
- Rieger, D., Steiner, A., Bachmann, V., Gasch, P., Förstner, J., Deetz, K. et al. (2017) Impact of the 4 April 2014 Saharan dust outbreak on the photovoltaic power generation in Germany. *Atmospheric Chemistry and Physics*, 17, 13391–13415.
- Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S. et al. (2002) An introduction to meteosat second generation (msg). *Bulletin of the American Meteorological Society*, 83, 977–992.
- Schröter, J., Rieger, D., Stassen, C., Vogel, H., Weimer, M., Werchner, S. et al. (2018) ICON-ART 2.1: a flexible tracer framework and its application for composition studies in numerical weather forecasting and climate simulations. *Geoscientific Model Development*, 11, 4043–4068.
- Seifert, A., Bachmann, V., Filipitsch, F., Förstner, J., Grams, C.M., Hoshyaripour, G.A. et al. (2023) Aerosol–cloud–radiation interaction during Saharan dust episodes: the dusty cirrus puzzle. *Atmospheric Chemistry and Physics*, 23, 6409–6430 <https://acp.copernicus.org/articles/23/6409/2023/>
- Shao, Y., Wyrwoll, K.-H., Chappell, A., Huang, J., Lin, Z., McTainsh, G.H. et al. (2011) Dust cycle: an emerging core theme in earth system science. *Aeolian Research*, 2, 181–204.
- Textor, C., Schulz, M., Guibert, S., Kinne, S., Balkanski, Y., Bauer, S. et al. (2006) Analysis and quantification of the diversities of

aerosol life cycles within AeroCom. *Atmospheric Chemistry and Physics*, 6, 1777–1813.

Weger, M., Heinold, B., Engler, C., Schumann, U., Seifert, A., Föbfig, R. et al. (2018) The impact of mineral dust on cloud formation during the Saharan dust event in April 2014 over Europe. *Atmospheric Chemistry and Physics*, 18, 17545–17572.

Zängl, G., Reinert, D., Rípodas, P. & Baldauf, M. (2015) The ICON (ICOsahedral non-hydrostatic) modelling framework of DWD and MPI-m: Description of the non-hydrostatic dynamical core. *Quarterly Journal of the Royal Meteorological Society*, 141, 563–579.

How to cite this article: Hermes, K., Quinting, J., Grams, C.M., Hoese, C. & Hoshyaripour, G.A. (2024) Impact of Saharan dust outbreaks on short-range weather forecast errors in Europe. *Quarterly Journal of the Royal Meteorological Society*, 1–20. Available from: <https://doi.org/10.1002/qj.4666>

APPENDIX

In the case study we rely on the quasi-operational dust forecast from ICON-ART for assessing the extent of the dust plume on March 3, 2021. Here we provide additional material for validation of the forecast for this particular day. For this we compare the vertical profile from ICON-ART model data along the CALIPSO overpass (Figure A1) with the vertical feature mask (VFM) from the CALIPSO retrieval (Figure A2). Comparison of the model profile with the VFM suggests agreement for the zonal extent of the dust plume between model simulations and the CALIPSO retrieval. Besides the cirrus cloud layer, key features in the VFM are reproduced by ICON-ART. The cold cirrus cloud that is recorded in GridSat BT (Figure 2) and visible in the SEVIRI dust RGB (Figure A3) as dark red colours is consistent with the CALIPSO profile. Dust is not directly visible in the RGB product, as the cirrus cloud shield obscures any signal from dust. CALIPSO does not retrieve vertical information from below the cloud layer. Individual dust aerosol pixels north from the cirrus cloud, however, suggest the presence of dust in higher altitudes up to 9 km, which means a vertical overlap of dust with the cirrus cloud layer. As model data show agreement in the horizontal extent of the dust plume compared with CALIPSO measurement data, as well as in the vertical extent of the dust plume south from 42° and north from 48°, we conclude that ICON-ART shows reasonable skill in reproducing the spatial structures of the dust plume.

To investigate further whether ICON-ART reproduces the general vertical profile of atmospheric variables in the area of the observed cirrus cloud with reasonable skill, we provide an additional comparison against measured values from the radiosonde launched at Payerne (Switzerland,

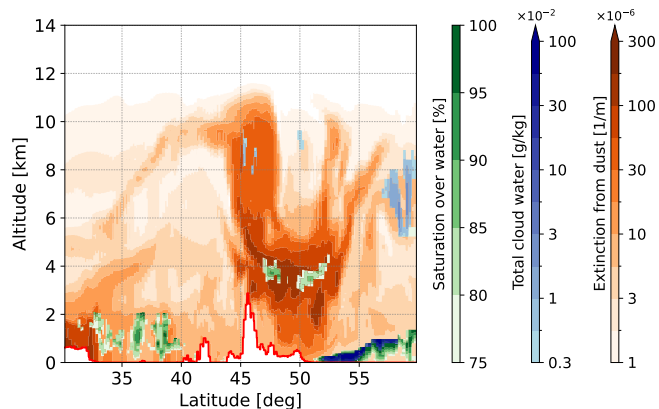


FIGURE A1 Vertical cross-section on March 3, 1200 UTC along the CALIPSO track from ICON-ART quasi-operational +12 hr forecast. Marked with color shading are extinction from dust aerosol, total cloud water (liquid water + ice), and saturation over water. Surface marked with a solid line. [Colour figure can be viewed at wileyonlinelibrary.com]

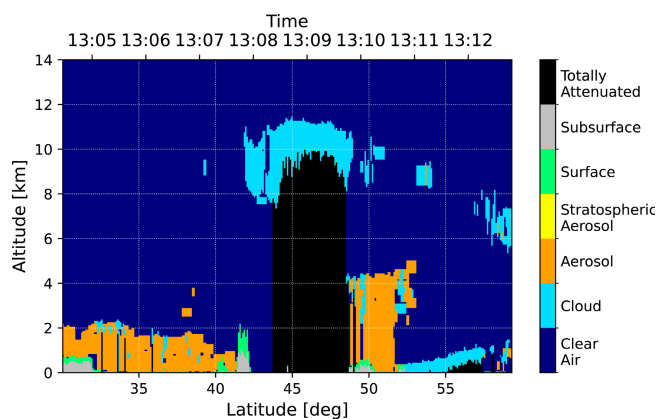


FIGURE A2 Vertical cross-section with scene identification as VFM. Data from the CALIPSO overpass on March 3, 2021, approximately 1304–1313 UTC. [Colour figure can be viewed at wileyonlinelibrary.com]

46.81° N, 6.94° E) on March 3, 2021, 1200 UTC. Figure A4 shows the radiosonde profile and the values from the closest model cells of ICON-ART as a shaded range. The modelled temperature profile fits measurements well, but shows a slight underestimation of the tropopause height. The modelled moisture profile fits with measurements in most sections of the profile, but shows deviations close to the cloud top of the observed cirrus at around 230 hPa (about 11 km), where moisture is slightly overestimated, and below the cloud top at around 400 and 700 hPa, where moisture is underestimated by the model. A similar difference at 350–400 hPa between model and measurements is shown in figure 7a of Seifert et al. (2023). With their proposed subgrid parametrisation for dust-induced cloud decks this difference vanishes, indicating improved dust

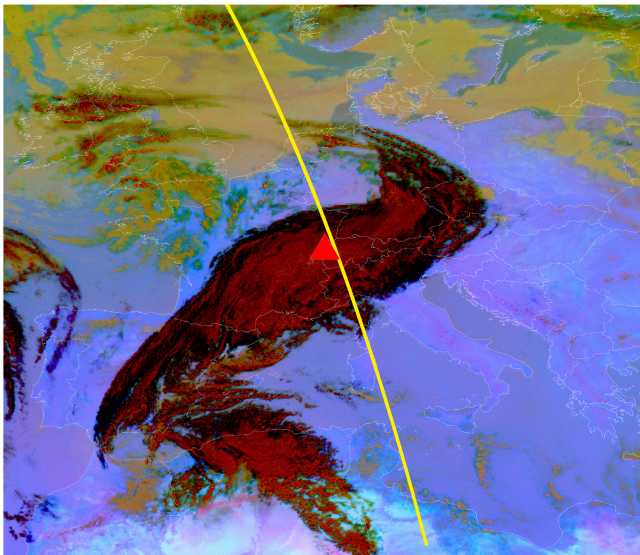


FIGURE A3 *Meteosat* SEVIRI dust RGB (Lensky & Rosenfeld, 2008) for March 3, 2021, 1200 UTC. Surface track of CALIPSO overpass (approximately 1304–1313 UTC) marked with a solid line, Payerne radiosounding station marked with a triangle. High clouds appear dark red, elevated dust over warm surfaces appears pink. [Colour figure can be viewed at wileyonlinelibrary.com]

representation also improving vertical moisture transport. Taking into account the uncertainty in the exact location of measurement due to balloon drift, the model shows very good agreement with the measurement data. Underestimations from the measured moisture profile occur at altitudes below the cloud top, while the model slightly overestimates moisture around the cloud-top height. Summarising, the model is generally capable of reproducing the temperature and moisture conditions that lead to the formation of the observed cirrus cloud layer.

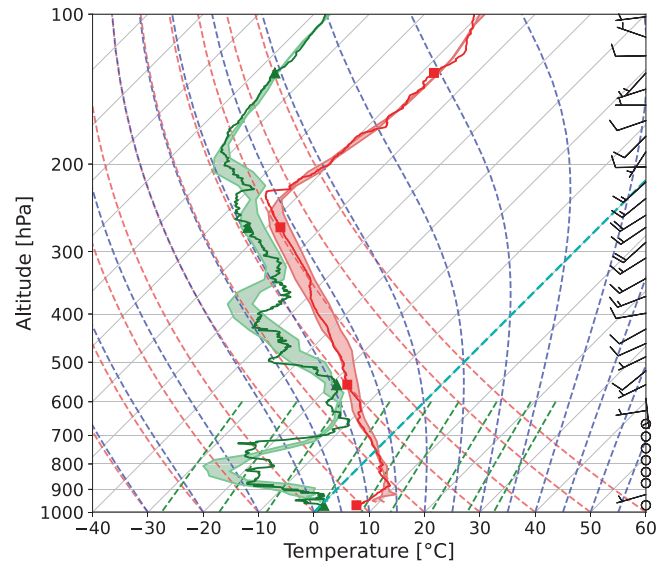


FIGURE A4 Radiosounding for Payerne (Switzerland) at March 3, 2021, 1200 UTC in a skew T-log P diagram. Grey solid horizontal lines show isobars, grey solid lines from bottom left to top right show isotherms. Green dashed lines show equal water-vapour mixing ratio, red dashed lines show the dry adiabatic lapse rate, and blue dashed lines show the moist adiabatic lapse rate. Measurement or model values for air temperature in red and marked with squares, for dew point in green and marked with triangles. Radiosonde measurement data as solid lines, range of ICON-ART model values from the closest six cells to the radiosounding launch location as shaded area. Model data for March 3, 2021, 1200 UTC from a +12 hr forecast. [Colour figure can be viewed at wileyonlinelibrary.com]