Impacts of increasing the aerosol complexity in the Met Office global numerical weather prediction model.

Article

Published Version

Creative Commons: Attribution 3.0 (CC-BY)

Open Access


It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

Published version at: http://www.atmos-chem-phys.net/14/4749/2014/acp-14-4749-2014.html
To link to this article DOI: http://dx.doi.org/10.5194/acp-14-4749-2014

Publisher: Copernicus Publications

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur
Impacts of increasing the aerosol complexity in the Met Office global numerical weather prediction model

J. P. Mulcahy¹, D. N. Walters¹, N. Bellouin¹,* and S. F. Milton¹
¹Met Office, FitzRoy Road, Exeter, EX1 3PB, UK
*now at: Department of Meteorology, University of Reading, Reading, RG6 6BB, UK

Correspondence to: J. P. Mulcahy (jane.mulcahy@metoffice.gov.uk)

Received: 9 October 2013 – Published in Atmos. Chem. Phys. Discuss.: 21 November 2013
Revised: 21 March 2014 – Accepted: 25 March 2014 – Published: 13 May 2014

Abstract. The inclusion of the direct and indirect radiative effects of aerosols in high-resolution global numerical weather prediction (NWP) models is being increasingly recognised as important for the improved accuracy of short-range weather forecasts. In this study the impacts of increasing the aerosol complexity in the global NWP configuration of the Met Office Unified Model (MetUM) are investigated. A hierarchy of aerosol representations are evaluated including three-dimensional monthly mean speciated aerosol climatologies, fully prognostic aerosols modelled using the CLASSIC aerosol scheme and finally, initialised aerosols using assimilated aerosol fields from the GEMS project. The prognostic aerosol schemes are better able to predict the temporal and spatial variation of atmospheric aerosol optical depth, which is particularly important in cases of large sporadic aerosol events such as large dust storms or forest fires. Including the direct effect of aerosols improves model biases in outgoing long-wave radiation over West Africa due to a better representation of dust. However, uncertainties in dust optical properties propagate to its direct effect and the subsequent model response. Inclusion of the indirect aerosol effects improves surface radiation biases at the North Slope of Alaska ARM site due to lower cloud amounts in high-latitude clean-air regions. This leads to improved temperature and height forecasts in this region. Impacts on the global mean model precipitation and large-scale circulation fields were found to be generally small in the short-range forecasts. However, the indirect aerosol effect leads to a strengthening of the low-level monsoon flow over the Arabian Sea and Bay of Bengal and an increase in precipitation over Southeast Asia. Regional impacts on the African Easterly Jet (AEJ) are also presented with the large dust loading in the aerosol climatology enhancing of the heat low over West Africa and weakening the AEJ. This study highlights the importance of including a more realistic treatment of aerosol–cloud interactions in global NWP models and the potential for improved global environmental prediction systems through the incorporation of more complex aerosol schemes.

1 Introduction

While the important role of aerosols in climate prediction studies has long been recognised (Forster et al., 2007; Haywood and Boucher, 2000; Houghton et al., 1996) the impact of aerosol–cloud–radiation feedbacks in global numerical weather prediction (NWP) models is less well understood. Aerosol particles modify the Earth’s radiation balance through the scattering and absorption of solar and infra-red radiation (the direct aerosol effect). They also act as cloud condensation nuclei on which cloud droplets can form. Consequently, increasing concentrations of aerosols can alter the microphysical and optical properties of clouds (the indirect aerosol effect) (Lohmann and Feichter, 2005), such as cloud albedo (Twomey, 1977), cloud lifetimes and precipitation efficiency (Rosenfeld, 2000). Increased absorption in the long-wave (LW) and short-wave (SW) spectral regions due to the presence of absorbing aerosol species, such as mineral dust and black carbon (BC), modifies the atmospheric heating profile affecting cloud cover and atmospheric stability which can subsequently impact large-scale circulation patterns (semi-direct effects). Aerosol particles can also be hazardous to human health, severely reducing visibility and affecting air quality (Prospero et al., 2001; Prospero, 2001).
has also been proposed that mineral dust particles provide an effective medium for the efficient transport of bacteria, and potentially pathogens, across large regions of the globe (Prospero et al., 2005). Furthermore, dust can adversely impact aviation, military operations and technology in dust-prone locations.

Most global climate models include increasingly complex aerosol schemes and their influences on climate (Belouin et al., 2011; Stier et al., 2005; Ghan et al., 2001). In contrast, aerosols have historically been poorly represented in global NWP models. Most operational systems prescribe fixed aerosol distributions for the direct aerosol effect and few, if any, incorporate the aerosol indirect effects. This is mainly due to the additional complexity and large computational resources required to include fully prognostic aerosol schemes in global high-resolution operational forecasting systems but also due to a limited understanding of aerosol–cloud–radiation feedbacks in short-range forecasts. NWP systems are constrained at short lead times by the data assimilation of near-real-time weather observations which implicitly include, for example, the impact of aerosols on atmospheric temperature profiles. It has been previously assumed in NWP communities that such advanced data assimilation systems remove the requirement of including interactive aerosol schemes but this view is rapidly changing and the potential benefits of increasing aerosol complexity and fully coupling to the meteorology in high-resolution systems is being advocated (Zhang, 2008).

On NWP timescales (5–10 days), Rodwell and Jung (2008) show an improvement in forecast skill and general circulation patterns in the tropics and extra-tropics by using a monthly varying aerosol climatology rather than a fixed climatology in the European Centre for Medium-Range Weather Forecasting (ECMWF) global forecasting system. Mineral dust aerosol is the largest contributor to the global aerosol load and therefore many NWP studies to date have focused on the radiative impact of mineral dust in the major dust source regions. Dust induces a thermal dipole effect, namely a warming within the dust layer and a cooling of the surface below (Reale et al., 2011; Perez et al., 2006). Such thermal stratification leads to increased atmospheric stability during the day and decreased stability at night, affecting the diurnal cycle of precipitation and wind speed (Zhao et al., 2011; Heinhold et al., 2008). Improved representations of dust have also been shown to lead to a northward and upward shift in the African Easterly Jet (AEJ) in better agreement with observations (Reale et al., 2011; Wilcox et al., 2010; Tompkins et al., 2005).

The omission of mineral dust in the global NWP configuration of the Met Office Unified Model (MetUM) results in a large bias of up to 55 W m\(^{-2}\) in outgoing long-wave radiation (OLR) fields over West Africa (Haywood et al., 2005). Milton et al. (2008) attributed biases of a similar magnitude in the surface and top-of-atmosphere (TOA) SW fluxes to missing absorbing aerosol species (dust and biomass burning aerosol) over the same region. These model biases are not solely limited to the MetUM and similar biases have been found over Northwest Africa using an ensemble of eight different models (Allan et al., 2011).

Reale et al. (2011) investigate the direct impact of aerosols in the GEOS-5 (Goddard Earth Observing System) global high-resolution operational forecasting system using the GOCART bulk aerosol model (Colarco et al., 2010). They argue that standard metrics used to evaluate NWP model skill, such as the global mean 500 hPa geopotential height anomaly, are not suitable to assess the impact of aerosols as benefits can often be on local scales associated with significant aerosol events such as large dust outbreaks or wild fire episodes. Grell and Baklanov (2011) advocate the use of fully coupled chemical and weather forecasting systems for both air quality and weather forecasting applications, highlighting benefits such as improved transport of chemical species in the former and a better representation of chemical species required by the latter. They and other studies (e.g. Grell et al., 2011) show that coupling aerosols to radiation and microphysics schemes in high-resolution weather forecasting models may improve forecasts of temperature and wind during a significant wild fire event in Alaska.

In this study we investigate the direct and indirect effects of aerosols in the global NWP configuration of the MetUM. The aim of this paper is to determine the appropriate level of aerosol complexity required in operational NWP systems by evaluating a hierarchy of different aerosol representations. These aerosol representations span different levels of complexity ranging from a simple climatological representation of aerosols to fully online aerosols coupled to the model dynamics and radiation. The ability to accurately forecast aerosol spatial and vertical distributions is of key importance in any attempt to demonstrate the benefits of using online interactive aerosols over an aerosol climatology. To this end the development of aerosol data assimilation techniques has substantially furthered our near-real-time aerosol forecasting capability. The GEMS (Global and regional Earth System Monitoring using satellite and in situ data) project successfully developed an integrated global and regional monitoring system of the key greenhouse gases, reactive gases and aerosols (Hollingsworth et al., 2008), which has been further developed in the follow-on project MACC (Monitoring Atmospheric Composition and Climate). In this study we use speciated data assimilated aerosol fields from the GEMS system to initialise our interactive aerosol simulations.

We aim to further our understanding of the direct and indirect aerosol effects on the predictive skill of the model by examining the relative importance of these effects at each level of aerosol complexity on the radiation budget, hydrological cycle and global circulation patterns in the model. Section 2 describes in detail the model and aerosol configurations used. Section 3 gives an overview of the observations used to evaluate the model’s aerosol, radiation and meteorological fields. The aerosol forecasts are evaluated in Sect. 4.
Section 5 presents and discusses the impacts of aerosols on the model radiation fields. Section 6 evaluates the impact on aerosols on the predictive skill of the model as well as impacts on the model precipitation and circulation patterns. Section 7 discusses the significance of these findings for global NWP models while a summary and conclusions are provided in Sect. 8.

2 Model description and experimental design

The atmospheric simulations in this study use a deterministic global NWP configuration of the Met Office Unified Model (MetUM) (Cullen, 1993) based on that in the Met Office’s operational NWP suite between 14 July and 2 November 2010; this is designated global NWP cycle G53. Its dynamical core uses a semi-implicit semi-Lagrangian formulation to solve the non-hydrostatic, fully compressible deep atmosphere equations of motion discretised onto a regular longitude/latitude grid (Davies et al., 2005). The radiation scheme employed is the two-stream radiation code of Edwards and Slingo (1996) with six and nine bands in the short-wave (SW) and long-wave (LW) parts of the spectrum respectively. The atmospheric boundary layer is modelled with the turbulence closure scheme of Lock et al. (2000) with modifications described in Lock (2001) and Brown et al. (2008), whilst the land surface and its interaction with the atmosphere are modelled using the Met Office Surface Exchange Scheme (MOSES, Cox et al., 1999). Convection is represented with a mass flux scheme based on Gregory and Rowntree (1990) with various extensions to include downdraughts (Gregory and Allen, 1991) and convective momentum transport. Large-scale precipitation is modelled with a single-moment scheme based on Wilson and Ballard (1999), whilst cloud is modelled using the prognostic cloud fraction and prognostic condensate (PC2) scheme (Wilson et al., 2008a, b).

The horizontal grid spacing used is $0.375^\circ \times 0.5625^\circ$, which corresponds to a resolution of approximately 40 km in the mid-latitudes. This is slightly coarser than the 25 km resolution used operationally in G53, due to the additional computational expense in the experiments using prognostic aerosol. In the vertical, we use the 70 vertical level set $L70(50, 20)_80$, which has 50 levels below 18 km, 20 levels above this and a fixed model lid 80 km from the surface.

All experiments presented cover the period from 3 June to 24 July 2009, although results presented below do not include the first 2 weeks, to allow simulations using prognostic aerosol some time to “spin up”. We perform two main forecasts per day, each 5 days in length, initialised at 00:00 UTC (00Z) and 12:00 UTC (12Z). The atmospheric state in these forecasts is initialised using a continuous 6-hourly cycle of four-dimensional variational data assimilation (4D-Var) (Rawlins et al., 2007), the land surface using a soil moisture analysis scheme (Best and Maisey, 2002) and the sea-surface temperature and sea ice using OSTIA (Operational Sea-surface Temperature and sea Ice Analysis) (Donlon et al., 2012), which is continued throughout the forecast. The first forecast in the cycle for each experiment is initialised from data held in the Met Office operational archive.

In order to investigate the role of aerosols, and the appropriate level of complexity required in a global deterministic NWP system, we have devised a hierarchy of experiments ranging from a very simple parametrisation based solely on the surface type in each grid box through to a speciated prognostic aerosol scheme initialised via near-real time data assimilation. These experiments are described in more detail in Sect. 2.1. A significant advantage in using the MetUM for these experiments is that the model includes the option to use a relatively comprehensive prognostic aerosol scheme: CLASSIC (Coupled Large-scale Aerosol Simulator for Studies in Climate), originally designed for use in climate change simulations, and described in more detail in Sect. 2.1.1. Finally, the Met Office does not currently have an assimilation system for atmospheric composition, so for experiments with initialised aerosol fields we use aerosol fields from the GEMS aerosol forecasting system as described in Sect. 2.1.3.

2.1 Aerosol representations in the model

Table 1 lists the hierarchy of model experiments that were conducted. The control (CNTRL) simulation includes the simplest treatment of aerosol and consists of a very basic, two-dimensional climatology with fixed aerosol properties over land and ocean (Cusack et al., 1998). This was the operational aerosol configuration in the NWP model prior to July 2010, after which it was replaced by the more realistic monthly-mean aerosol climatologies described in Sect. 2.1.2. These aerosol climatologies and their direct radiative effect are included in the aerosol climatology (CLIM) simulation. Fully prognostic aerosols using CLASSIC are included in the AER simulations, while these prognostic aerosol fields are initialised with GEMS aerosol data in the INIT simulations. For the prognostic aerosol simulations, two separate simulations were carried out in each case, modelling the direct aerosol effect only (AER_DIR and INIT_DIR) and the combined direct and indirect aerosol effects (AER_DIR_INDIR and INIT_DIR_INDIR). In the results presented in Sects. 5 and 6 all aerosol experiments are evaluated against CNTRL.

2.1.1 CLASSIC aerosol scheme

The CLASSIC aerosol scheme was developed by the Met Office Hadley Centre for use in climate model studies and is described in detail by Bellouin et al. (2011). Up to eight aerosol types can be explicitly modelled in CLASSIC as external aerosol species: sulfate, biomass burning, fossil fuel black carbon (FFBC), fossil fuel organic carbon (OCFF), sea salt, mineral dust, nitrate and secondary organic aerosol (termed biogenic aerosol). The mineral dust scheme is based
on the mineral dust parametrisation described by Woodward (2001), with a number of modifications described by Woodward (2011). In addition, we use a spatially fixed size distribution of the emitted dust following the approach of Zender et al. (2003). While CLASSIC does have the capability of modelling nitrate aerosol, we do not include it here as it requires a fully online chemistry scheme which is currently too costly to run at NWP resolutions. Biogenic aerosol is represented by a monthly mean climatology derived from modelled terpene emissions using the STOCHEM chemistry transport model (Derwent et al. (2003)). Sea salt aerosol is a diagnostic species only, with concentrations of the jet and film mode diagnosed from near-surface wind speeds using parametrisations developed by O'Dowd et al. (1997), and does not undergo advection or deposition in the model. All other aerosol species are transported and deposited as prognostic tracers by the model’s tracer advection and deposition schemes respectively. Aerosols are removed by wet and dry deposition processes. Dry deposition is parametrised analogous to electrical resistance (Seinfeld and Pandis, 2006). The mechanisms for wet removal of hydrophilic aerosols are via in-cloud scavenging by large-scale and convective precipitation and below-cloud (washout) large-scale and convective scavenging of hydrophobic aerosol. In order to account for re-evaporation of precipitation, in-cloud aerosol is transferred to the accumulation mode in proportion of the amount of precipitation that re-evaporates in each model level. Aerosol emissions used to drive the CLASSIC scheme are taken from the AeroCom-2 hindcast data set (Diehl et al., 2012) and are based on the year 2005. The emissions are updated daily throughout the model simulations. The aerosols are free-running in the forecasts and are not constrained by any observations.

### 2.1.2 Aerosol climatologies

The current operational global NWP model uses monthly mean three-dimensional climatologies of mass mixing ratio of all CLASSIC aerosol species. Multi-year monthly means of the aerosol fields are derived from a 20 yr HadGEM-2 climate model (Martin et al., 2011) simulation using the CLASSIC scheme, except for mineral dust which was derived from a 10 yr run. The aerosol fields from these model simulations have been shown to compare well with observations (Belouin et al., 2011), capturing key aerosol features such as increased biomass burning loadings over central Africa in January, dust emissions over West Africa and subsequent transport across the Atlantic Ocean and pollution over South and East Asia in the summer.

#### 2.1.3 Initialised aerosol fields

The GEMS forecasting system includes an operational data assimilation and forecasting system for tropospheric aerosols within the ECMWF Integrated Forecasting System (IFS). Both natural (sea salt, mineral dust) and anthropogenic (organic carbon, BC and sulfates) aerosol components are represented. The forecasting system includes the near-real-time data assimilation of MODIS total aerosol optical depth (AOD) at 550 nm. The aerosol assimilation has been shown to be most effective in capturing high AOD events (Morton et al., 2009; Mangold et al., 2011), such as large dust plumes off the West African coast, which might otherwise be missed by the free-running aerosol forecast (Benedetti et al., 2009). In the absence of an aerosol data assimilation system in the Met Office we use assimilated aerosol fields from the GEMS forecasting system to initialise the CLASSIC aerosol species in the INIT simulations. We use data from the GEMS near-real-time experiment rather than the more recent MACC forecasts as neither the MACC near-real-time forecasts nor the reanalysis were available for the June to July 2009 time period at the time the model experiments were being conducted. The resolution of the GEMS forecasts is T159L60, which corresponds to an approximately 1.125° horizontal resolution and 60 vertical levels.

In order to use the GEMS aerosol fields to initialise the CLASSIC scheme in the MetUM it was necessary to regrid the fields to the MetUM vertical and horizontal grids as well as map them to the appropriate CLASSIC aerosol species. The aerosol model used in the GEMS aerosol forecasting systems is based on the LOA/LMD-Z model (Reddy et al., 2005) and contains five tropospheric aerosol types: sulfate, sea salt, dust, organic matter and BC. Figure 1 is a schematic outlining the procedure used to map the GEMS aerosol species to the equivalent species in CLASSIC. The model variable mapped was the aerosol mass mixing ratio. Sea salt being a diagnostic variable in CLASSIC was not initialised. GEMS dust has a smaller number of size bins (three...
bins representing 0.03 to 20 microns) than CLASSIC and so was split between the six CLASSIC dust size bins (representing 0.03–31.6 µm). GEMS contains only one sulfate (SO$_4$) variable which was split with a ratio of 9 : 1 into the accumulation and Aitken SO$_4$ modes of CLASSIC, based on the global mean ratio of these species in CLASSIC climate simulations. Sulfur dioxide (SO$_2$) which is an important precursor to SO$_4$ formation along with dimethyl sulfide (DMS) was not available from the GEMS archive and so could not be initialised. While CLASSIC has separate dedicated schemes for OCFF, FFBC, and biomass burning aerosol, the GEMS aerosol system represents these aerosols as two species: organic matter (OM) and BC. The GEMS OM aerosol includes organic carbon contributions from fossil fuel and biomass burning sources as well as having a biogenic aerosol component, while the GEMS black carbon represents both fossil fuel and biomass burning BC aerosol. Consequently, the CLASSIC biomass burning scheme was not used in the initialisation experiments. Instead the biomass burning emissions were split into their OC and BC parts and added to the OC and BC emissions used in the OCFF and FFBC schemes respectively. This was deemed acceptable as the OCFF and biomass burning parametrisations in CLASSIC are very similar, as are the microphysical and optical properties used for each species (Bellouin et al., 2011). Thus, the CLASSIC OCFF and FFBC schemes are, in this configuration at least, representative of the GEMS organic matter and black carbon aerosol species.

The aerosol fields were initialised at the start of the 00Z and 12Z forecasts and were subsequently free-running through the remainder of the 5-day forecast. As the NWP forecast effectively runs from $T - 3$ h, $T + 9$ h aerosol mass mixing ratios from the 00Z and 12Z GEMS aerosol forecasts were used to initialise the 12Z and 00Z forecasts respectively. Comparison of the aerosol optical depth (AOD) at 550 nm before and after initialisation along with the corresponding derived AOD from the GEMS aerosol forecast (not shown) show that the main impact of the initialisation is to increase the aerosol loading in the Northern Hemisphere (NH), where previously there was a factor of 2 difference between the two models. After initialisation, there is a much improved agreement between the two models, with a global mean difference in AOD of 0.01. This good agreement gives us confidence that the mapping procedure employed is fit for purpose.

2.2 Treatment of the direct and indirect aerosol effects

The aerosols are coupled to the model’s radiation and cloud microphysics schemes allowing the direct and indirect effects to be modelled. Each CLASSIC aerosol species has a prescribed log-normal number size distribution and a set of wavelength-dependent refractive indices as detailed in Table A1 of Bellouin et al. (2011). The aerosol optical properties are calculated offline using Mie calculations and are then averaged across the short-wave and long-wave bands of MetUM. Hygroscopic growth is parametrised as a function of relative humidity (Fitzgerald (1975), Haywood et al. (2003b)) for hygroscopic species. The aerosol optical properties of the aerosol climatologies are the same as those of the full CLASSIC scheme. The resulting optical properties are stored in look-up tables for use during the model integration by the radiation scheme. The semi-direct effect is implicitly included, as the change in temperature profile due to heating from absorbing aerosol, such as BC, feeds back onto the modelled cloud fields.

Indirect radiative effects are included for all species except mineral dust and BC which are assumed to be hydrophobic. The total cloud condensation nuclei (CCN) concentration is calculated from the accumulation and dissolved aerosol modes of the hydrophilic species using the aerosol mass and prescribed size distributions. The cloud droplet number concentration (CDNC) is then calculated using the Jones et al. (2001) relationship:

$$\text{CDNC} = 3.75 \times 10^8 (1.0 - \exp(-2.5 \times 10^{-9} \text{CCN})).$$

A minimum value for CDNC of $5 \times 10^6 \text{m}^{-3}$ is assumed.
For the first indirect effect, the radiation scheme uses the CDNC to obtain the cloud effective radius and cloud albedo. For the second indirect effect, the large-scale precipitation scheme uses the CDNC to calculate the autoconversion rate of cloud water to rainwater based on the Tripoli and Cotton (1980) autoconversion scheme. The indirect aerosol effects are described in detail by Jones et al. (2001). The indirect effects from the aerosol climatologies are currently not included in the model. In this case, the MetUM assumes a constant CDNC of 100 cm$^{-3}$ and 300 cm$^{-3}$ over ocean and land respectively. Aerosol impacts on ice clouds are currently not included.

3 Observations

3.1 AERONET global surface aerosol measurements

The Aerosol Robotic Network (AERONET) (Holben et al., 1998) is an extensive ground-based observational network of sun- and sky-scanning radiometers, which provide quality assured measurements of aerosol optical properties from a range of different aerosol regimes across the globe (Holben et al., 2001). In this study, post-processed and quality assured (Level 2.0) AOD at 440 nm is used from a number of selected stations to evaluate the model-derived total AOD in disparate aerosol regimes.

3.2 MODIS AOD

The MODerate resolution Imaging Spectroradiometer (MODIS) onboard the Terra and Aqua satellites makes continuous global measurements of the upwelling radiance at the top of the atmosphere in 36 spectral bands. Its wide spectral range, high spatial resolution and near-daily global coverage make it an ideal platform to observe the changes in global, tropospheric aerosol distributions. A number of aerosol products at a resolution of 10 km $\times$ 10 km, including the spectral total AOD, Ångström exponent and fine-mode AOD, are derived over both land and ocean using seven of these well-calibrated channels in the region of 470 nm to 2.1 $\mu$m (Remer et al., 2005). AOD measurements from MODIS have been available since 1999 for the Terra platform and since 2002 from the Aqua platform. The standard MODIS aerosol retrieval algorithm for land surfaces (Kaufman et al., 1997) does not retrieve aerosol information over bright surfaces such as deserts. The “Deep Blue” algorithm has been subsequently developed by Hsu et al. (2004) for this purpose. In this study we use a merged Collection 5.1 Level 2 daily AOD product, which includes AOD from the “Deep Blue” algorithm over bright land surfaces and AOD from the standard land and ocean algorithms elsewhere.

3.3 MISR AOD

The Multi-angle Imaging SpectroRadiometer (MISR) instrument on board the Terra satellite observes the TOA radiances at nine different zenith angles and four narrow spectral bands (446, 558, 672 and 866 nm) (Diner et al., 1998). The different viewing angles allows MISR to retrieve aerosol information over bright surfaces (Martonchik et al., 1998). Near global coverage is obtained every 9 days at the equator, and 2 days at the poles. Measurements of AOD have been available since the year 2000. In this study the level-3 daily global MISR AOD product was used at 0.5° $\times$ 0.5° resolution.

3.4 TRMM precipitation measurements

Precipitation measurements are taken from the Tropical Rainfall Measuring Mission (TRMM) ( Huffman et al., 2001), which provides high-quality infrared precipitation measurements and error estimates. The data are available with a 3-hourly temporal resolution and a spatial resolution of 0.25°. Coverage extends from 50° N to 50° S.

3.5 GERB radiation measurements

The GERB (Geostationary Earth Radiation Budget) instrument onboard the Meteosat-8 geostationary satellite routinely measures broadband SW and LW radiative fluxes from TOA radiances (Harries et al., 2005). Calibrated SW radiances are converted to fluxes using angular distribution models, which are themselves a function of the identified scene type derived from the Spinning Enhanced Visible and Infrared Imager (SEVIRI, Schmetz et al., 2002) on Meteosat-8. GERB has a temporal resolution of approximately 17 min and spatial scales of 50 km. The GERB data has an accuracy of 2.25 % for SW irradiance and 0.96 % for LW irradiance. Further details of the GERB instrument and its use in the evaluation of the MetUM are given by Allan et al. (2011) and Allan et al. (2007).

3.6 ARM radiation measurements

The US Atmospheric Radiation Measurement (ARM) programme (Ackerman and Stokes, 2003) produces a wide range of atmospheric measurements at a number of locations across the globe. We use surface radiative flux measurements at broadband radiometers to evaluate the SW and LW downwelling and upwelling components. Radiative flux measurements are made every minute. In this study we use the ARM climate modelling best estimate products (Xie et al., 2010), where the data have been quality controlled and processed to a 1 h temporal resolution.
In this section we evaluate the skill of the forecasts of AOD produced in the prognostic aerosol and CLIM simulations. Figure 2 shows the global mean spatial distribution of simulated AOD (550 nm) at $T + 120$ h into the forecast (day 5) from all simulations. Also shown are the corresponding mean spatial AOD distributions from MISR and MODIS. Qualitatively, the spatial distributions from all model simulations compare well with the satellite observations. All models capture the mineral dust plume observed over West Africa and the tropical Atlantic Ocean but there are some significant differences in the details of this distribution. The CLIM mineral dust is located too far north compared with the prognostic dust simulations and satellite observations. It has a significant localised dust loading over the northwest of Africa, while the prognostic simulations have a more widespread distribution of dust across the whole of North Africa advecting westwards across the Atlantic Ocean and eastwards from the Horn of Africa across the Arabian Sea in good agreement with both MODIS and MISR observations. CLIM also has a significantly higher aerosol loading over China and over the tropical Atlantic Ocean.
The elevated AODs are not observed over China, although satellite observations over this region are relatively sparse. In contrast, both AER and INIT simulations appear to slightly underestimate the AOD over the Indo-Gangetic Plain compared with both AERONET (see comparison at Kanpur, India in Fig. 4d) and MODIS observations in this region. The elevated AOD in CLIM can be attributed to the inclusion of the nitrate aerosol climatology, while there is no representation of nitrate in the prognostic aerosol simulations. While anthropogenic emissions of nitrogen and ammonia are known to be increasing in this region (Bellouin et al., 2011; Galloway et al., 2004) and are therefore an important aerosol source, carbonaceous emissions in this region are also likely to be underestimated (Bond et al., 2013) in the prognostic aerosol simulations.

The AER simulations tend to underestimate the AOD over the high latitudes in the NH. Again, this is most likely caused by the lack of nitrate aerosol in the CLASSIC scheme, which is an important aerosol source in industrialised areas such as Europe and China. This bias is improved in CLIM, and also in the INIT experiments where the main effect of the initialisation, as already stated in Sect. 2.1.3, is the addition of aerosol in the NH.

Figure 3 presents the time series of AOD from all aerosol simulations at \( T + 120 \) at a number of selected AERONET sites where mineral dust is the predominant aerosol species. Figure 3a and b considers the model performance at two sites relatively close to dust sources (Saada (31° N, 8° W) and Solar Village (24° N, 46° E)). Capo Verde (16° N, 22° W), off the West African coast, and La Parguera (17° N, 67° W) in the Caribbean Islands, shown in Fig. 3c and d, are sites which frequently observe large amounts of transported dust. In dust source locations, the prognostic model simulations capture the temporal variation and magnitude of the dust remarkably well. By design, CLIM exhibits little diurnal variation in these dry, arid regions and is therefore not able to capture the day-to-day variability and misses a number of large dust events, notably between the 3 and 5 July. Dust transport across the Atlantic Ocean also appears to be well represented in the AER and INIT simulations, particularly at La Parguera. At Capo Verde, which is close to the major dust source region of the Sahara Desert, the prognostic simulations appear to advect too much dust from West Africa and CLIM has a smaller overall mean bias at this location.

A similar comparison is shown in Fig. 4 for regions of predominantly anthropogenic aerosol. Over Europe, the large urban centres of Paris (48° N, 2° E) and Rome (41° N, 12° E) are presented and for Asia, Kanpur in India (26° N, 80° E) and Hong Kong in China (22° N, 114° E) are shown. Over Europe, the initialised simulations exhibit a small positive bias (mean bias of 0.02 and 0.07 at Rome and Paris respectively) while the uninitialised simulations tend to underestimate the AOD by a similar magnitude (mean bias of −0.02 and −0.03 respectively). Overall, the agreement between model and observations is good for the prognostic simulations, although a few large AOD events are missed at the Asian sites, for example at the end of June at Kanpur and on 14 July at Hong Kong. The temporal variability of the AOD
in CLIM, particularly at Kanpur and Paris, is due solely to the large variations in the relative humidity in this area and leads to large positive biases in these regions (mean bias of 0.24 and 0.12 at Kanpur and Paris respectively).

It is worth noting that the mean bias in AOD from the INIT simulations is smaller than the bias in AER AOD at the beginning of the forecast. Figure 5 shows the time series of AOD at $T + 24$ at Saada, Cape Verde, Rome and Kanpur and highlights the impact of initialising the aerosol fields.
AER gives better agreement in dust source regions such as Saada, but away from source at Capo Verde INIT advects less dust than AER over the Atlantic Ocean in better agreement with the AERONET observations. It should be noted that the GEMS forecasting system does not assimilate observations of AOD over bright surfaces such as the Sahara Desert due to the limitations of the standard MODIS land retrieval algorithm (Kaufman et al., 1997). Thus, the initialisation is not expected to have a positive impact in these regions. At anthropogenic locations, such as Rome and Kanpur, while AER underestimates the AOD in most cases INIT has a smaller bias. Overall, such good agreement with the observations highlights the positive impact of initialising the CLASSIC fields with GEMS assimilated data at shorter forecast times.

5 Impacts on radiation

The impact of aerosol on the model radiation fields was assessed by comparison against all-sky radiative flux measurements derived from GERB measurements. Figure 6 shows the mean model bias in OLR at the start of the forecast over West Africa from the 22 June to the 22 July 2009 from all model simulations. Both model and observations have been cloud-screened by removing values where both the observations and model report cloud amounts greater than 50%. However, it is likely that some residual cloud contamination remains along the edge of the 50% contour line in Fig. 6. As the region south of 15°N is predominantly cloud-contaminated we focus our analysis to the north of this region. The impact of aerosols on the model radiation biases is evaluated early in the forecast in order to minimise the role of other model errors such as general circulation and temperature errors.

A significant positive bias of between 20 and 40 W m\(^{-2}\) is clearly seen over northwest Africa in the CNTRL simulation (Fig. 6a), in agreement with the Haywood et al. (2005) study which also used the Cusack aerosol climatology. With the inclusion of mineral dust (Fig. 6b–f), the positive bias is notably reduced in regions dominated by mineral dust aerosol. CLIM (Fig. 6b) reduces the OLR by too much in the regions of highest dust concentrations to the east of the Atlas Mountains and Algeria, leading to a large negative bias. The AER simulations (Fig. 6c and d) largely remove the bias with
Fig. 7. Reflected short-wave radiation differences at TOA between model and GERB observations from (a) CNTRL, (b) CLIM, (c) AER_DIR, (d) AER_DIR_INDIR, (e) INIT_DIR and (f) INIT_DIR_INDIR. Areas where the cloud fraction exceeds 0.5 have been removed from the comparison.

...a residual negative bias persisting over Algeria and stretching east towards the Bodele depression where the model produces more dust than in the other aerosol simulations. In the INIT simulations (Fig. 6e and f) the improvement in the OLR bias is smaller than in the AER simulations. This is due in part to smaller dust loadings in the initialised simulations at short forecast lead times.

The dust size distributions in all three aerosol representations (CLIM, AER, INIT) are also different (not shown). This has important implications for the dust radiative forcing. The size distribution of the dust climatology used in CLIM peaks towards the larger size bins which subsequently gives a larger impact in the LW spectrum. The fixed size distribution of the emitted dust in the AER simulations is based on the observations of D’Almeida (1987) and peaks in bin 4 (1–3.16 µm). The INIT dust has more mass in the sub-micron particle range following the GEMS dust size distribution and subsequently is less radiatively efficient in the LW and more efficiently scatters SW radiation (see Fig. 7). This size distribution has been subsequently revised in the latest operational MACC forecasts. As already discussed, the impact of the GEMS initialisation reduces with forecast lead time, so that by day 5, the dust distribution is very similar to the uninitialised dust distribution.

Figure 7 shows the mean model bias in the reflected short-wave radiation for the same period. The inhomogeneity of the underlying surface albedo over North Africa leads to a less well defined error signal, but the CNTRL simulation has a predominantly positive bias over this region. The impact of the dust can be seen in Fig. 7b and c, with the CLIM dust reducing the positive bias over northwest Africa but increasing the bias over the dark ocean surface of the Mediterranean Sea. Conversely, the prognostic simulations increase the outgoing reflected SW radiation, increasing the positive bias across North Africa.

Figure 8 presents the direct radiative impact of aerosols on the clear-sky net TOA radiation (sign convention used is positive downwards). The clear-sky direct radiative effect is calculated as the difference between the top of atmosphere clear-sky net radiation from the direct only simulations with and without aerosol, i.e. $\text{DRE}_\text{net} = \text{TOA}_\text{net} - (\text{AEROSOL}) - \text{TOA}_\text{net} - (\text{CNTRL})$, where...
AEROSOL can be from CLIM, AER_DIR or INIT_DIR simulations. Note this is not a true aerosol radiative perturbation, as CNTRL still contains a basic aerosol representation. The main direct impacts of aerosol are the reduction in OLR over West Africa leading to a local net TOA radiative warming and an increase in the backscattered SW radiation due to dust over the tropical Atlantic, biomass burning over Central Africa, and pollution over China, leading to a localised net TOA radiative cooling in these regions. The radiative warming seen over northern Canada and Russia is due to lower aerosol loadings in these relatively clean-air regions compared with the Cusack climatology, resulting in a reduction in the upwelling SW at TOA in these regions. INIT_DIR exerts the largest radiative perturbation with a net global mean TOA cooling of $-1.25 \text{ W m}^{-2}$.

The combined direct and indirect effect of aerosols on the all-sky net TOA radiation is shown in Fig. 9. Inclusion of the aerosol indirect effects leads to a much larger global mean impact than including the direct aerosol effect on its own (see Table 2). A net TOA radiative cooling is found over Northern and Southern Hemisphere (SH) ocean regions where low-level stratocumulus clouds are prevalent, namely off the West African and Chilean coastlines. A radiative cooling is also found over regions influenced by anthropogenic emissions such as the Mediterranean Sea and off the northeast coast of USA. A significant net radiative warming of up to $30 \text{ W m}^{-2}$ is found at high latitudes in the NH and south of $20^\circ$ S. These regions are typically pristine clean-air regions, with relatively low aerosol concentrations and correspondingly low cloud droplet number concentrations. Figure 10 shows the mean $T + 120$ potential cloud droplet number concentration (CDNC) in the troposphere from the CLIM and AER_DIR_INDIR simulations. Currently, UM configurations which do not include the indirect effects of aerosols assume a fixed value of CDNC of $100 \text{ cm}^{-3}$ and $300 \text{ cm}^{-3}$ over ocean and land surfaces respectively. Such assumptions are gross over-simplifications of the true distribution of CDNC which is closer to Fig. 10b most notably in the high-latitude regions. This subsequently leads to effectively cleaner air and a reduction in low-level cloud (approximately 20%) in these regions in the DIR_INDIR simulations and leads to a strong radiative warming at both the TOA and...
the surface. A number of recent studies report aerosol cloud condensation nuclei (CCN) values in the Arctic summer to be less than 100 cm$^{-3}$ and in many cases to be as low as 1 cm$^{-3}$ (Birch et al., 2012; Mauritsen et al., 2011; Bennartz, 2007; Bigg and Leck, 2001), which would support the low values found in the current prognostic aerosol experiments. AER_DIR_INDIR has in general smaller values of CDNC compared with the INIT simulation, due to the latter having higher aerosol concentrations particularly in the NH, with the exception of lower concentrations over dust and biomass burning regions of Africa. AER_DIR_INDIR also has low concentrations over high-altitude locations such as the Tibetan Plateau (TP) and the Andes.

The impact of aerosol on the net surface radiation from both AER simulations is shown in Fig. 11. Milton et al. (2008) highlighted large positive biases ($\sim$ 50 W m$^{-2}$) in the downwelling SW surface radiation over West Africa and linked this bias to the omission of absorbing aerosol species in the NWP model. With the inclusion of a more realistic aerosol distribution the net SW flux at the surface is reduced in dust regions by 20 to 30 W m$^{-2}$ (not shown). However, this is generally offset by an enhanced downwelling LW emission from the absorbing mineral dust. A surface radiative cooling is seen in areas of high aerosol loading over ocean, polluted regions over China and Korea, and in regions of increased amounts of stratocumulus cloud in the simulations which include the indirect aerosol effects.

A significant positive warming at both the TOA and surface is also found over South Asia and India, particularly in the AER_DIR_INDIR simulation. An initial reduction in CDNC over the TP results in a positive SW cloud forcing at short forecast lead times. This is replaced by a positive LW forcing at longer lead times due to the subsequent increase in deep convective activity and associated high cloud amount that occurs in response to the SW warming.

All global mean net radiative perturbations due to aerosols are summarised in Table 2. A larger cooling at TOA and smaller warming at the surface is found in INIT_DIR due to the enhanced scattering of the incoming SW radiation in these simulations particularly in the NH where the INIT simulations have a larger aerosol loading than AER. The TOA and surface perturbations due to the indirect aerosol effects in the INIT simulations have the same sign as AER simulations, but are smaller in magnitude due to the reduced cloud sensitivity in the NH in the INIT simulations. The global mean TOA forcing is 4.18 W m$^{-2}$ in AER_DIR_INDIR compared with 2.26 W m$^{-2}$ in INIT_DIR_INDIR. Corresponding surface forcings are 6.80 and 4.07 W m$^{-2}$ respectively. The direct aerosol effect leads to a cooling at the TOA and a warming at the surface, while inclusion of the indirect effects leads to an overall warming at TOA and an even larger warming at the surface. Overall, the prognostic aerosol simulations have a larger impact on the net radiative fluxes than CLIM.

### Table 2. The global mean day 5 all-sky net radiative forcings (W m$^{-2}$) from the different aerosol representations.

<table>
<thead>
<tr>
<th></th>
<th>TOA</th>
<th>ATM</th>
<th>SFC</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLIM</td>
<td>−0.30</td>
<td>−1.72</td>
<td>1.42</td>
</tr>
<tr>
<td>AER_DIR</td>
<td>−0.87</td>
<td>−2.91</td>
<td>2.04</td>
</tr>
<tr>
<td>AER_DIR_INDIR</td>
<td>4.18</td>
<td>−2.62</td>
<td>6.80</td>
</tr>
<tr>
<td>INIT_DIR</td>
<td>−1.48</td>
<td>−2.29</td>
<td>0.89</td>
</tr>
<tr>
<td>INIT_DIR_INDIR</td>
<td>2.26</td>
<td>−1.81</td>
<td>4.07</td>
</tr>
</tbody>
</table>

5.1 Evaluation against ARM measurements

Observations from the ARM site at Barrow, Alaska (156.6° W, 71.3° N) are used to further evaluate the indirect aerosol impacts on the model surface radiative balance in the NH high-latitude regions. T+24 h forecasts of SW and LW downwelling (SW$_{down}$, LW$_{down}$) and upwelling (SW$_{up}$, LW$_{up}$) radiative fluxes at the surface valid at 00Z and 12Z are evaluated. Again the sign convention of positive for downward radiative fluxes is used. Figure 12a compares the SW$_{down}$ radiative flux at the surface from the CNTRL and AER_DIR_INDIR simulations with ARM observations for the full simulation period. The CNTRL simulation clearly underestimates SW$_{down}$ throughout the period with a mean bias of 63.6 W m$^{-2}$ and was also found to underestimate the SW$_{up}$ component (not shown). The CNTRL model overestimates the LW$_{down}$ component at the surface.
(mean bias of 18.38 W m\(^{-2}\)) and also the emitted LW\(_{\text{up}}\) component. The underestimation of SW\(_{\text{down}}\) and overestimation of LW\(_{\text{down}}\) at surface suggests that the model is predicting too much cloud in this region and is consistent with Milton and Earnshaw (2007) who postulated that similar errors found in the surface radiation balance at the same site in boreal winter are most likely caused by an overprediction of cloud or inadequate representation of cloud properties. When the aerosol indirect effects are included the SW\(_{\text{down}}\) bias is greatly reduced to \(-0.59\) W m\(^{-2}\) and a much improved correlation with the ARM observations is found. The improvement in the SW\(_{\text{down}}\) bias is found to correlate with days when the AER\(_{\text{DIR\_INDIR}}\) predicts less cloud than CNTRL, as can be seen in Fig. 12c. In particular, between 27 June and 4 July AER\(_{\text{DIR\_INDIR}}\) predicts between 20 and 60 % less cloud than CNTRL with concomitant increases in the SW\(_{\text{down}}\) fluxes of between 100 and 300 W m\(^{-2}\), improving the model bias. A smaller impact in the LW\(_{\text{down}}\) component was found (see Fig. 12b). The upward surface SW bias is also improved but not totally removed with the inclusion of the aerosol radiative effects. This is potentially due to an inadequate representation of the surface properties in the model such as albedo, snow cover and vegetation type. The additional surface SW warming combined with the small impact on the LW\(_{\text{down}}\) increases the surface temperature and subsequently the LW radiative emission from the surface is increased.

6 Impacts on meteorological fields

6.1 Verification against model analysis

Figure 13 presents a comparison of the \(T + 120\) mean error in temperature, geopotential height, and relative humidity profiles over the NH and the tropics from 17 June to 17 July 2009 from all simulations. The model mean error is calculated relative to each simulation’s own analysis. The large radiative warming over the NH high latitudes, in response to the inclusion of the aerosol indirect effects discussed above, leads to an overall improvement in the mean error of temperature and height profiles in both AER\(_{\text{DIR\_INDIR}}\) and INIT\(_{\text{DIR\_INDIR}}\) simulations. The larger temperature change found in AER\(_{\text{DIR\_INDIR}}\) is presumably in response to the larger change in CDNC in the uninitialised CLASSIC aerosol simulations. A similar but smaller impact is seen in the tropics. Figure 14 shows the zonal mean temperature error in the CNTRL and the AER\(_{\text{DIR\_INDIR\_CNTRL}}\) difference. Figure 14a highlights a significant cold bias of \(\sim 2\) K in the NH below 700 hPa in the CNTRL simulation. The aerosol indirect effect acts to reduce this error by \(\sim 1\) K. The indirect aerosol effect also has a positive impact on temperature errors in the SH where the mid-tropospheric cold bias has reduced by up to \(\sim 0.5\) K. The location of these beneficial changes in temperature are primarily observed over remote NH land regions of northern Canada and Siberia and SH ocean regions where CDNC concentrations of 300 and 100 cm\(^{-3}\) are obviously too high in CNTRL. These findings confirm that the more accurate treatment of aerosols and their indirect effects on cloud formation in clean air regions in global NWP leads to a more accurate simulation of the atmosphere.

6.2 Clouds and precipitation

We have already seen in Sect. 5 that the indirect aerosol effects lead to a large positive TOA forcing over the NH high latitudes and a negative forcing over the subtropical and northern Pacific and Atlantic Oceans. The regions subjected to the largest TOA aerosol forcing are dominated by low-level cloud as shown in Fig. 15a. The impact of the direct and indirect aerosol effects on model mean \(T + 120\) cloud amount is shown in Fig. 15b–f. The primary impact of the direct aerosol effect is to reduce low-level cloud amount across the tropical Atlantic Ocean along the path of the dust plumes being transported west from the Sahara Desert (Fig. 15b and f). When the aerosol indirect effect is included, the change in CDNC from the fixed values used in the DIR and CNTRL simulations to variable CDNC diagnosed from the CLASSIC scheme leads to an increase in low-level maritime cloud over much of the NH oceans and a reduction.
Fig. 12. Time series of (a) SW downward, (b) LW downward fluxes at the surface from CNTRL, AER_DIR_INDIR and ARM surface measurements at the North Slope of Alaska site, (c) low cloud amount from CNTRL and AER_DIR_INDIR.
in clean polar regions. Marine stratocumulus clouds off the coasts of Namibia in Africa and Chile in South America are increased by up to 20 % in both AER_DIR_INDIR and INIT_DIR_INDIR simulations (Fig. 15d and f) leading to brighter clouds and a negative TOA forcing. In contrast, over the NH and SH high latitudes the low-level cloud decreases, most significantly over northern Canada where the low-level cloud is reduced by up to 30 %. This leads to a reduction in the outgoing SW at TOA and increases the solar radiation reaching the surface. Consequently, this reduction in cloud in both the NH and SH high latitudes leads to the improved temperature biases in these regions discussed above.

Notable increases of ~10 % in high-level cloud amount (not shown) were also found over West Africa and Southeast Asia in AER_DIR_INDIR and INIT_DIR_INDIR. This results from a positive surface forcing exerted by the aerosols in these regions. Over West Africa, the net positive forcing at the surface (see Fig. 11b) is due predominantly to the enhanced downwelling LW emission from mineral dust. Over the Indo-China Peninsula and the TP the aerosol loading is relatively low with an AOD (550 nm) of less than 0.2 and 0.1 in these regions respectively (see Fig. 19). This leads to large reductions in CDNC of more than 50 % relative to CNTRL and a subsequent large positive surface warming due to the increased SW radiation reaching the surface at short forecast lead times. This large radiative warming triggers additional convective activity in these regions leading to increases in high cloud amount and convective precipitation at longer forecast lead times. Comparisons of the mean AOD (550 nm) against a number of AERONET observations on the Indo-China Peninsula (not shown) indicate that the uninitialised CLASSIC simulations underestimate the AOD in this region, with the higher aerosol loading in INIT comparing better with the observations. So it is likely that the large sensitivities seen in this region are due to an inaccurate representation of the aerosol emissions or aerosol optical properties in this region in the model.

Figure 16 presents the change in the $T + 120$ mean global precipitation rate from each experiment relative to CNTRL. The CNTRL error shown in Fig. 16b and calculated using TRMM observations shows that the model overestimates precipitation over Southeast Asia, the North Indian Ocean, eastern parts of Africa and over much of the tropical Pacific Ocean. The model has a negative bias over continental India.
and over the tropical North Atlantic. While the global mean impact of the aerosol on the precipitation rate is small at $T + 120$ some beneficial impacts are found. Dust and smoke aerosol over the Sahel region exerts a semi-direct effect reducing the precipitation over the region by 1–3 mm day$^{-1}$ in the DIR simulations and also acts to reduce the positive biases over Venezuela and Colombia. Precipitation is also reduced in the Indo-Gangetic Plain along the southern edge of the Himalaya mountain range, an area which shows a significant positive bias in the CNTRL simulation. With the inclusion of the direct and indirect effects, the trend in rainfall change over the Sahel is not as clear between the different aerosol simulations (a small increase is found in AER but a decrease is found in the INIT simulation). An increase in precipitation over the tropical North Atlantic partially offsets the negative bias in this region. It also reduces the precipitation over the Indo-Gangetic Plain and Northeast Pakistan but notably increases the positive bias over Southeast Asia and China. Most of the precipitation over the Indian continent in the model is in the form of convective precipitation. The indirect aerosol effects as currently parametrised in the MetUM interacts with the large-scale precipitation scheme only and thus the impacts on the convective precipitation in this region are due solely to a temperature response to the aerosol. Inaccuracies in both the representation of precipitation in the model as well as the representation of the aerosol indirect effects will lead to unrealistic responses here.

### 6.3 Global circulation patterns

Figure 17 shows the mean change in mean sea level pressure (MSLP) and 925 hPa wind fields at $T + 120$ between the aerosol and control simulations. The CLIM simulation enhances the heat low over Northwest Africa by up to 2 hPa, however the prognostic dust simulations have a much smaller impact in this region. Simulations including the direct aerosol effect only increase the MSLP across NH land surfaces with the largest increase found in the INIT simulations in response to the larger NH aerosol loading and subsequent reduction in downwelling SW flux at the surface in these simulations. The most notable impacts are the reductions in MSLP over the South Asian monsoon region, North
Fig. 15. Impact of different aerosol representations on low cloud amount at $T + 120$: (a) mean low cloud amount from CNTRL simulation, (b) CLIM-CNTRL, (c) AER_DIR-CNTRL, (d) AER_DIR_INDIR-CNTRL, (e) INIT_DIR-CNTRL and (f) INIT_DIR_INDIR-CNTRL.

Canada and Arctic regions in the INDIR simulations which are co-located with the regions of significant SW warming discussed in Sect. 5. In the following section we focus on the more notable regional impacts found over South Asia and West Africa in the form of changes to the important circulation patterns associated with the South Asian monsoon and the African easterly jet.

6.4 Regional circulation patterns

Figure 18a shows the 850 hPa wind field over the South Asian monsoon region from the CNTRL simulation. The bias in the CNTRL wind field is shown in Fig. 18b and highlights an overly strong monsoon flow over North India and Southeast Asia. Aerosol-induced wind changes are shown in Fig. 18c–f. The CLIM simulation was found to have a negligible impact on the low-level circulation and so is not shown here. A consistent response to the aerosol is seen in Fig. 18c–f with an induced cyclonic circulation opposing the anticyclonic flow over India. This response appears to be due to the direct aerosol effect increasing the convergence over the North Indian Ocean. Figure 19 highlights the different AOD distributions from CLIM, AER and INIT simulations in this region. The lack of a circulation response in CLIM can be explained by the low aerosol loadings in CLIM over the North Indian Ocean compared to the prognostic aerosol simulations. The impact is strongest in the INIT simulations due to the higher AOD over the Arabian Sea compared to the AER simulations. Dust is the predominant aerosol species in the model in this region and a warming of between 0.1–0.2 K is found within the dust layer.

When the indirect aerosol effect is included the large SW forcing induced over southeast Asia enhances the heat low across the TP, increasing the anomalous convergent flow around the TP, over Southeast Asia and China. The result is an enhanced monsoon flow over the Bay of Bengal. The impact is largest in the AER_DIR_INDIR simulations due to the higher sensitivity to the change in CDNC in this simulation (as discussed in Section 6.2). The associated SW heating increases atmospheric temperatures by up to 0.5 K over South Asia. This warming begins in the lower troposphere at short lead times and propagates to the upper troposphere by $T + 120$ due to the resultant increase in convection.
Figure 20a shows the $T + 120$ mean zonal wind at 600 hPa over West Africa from the CNTRL simulation and highlights a strong African easterly jet (AEJ) between the equator and 20° N, with maximum intensity between 10 and 15° N. The mean $T + 120$ zonal wind bias, shown in Fig. 20b illustrates that the AEJ is located too far south and the maximum intensity is also too strong off the West African coast. The impact of the different aerosol representations on the AEJ is shown in Fig. 20b–d. Dust is the primary contributing aerosol species over West Africa and as the indirect effects of mineral dust are not currently included in the MetUM we present the DIR-only simulations here in order to examine the impacts of the direct radiative effect of mineral dust on the circulation in this region. The CLIM simulation has the largest impact on the AEJ, weakening both the southern and northern flanks of the jet. The impacts are smaller in the prognostic simulations particularly along the southern edge of the jet. However, a small enhancement of the jet maximum is found and a beneficial northward shift is also seen in the INIT_DIR simulation.

Figure 21 presents the impact of aerosol on the vertical temperature profile averaged between 15° W and 10° E with dust and biomass burning aerosol concentrations overlaid as contours. All simulations show a significant warming of up to 1 K within the dust layer which is concentrated close to the surface and highlights that most of the region north of 15° N is an active dust source region in the model. The warming is larger and more regionally extensive in the CLIM simulations owing to the much higher localised dust concentrations in the northwest of the region between 20 and 35° N. The warming begins near the surface (below 900 hPa) at short forecast lead times and propagates to above the boundary layer by $T + 120$. This results in an enhanced heat low over Northwest Africa in CLIM as well as an increase in the upper level convergence decreasing the strength of the AEJ at 600 hPa. In AER_DIR, which has a smaller dust loading than CLIM, the warming is more confined further south than CLIM between 17 and 22° N. The upper level warming is also smaller than CLIM and as a result a reduced impact on the AEJ is found. The INIT simulations, as already discussed in Sect. 5, have a much smaller dust loading as well as a different size distribution to CLIM and AER. This leads to enhanced scattering above the dust layer and a subsequent cooling. By $T + 120$ the impact of the aerosol initialisation has reduced and the INIT dust more closely resembles the AER dust distribution than the GEMS dust forecasts. However,
overall concentrations early in the forecast are much lower, leading to a smaller warming near the surface and an elevated cooling which persists through the forecast. This leads to a smaller mean impact on the jet structure north of 15° N.

The impact of dust on the circulation over Africa is extremely complex and very much dependent on the altitude of the dust layer, the dust size distribution and dust optical properties employed in the model. Elevated dust layers tend to cool the surface below by reducing the incoming SW radiation reaching the surface while warming within the dust layer itself. This thermal dipole effect tends to stabilise the atmosphere during the day and destabilise during the night-time through LW warming (Reale et al., 2011; Lavaysse et al., 2011). Many of these studies report a beneficial enhancement of the jet with the inclusion of a representative dust distribution. In the simulations presented here most of the dust is concentrated close to the surface and the thermal dipole impact is not observed. Consequently the resulting impact on the jet structure is not as large as in other studies. The simulated impact of dust on West African circulation and therefore on the West African monsoon is also highly dependent on the absorption properties of the dust (Zhao et al., 2011; Solmon et al., 2008). The Balkanski et al. (2007) dust optical properties used here are less absorbing than other dust optical properties reported in the literature and used in modelling studies (see Ryder et al., 2013b; Zhao et al., 2011; Sokolik et al., 1993). This will also reduce the dust direct radiative effect in the present simulations. Dust optical properties vary depending on the particle size, chemical composition and distance from source (Ryder et al., 2013a). Using a single set of dust optical properties to model the global radiative impacts is therefore a known limitation in global high-resolution dust modelling but is currently necessary given the computational requirements of using regionally varying optical properties, not to mention the uncertainties in the characterisation of dust physical and optical properties.

Fig. 17. Impact of different aerosol representations on MSLP and 925 hPa winds at T + 120: (a) CNTRL, (b) CNTRL error, (c) AER_CLIM-CNTRL, (d) AER_DIR_INDIR-CNTRL, (e) INIT_DIR-CNTRL, (f) INIT_DIR_INDIR-CNTRL.
7 Discussion

The aim of this paper is to establish the potential benefits from increasing the aerosol complexity on global model forecasting skill on NWP timescales. Here we summarise our findings and discuss the benefits gained from each level of additional aerosol complexity.

The first level of increased aerosol complexity involves the use of monthly mean aerosol climatologies. This is what is currently implemented operationally in the global NWP configuration of the MetUM and in the Met Office seasonal forecasting system, GLOSEA (Arribas et al., 2011). One of the main attractions of using aerosol climatologies is that they do not increase the operational running costs of the model, a key requirement in any operational high-resolution global forecasting system. This study shows that the climatology gives a reasonable representation of mean aerosol loading but is unable to capture the large temporal variation in AOD. This is not surprising given the highly inhomogeneous nature of aerosol distributions (tropospheric aerosols have average lifetimes ranging from a couple of days to ~1 week) and that the climatologies bear little resemblance to the driving model’s atmospheric state. The impact of including the direct effect from the aerosol climatologies on the model forecast
skill is found to be generally small or neutral. The largest impacts are found predominantly in dust-dominated regions such as the Sahara Desert, where the large dust loading enhances the heat low over Northwest Africa and through the semi-direct effect reduces cloud over tropical Atlantic. However, the comparison of model OLR against GERB observations in this study suggests that the dust climatology leads to a significant reduction in the OLR over West Africa resulting in a negative bias of up to 30 W m$^{-2}$. This is most likely due to inaccuracies in the representation of mineral dust in the climatology.

The next level of aerosol complexity involves coupling a fully prognostic aerosol scheme to the global NWP model. The additional tracers required by prognostic aerosol schemes obviously lead to a substantial increase in the running costs of the model. However, this study shows how the use of prognostic aerosol schemes in the AER and INIT simulations gives a much better temporal and spatial distribution of the aerosol. Initialising the aerosol with GEMS assimilated aerosol fields further improves aerosol forecasts particularly in complex aerosol regions such as South Asia (see Fig. 19d and e). There are uncertainties in the aerosol emissions and processes as well as the aerosol optical properties used in these simulations. For instance, nitrate aerosol is a notable omission in the simulations carried out here. Nitrate aerosol is also not included in the GEMS aerosol model (Morcrette et al., 2009). Therefore, even though the aerosol assimilation improves total AOD comparisons against AERONET over India for example, the speciation and subsequent aerosol optical properties used could be in error. One key area where the initialisation does not improve aerosol forecasts is in dust-dominated regions. This is due to an excess of sub-micron size dust particles in the GEMS dust scheme. Using the improved MACC aerosol forecasts in future studies, which include a revised dust size distribution, improved aerosol assimilation and consequently a smaller AOD bias is expected to lead to significant improvements in the initialised aerosol forecasts. Apart from this, the overall
Fig. 20. Impact on 600 hPa zonal wind over West Africa at $T+120$: (a) CNTRL U wind, (b) CNTRL error, (c) CLIM-CNTRL, (d) AER_DIR-CNTRL, (e) INIT_DIR-CNTRL.

success of the initialised simulations highlights the skill of the GEMS and subsequent MACC assimilated aerosol forecasts. Overall, this increased level of aerosol complexity has an important benefit as it enables the development of a global environmental prediction capability. The potential use of the global NWP configuration of the MetUM to predict significant aerosol events such as large dust storms, volcanic ash events and an increasing number of wild fire episodes would be extremely important due to the high impact these events have on daily lives and health of the general public. Operational dust forecasts are now available from the operational global NWP configuration of the MetUM at a 25 km resolution. These dust forecasts use a simplified two-bin version of the Woodward (2001) scheme and include a 4D-Var data assimilation of dust AOD derived from MODIS observations. They are used routinely for military planning purposes and
are also provided daily to the World Meteorological Organisation Sand and Dust Storm Warning Advisory and Assessment System (WMO SDS-WAS) to provide dust forecasts for North Africa, Europe and the Middle East. Routine evaluation of the operational dust forecasts against AERONET observations as well as other operational dust models highlights the good skill of the Met Office dust forecasts (Terradellas et al., 2013). Further extending this capability to include other aerosol species would provide additional benefits for global environmental prediction.

In this study, the impact of including the direct effect from prognostic aerosols on the mean forecast skill is comparable to including the direct effect from the aerosol climatologies, with impacts found to be small or neutral. However, larger regional impacts are found primarily in dust-dominated regions. Due to the better dust representation over North Africa in the AER simulations compared to both the CLIM and INIT simulations, significant beneficial improvements are found in the model OLR biases over West Africa. Indeed, the very different model responses to the different dust representations shown in Fig. 21 highlight that despite numerous advances made in characterising the physical and optical properties of mineral dust aerosol in a number of aircraft and ground-based measurement campaigns (Haywood et al., 2011, 2003a; Johnson et al., 2011) large uncertainties still remain. Further constraining these uncertainties is key for an accurate representation of the dust radiative impacts in both NWP and climate models. It is also likely that the highly variable nature of biomass burning emissions due to forest fires will have strong regional short-term impacts on forecast skill. The study period analysed here is outside the primary biomass burning seasons in Amazonia and central Africa but these impacts are being evaluated in on-going separate studies.

Inclusion of the indirect aerosol effects from the prognostic aerosols have large global impacts on cloud formation and planetary albedo. The large reduction in low cloud amounts in high-latitude, clean-air regions improves model radiation and temperature biases. Similar large impacts found in South Asia are dominated by the large sensitivities to the change in CDNC rather than due to the direct absorption by the prevalent haze layers in this region. The role of absorbing aerosol in controlling the strength and variability of the South Asian monsoon has been hypothesised in a number of studies (e.g.: Lau et al., 2006; Ramanathan et al., 2001). The negligible impacts found in this study could be due to either missing aerosol sources in this region or unrepresentative optical properties for these absorbing aerosol species. The optical properties of both dust and biomass burning aerosol are based on aircraft measurements made over Africa (Balkanski et al., 2007; Haywood et al., 2003b) and might not be appropriate for this region. While investigations into these issues are beyond the scope of the current work, it is clear that a consensus on aerosol impacts on the South Asian monsoon is far from being reached and further research is required.
Large sensitivities to the change in CDNC from fixed values used in the current operational NWP system to values diagnosed from the CLASSIC aerosol scheme highlight gross errors in the current simple representation of aerosol indirect effects in the model. Consequently, as a first step to improving aerosol–cloud–radiation interactions in the global NWP model, inclusion of the indirect effects from the aerosol climatologies has been evaluated and implemented in the most recent global configuration of the MetUM. One of the key benefits of the MetUM is that the same physical model is used from short-range NWP to climate timescale predictions. Inclusion of both the direct and indirect effects from speciated aerosol climatologies derived from the CLASSIC scheme means that for the first time there is a consistent treatment of aerosol processes across the Met Office configuration of models from NWP to climate, an objective in the development of a seamless prediction system in the Met Office (Senior et al., 2010). Implementing a fully prognostic aerosol scheme in the operational global NWP configuration of the MetUM would even further benefit these seamless prediction aims. Furthermore, aerosol forecasts produced using short-range high-resolution forecasting systems can be more easily evaluated against a wide range of near-real-time aerosol observations. Findings from such routine evaluations would feedback into aerosol model development and lead to improved aerosol predictions on both NWP and climate timescales and subsequently improve our estimates of the direct and indirect aerosol forcing on climate.

8 Conclusions

A number of model simulations have been conducted to investigate the impact of increasing the aerosol complexity in short-range weather forecasts using the global NWP configuration of the Met Office Unified Model. A hierarchy of aerosol representations was evaluated. These included a simple one-dimensional climatology with fixed properties for land and ocean (CNTRL), three-dimensional monthly-varying speciated aerosol climatologies (CLIM), fully interactive aerosols modelled using the CLASSIC aerosol scheme (AER) and finally initialised aerosol simulations using assimilated aerosol fields from the GEMS aerosol forecasting system (INIT). Both direct and indirect radiative effects were evaluated and the impact of each aerosol representation on the NWP model performance was assessed.

Forecasts of AOD from all aerosol simulations were compared against ground-based AERONET and satellite observations. The prognostic aerosol simulations demonstrated a significant improvement in skill compared with CLIM in predicting the spatial and temporal variation of AOD in the short-range 5-day forecasts evaluated. Initialisation with the GEMS aerosol generally resulted in a higher aerosol loading in the NH except for dust source regions. The impact of the initialisation is relatively short and the mean global aerosol loading falls back to that of the model’s own climatology by $T + 120$ h of the forecast. The good predictive skill of the aerosol forecasts produced using the CLASSIC aerosol scheme highlights the potential capability of this system for the prediction of high-impact aerosol events, such as large dust storms or forest fires.

The most notable benefit of including an improved aerosol representation and hence more realistic aerosol–cloud–radiation interactions is the improvement in the radiation budget of the model. A significant improvement in model OLR biases over the desert regions of West Africa are found when mineral dust aerosol is included in the model. Due to the different mineral dust loadings and size distributions used in the different aerosol simulations the response to the direct radiative effect of dust correspondingly varies and the AER simulation gives the best improvement in model OLR bias. Evaluation of the impact on the SW radiation at TOA is more difficult due to the heterogeneity of the underlying surface, but dust generally increases the reflected SW radiation at TOA leading to a slight degradation in the existing SW bias at TOA. The indirect aerosol effect leads to a net TOA cooling in regions where low-level marine cloud has increased, particularly in the stratocumulus regions off the west coast of Africa and South America. A significant warming is found at high latitudes in the NH and to a smaller extent over open ocean regions in the SH. This is caused by sensitivities to the first indirect aerosol effect with a significant reduction in CDNC found in these regions compared to regionally fixed values of CDNC used in the control simulation. This reduction appears to be in much better agreement with observations in clean marine air regions (Birch et al., 2012; Bennartz, 2007) and improves the bias in the surface SW downward radiation at the North Slope of Alaska ARM site. The subsequent warming of the NH leads to a local improvement of $\sim 1$ K in the NH temperature cold bias during the study period. Height biases are also improved in the NH and tropics when the indirect aerosol effects are included in the model. The net TOA perturbation due to including the direct and indirect aerosol effects in the initialised simulations is approximately $2 \text{W m}^{-2}$ smaller than the corresponding uninitialized experiment. This is due to the higher NH aerosol burden and consequently slightly higher CDNC values in the INIT simulation.

Aerosol impacts on the global mean model precipitation and circulation are found to be small in the short-range forecasts evaluated here. However, the direct aerosol effect reduces precipitation by $1–3 \text{mm day}^{-1}$ over the NH sub-tropics where the model has a significant positive bias compared with TRMM observations. All prognostic aerosol simulations weaken the anticyclonic flow over India. Further significant enhancement of the low level monsoon flow over the Indian Ocean and Bay of Bengal is found in simulations including the indirect aerosol effect. The large cloud forcing over the TP and Indo-China regions enhances the heat low over the TP increasing errors in the low-level monsoon flow.
and upper-level tropical easterly jet. These findings are consistent with a number of global climate model studies (Lau et al., 2006; Ramanathan et al., 2001) and further research is required to further our understanding of aerosol properties in this complex aerosol region and their impacts on weather and climate. Representation of aerosol indirect effects remains one of the largest uncertainties in estimates of aerosol forcing on climate (Lohmann et al., 2010; Quaas et al., 2009; Lohmann and Feichter, 2005). Therefore, lack of a coupling between aerosols and convective parametrisation in the present study could lead to inaccuracies in the findings particularly for the tropics. Overall, the impact of the initialised aerosol simulations on the global NWP model performance is smaller than in the initialised aerosol simulations. This is primarily due to a smaller direct radiative forcing from mineral dust and biomass burning over West Africa and a reduced sensitivity to the first indirect effect in clean-air regions due to the higher aerosol loading in the NH.

In a general sense, choosing an appropriate level of aerosol complexity for operational global NWP systems should be governed by the downstream applications and users of the system. It is clear that increasing the complexity of aerosols in a global NWP model to the extent of incorporating a fully interactive prognostic aerosol scheme has large benefits for developing a global environmental prediction capability within the Met Office. Furthermore, it enables a seamless treatment of aerosols across the Met Office configuration of models from high-resolution short-range NWP forecasts to long-range climate predictions. While impacts on the global mean skill of the forecast are small, aerosol impacts on these timescales are likely to be limited by the sophisticated data assimilation systems used in NWP systems. However, large regional benefits in radiation and temperature forecasts highlight for the first time the importance of including a realistic treatment of aerosol–cloud interactions in global NWP models.

Acknowledgements. This study was funded by the European Commission under the EU Seventh Research Framework Programme (grant agreement No. 218793, MACC). The authors would like to thank the following people and projects for providing data and assistance for this study: Angela Benedetti, Jean-Jacques Morcrette and Miha Razinger at the ECMWF for assistance with the GEMS data assimilation system used in NWP systems. However, large regional benefits in radiation and temperature forecasts highlight for the first time the importance of including a realistic treatment of aerosol–cloud interactions in global NWP models.

References


