Reviews and syntheses: the GESAMP atmospheric iron deposition model intercomparison study

Stelios Myriokefalitakis1,a, Akinori Ito2, Maria Kanakidou3, Athanasios Nenes4, Maarten C. Krol1,
Natalie M. Mahowald5, Rachel A. Scanza5, Douglas S. Hamilton5, Matthew S. Johnson6, Nicholas Meskhidze7,
Jasper F. Kok8, Cecilie Guieu9, Alex R. Baker10, Timothy D. Jickells10, Mannohan M. Sarin11, Srinivas Bikkina11,
Rachel Shelley12, Andrew Bowie13,14, Morgane M. G. Perron13, and Robert A. Duce15

1Institute for Marine and Atmospheric Research (IMAU), Utrecht University, 3584 CC Utrecht, the Netherlands
2Yokohama Institute for Earth Sciences, JAMSTEC, Yokohama, Kanagawa 236-0001, Japan
3Environmental Chemical Processes Laboratory (ECPL), Department of Chemistry, University of Crete, 70013 Heraklion, Greece
4School of Earth & Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA 30332, USA
5Department of Earth and Atmospheric Sciences, Cornell University, Ithaca, NY 14853, USA
6Earth Science Division, NASA Ames Research Center, Moffett Field, CA 94035, USA
7Marine, Earth, and Atmospheric Sciences, North Carolina State University, Raleigh, NC 27695, USA
8Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, CA 90095, USA
9Laboratoire d’Océanographie de Villefranche (LOV), UMR7093, CNRS-INSU-Université Paris 6, Villefranche sur Mer, France
10Centre for Ocean and Atmospheric Sciences, School of Environmental Sciences, University of East Anglia, Norwich, UK
11Geosciences Division, Physical Research Laboratory, Ahmedabad, India
12Earth, Ocean and Atmospheric Science, Florida State University, Tallahassee, Florida 32306-4320, USA
13Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania, Australia
14Antarctic Climate & Ecosystems CRC, Hobart, Tasmania, Australia
15Departments of Oceanography and Atmospheric Sciences, Texas A&M University, College Station, TX 77843, USA
a now at: Institute for Environmental Research & Sustainable Development (IERSD), National Observatory of Athens, 15236 Palea Penteli, Greece

Correspondence: Stelios Myriokefalitakis (steliosm@noa.gr) and Akinori Ito (akinori@jamstec.go.jp)

Received: 12 June 2018 – Discussion started: 10 July 2018
Revised: 3 October 2018 – Accepted: 18 October 2018 – Published: 9 November 2018

Abstract. This work reports on the current status of the global modeling of iron (Fe) deposition fluxes and atmospheric concentrations and the analyses of the differences between models, as well as between models and observations. A total of four global 3-D chemistry transport (CTMs) and general circulation (GCMs) models participated in this intercomparison, in the framework of the United Nations Joint Group of Experts on the Scientific Aspects of Marine Environmental Protection (GESAMP) Working Group 38, “The Atmospheric Input of Chemicals to the Ocean”. The global total Fe (TFe) emission strength in the models is equal to \( \sim 72 \text{Tg Fe yr}^{-1} \) (38–134 Tg Fe yr\(^{-1}\)) from mineral dust sources and around \( 1.8–2.7 \text{Tg Fe yr}^{-1} \) (1.8–2.7 Tg Fe yr\(^{-1}\)) from combustion processes (the sum of anthropogenic combustion/biomass burning and wildfires). The mean global labile Fe (LFe) source strength in the models, considering both the primary emissions and the atmospheric processing, is calculated to be 0.7 (±0.3) Tg Fe yr\(^{-1}\), accounting for both mineral dust and combustion aerosols. The mean global deposition fluxes into the global ocean are estimated to be in the range of 10–30 and 0.2–0.4 Tg Fe yr\(^{-1}\) for TFe and LFe, respectively, which roughly corresponds to a respective 15 and 0.3 Tg Fe yr\(^{-1}\) for the multi-model ensemble model mean.
The model intercomparison analysis indicates that the representation of the atmospheric Fe cycle varies among models, in terms of both the magnitude of natural and combustion Fe emissions as well as the complexity of atmospheric processing parameterizations of Fe-containing aerosols. The model comparison with aerosol Fe observations over oceanic regions indicates that most models overestimate surface level TFe mass concentrations near dust source regions and tend to underestimate the low concentrations observed in remote ocean regions. All models are able to simulate the tendency of higher Fe concentrations near and downwind from the dust source regions, with the mean normalized bias for the Northern Hemisphere (∼14), larger than that of the Southern Hemisphere (∼2.4) for the ensemble model mean. This model intercomparison and model–observation comparison study reveals two critical issues in LFe simulations that require further exploration: (1) the Fe-containing aerosol size distribution and (2) the relative contribution of dust and combustion sources of Fe to labile Fe in atmospheric aerosols over the remote oceanic regions.

1 Introduction

Oceans are important for the Earth system’s functioning, currently absorbing roughly 27 % of total CO₂ emissions (e.g., Le Quéré et al., 2013), providing half of atmospheric oxygen and being a source of biomass that helps sustain life on our planet. Iron (Fe) is a key element for marine life (Duce and Tindale, 1991; Fung et al., 2000) and is required for photosynthesis and respiration. As an essential micronutrient, Fe (co-)limits ocean productivity over large regions (Boyd et al., 2005; Jickells et al., 2005; Martin et al., 1991; Moore et al., 2013), influences the nitrogen fixation capability of diazotrophs in oligotrophic regions (Falkowski, 1997; Falkowski et al., 2000) and generally affects the transport and sequestration of carbon into the deep ocean (Maher et al., 2010). Atmospheric deposition is considered to be an important external Fe source for the open ocean (Jickells et al., 2005; Tagliabue et al., 2017). Micronutrient Fe delivered via atmospheric pathways may influence the primary and export production of carbon over the high-nutrient low-chlorophyll (HNLC) oceanic regions (i.e., the oceanic regions where Fe is the limiting factor for phytoplankton productivity). However, significant Fe inputs from continental margins and hydrothermal vents are also supplied to the global ocean, regulating the ocean biogeochemical cycles. Moreover, riverine Fe inputs are currently estimated to be 1–2 orders of magnitude smaller that the atmospheric pathway (e.g., Tagliabue et al., 2016), affecting mainly coastal regions, while icebergs and glaciers could also be important for the polar oceans (Raiswell et al., 2016).

An understanding of the impact of Fe on global marine productivity requires a knowledge of the rates and locations of Fe supply to the ocean, and of the physicochemical forms of Fe that can be utilized by marine biota (i.e., those that are bioavailable). The bioavailability of Fe is a complex issue (e.g., Lis et al., 2015; Morel et al., 2008) and several naming conventions and abbreviations have been used to characterize the atmospheric supply of potentially bioavailable Fe to the global ocean (Baker and Croot, 2010; Shi et al., 2012). It has been widely assumed that soluble Fe can be considered, as a first approximation, to be bioavailable (Baker et al., 2006a, b). Therefore, a common experimental practice to determine the bioavailable Fe fraction in Fe-containing aerosols is the quantification of Fe in a leachate solution that passes through a 0.45, 0.2 or 0.02 µm sized filter (see Meskhidze et al., 2016, and references therein). However, due to its operational definition, it has been shown that this filterable Fe may contain both soluble and colloidal forms of Fe (Jickells and Spokes, 2001; Raiswell and Canfield, 2012). Upon deposition to the surface ocean, the soluble form of Fe delivered through atmospheric pathways can either enter the dissolved Fe pool or precipitate out as large (oxy)hydroxide particles (de Baar and de Jong, 2001; Boyd and Ellwood, 2010; Meskhidze et al., 2017; Turner and Hunter, 2001). Consequently, the impact of atmospheric Fe on marine biogeochemistry depends on both the total Fe (TFe) deposition and its solubility, keeping in mind that the bioavailable fraction of Fe in seawater will also change due to post-atmospheric deposition ocean processes (e.g., Baker and Croot, 2010; Chen and Siefert, 2004; Meskhidze et al., 2017; Rich and Morel, 1990).

On the global scale mineral dust is the dominant source of Fe to the atmosphere (∼95 %; Mahowald et al., 2009). The average Fe content in upper crustal minerals is 3.5 % (e.g., Duce and Tindale, 1991), but this can vary considerably, depending on the underlying mineralogy (and geography) of the dust source (Journet et al., 2014; Nickovic et al., 2012, 2013). According to Journet et al. (2014), the Fe content of various minerals is as follows: hematite (69.9 %), goethite (62.8 %), chloride (12.3 %), vermiculite (6.71 %), illite (4.3 %), smectite (2.6 %) and kaolinite (0.23 %) in clay- and silt-sized soil particles (i.e., soil particles with diameters < 2 and 2–50 µm, respectively). Fe-containing aerosols also originate from wildfires and biomass burning (Guieu et al., 2005; Mahowald et al., 2005; Oakes et al., 2012; Paris et al., 2010) and anthropogenic combustion processes, such as coal and oil fly ash (Luo et al., 2008), including ship oil combustion (Ito, 2013). Other sources of Fe oxides can be also identified in the atmosphere, which are attributed to volcanic eruptions (Benitez-Nelson et al., 2003; Langmann et al., 2010) and to a lesser extent meteors (Johnson, 2001).

The use of global biogeochemical numerical models and surface observations is an excellent way to better understand the past, present and future atmospheric supply to the oceans, as well as to quantify the resultant effect on the ocean biological productivity and the carbon uptake. Modeling the atmospheric supply of soluble Fe to the global ocean is chal-
lenging, due to the multitude and complexity of the forms under which Fe can be present in aerosol emitted to the atmosphere (Meskhidze et al., 2017), as well as the variety and complexity of processes which alter the solubility of Fe during its transport through the atmosphere (Baker and Croot, 2010). Indeed, the soluble fraction of Fe in atmospheric aerosols may include different Fe forms in the ferric oxidation state (Fe(III)) (Fu et al., 2012), ferrihydrite and amorphous precipitates, Fe-oxide nanoparticles (Shi et al., 2009), Fe-organic complexes (Cheize et al., 2012) and Fe in the ferrous oxidation state (Fe(II)) (Raiswell and Canfield, 2012). The atmospheric modeling community is mostly focused on the soluble fraction of the deposited Fe over the oceans and for this work the general term labile Fe (LFe) is used to represent the overall soluble Fe in simulated atmospheric aerosol. Both the TFe and LFe atmospheric deposition can be used in ocean biogeochemical modeling. For example, total Fe is used for the comparison of particulate Fe with the measurements in ocean biogeochemistry models (e.g., Ye and Völker, 2017), while LFe can be assumed to be readily available to the marine ecosystem. Note that the less labile fraction of TFe can be slowly dissolved from particulate Fe in the ocean during the sinking of mineral particles (e.g., roughly 0.01% per day; Bonnet, 2004); however, the dissolution of Fe is species dependant and affected by spatiotemporal variations in the ocean.

During recent decades, intensive research has been carried out to elucidate the origin, nature and magnitude of LFe fluxes to the surface ocean. Soils may include a small fraction of LFe – roughly 0.1% (e.g., Ito and Shi, 2016) – considered as impurities attached to minerals such as illite, smectite, kaolinite and feldspars (e.g., Ito and Xu, 2014). Fe-containing fly ash has been observed to be present as ferric sulfate salts or nanoparticulate Fe and is highly soluble (Fu et al., 2012; Schroth et al., 2009), as it is mainly formed via high-temperature combustion followed by sulfuric acid condensation (Sippula et al., 2009). However, the form and the chemical properties of Fe in emissions can vary substantially for each combustion source (Ito, 2013; Wang et al., 2015), with the initial soluble fraction in combustion emissions estimated to be about 77%–81% in oil fly ash (Schroth et al., 2009), 20%–25% in coal fly ash (Chen et al., 2012) and 18%–46% in biomass fly ash (Bowie et al., 2009; Oakes et al., 2012). Recently, Matsui et al. (2018) suggested, based on observed magnetite concentrations, that emissions of anthropogenic combustion Fe in global models could be significantly underestimated, and that the atmospheric burden of Fe is potentially up to 8 times greater than previous estimates have suggested (Luo et al., 2008). LFe can be also formed in the atmosphere during the atmospheric processing of mineral dust and combustion aerosols (Ito, 2012; Ito and Feng, 2010; Johnson and Meskhidze, 2013; Meskhidze et al., 2005; Myriokefalitakis et al., 2015). We use the general term “solubilization” here to describe the process that converts Fe from relatively insoluble minerals to soluble Fe during atmospheric transport and photochemical transformation in the aqueous solution of aerosols and clouds.

Iron solubility (i.e., the fraction of total Fe that is soluble) in atmospheric aerosols over the Atlantic and Pacific oceans has been observed to be in the range of 0.1%–67% during oceanographic cruises (Baker et al., 2006a; Furutani et al., 2010), with even higher solubilities (up to 80%) measured in precipitation samples in the Southern Ocean (Heimburger et al., 2013). During atmospheric transport, coating of Fe-containing dust particles by acidic compounds (e.g., sulfates and nitrates) increases the Fe solubility. When this process is taken into account in model simulations (e.g., Meskhidze et al., 2005) it aids in explaining the observations. Indeed, measurements of fresh dust particles present low (<1%) initial solubilities (Chuang et al., 2005; Fung et al., 2000; Hand et al., 2004; Sedwick et al., 2007), while high aerosol solubilities are commonly observed at lower dust concentrations far from sources (Baker and Jickells, 2006; Sholkovitz et al., 2012; Oakes et al., 2012). Atmospheric processing of dust (Kumar et al., 2010; Meskhidze et al., 2003; Srinivas et al., 2014) is considered to be the best candidate for explaining these observations. These processes may also alter the global pattern of LFe deposition (Fan et al., 2004), especially within remote regions, such as the Atlantic, the Pacific (e.g., Sedwick et al., 2007) and the Southern Ocean (Ito and Kok, 2017; Johnson et al., 2010, 2011).

There is clear experimental evidence that atmospheric acidity – which is mainly driven by air pollution over highly populated regions especially over the Northern Hemisphere (e.g., Seinfeld and Pandis, 2006) as well as natural sources such as volcanic sulfur emissions and the oceanic emissions of dimethylsulfide (DMS) in relatively pristine ecosystems (e.g., Benitez-Nelson et al., 2003) – increases the dust solubility. Laboratory studies indicate that Fe solubilization from minerals under acidic conditions in aerosol or rain droplets (Brandt et al., 2003; Shi et al., 2011; Spokes et al., 1994) occurs on different timescales: from hours to weeks depending on the size and the type of the Fe-containing minerals (Shi et al., 2011), with amorphous and ultrafine Fe solubilized much faster in acidic solutions (Brandt et al., 2003) compared to the aluminosilicates (roughly 10–14 days). Other laboratory studies also support the occurrence of photoinduced reductive Fe solubilization under acidic conditions (e.g., Fu et al., 2010), a mechanism that involves electron transfer to Fe(III) atoms on the particle surface to produce Fe(II) (Larsen and Postma, 2001). The reductive solubilization of minerals that are rich in Fe is also observed to be accelerated in the presence of Fe(II) or Fe(II)-ligand complexes (Litter et al., 1994). The oxalate-promoted solubilization (e.g., Paris et al., 2011) is controlled by the breaking of Fe–O bonds at the mineral’s surface due to the formation of a mononuclear bidentate ligand containing surface Fe (Yoon et al., 2004), with the solubilization rate significantly increased as the pH decreases. Luo et al. (2010) further prescribed pH and oxalate/hematite ratio dependent solubilization rates for mineral dust, based
on the laboratory experiments of Xu and Gao (2008). For weakly acidic conditions (pH of 4.7) and various oxalate concentrations, a positive linear correlation between oxalate concentrations and the released LFe from different minerals has also been observed (Paris et al., 2011; Paris and Desboeufs, 2013). Laboratory investigations (Chen and Grassian, 2013; Siffert and Sulzberger, 1991) have indicated that even under highly acidic solutions (pH of 2–3), oxalic acid can be more important for the Fe solubilization process of dust and combustion aerosols than sulfuric acid through the formation of Fe(III)–oxalate complexes. Thus, the minerals’ solubilization mainly depends on the proton concentration, the mineral surface concentration of organic ligands (such as oxalate), the sunlight and the ambient temperature (e.g., Hamer et al., 2003; Lanzl et al., 2012; Lasaga et al., 1994; Zhu et al., 1993).

The first modeling efforts that took the mixing of mineral dust with such anthropogenic acidic trace gases like sulfur dioxide (SO₂) into account (Fan et al., 2004; Meskhidze et al., 2003, 2005; Solmon et al., 2009) showed considerable enhancements of atmospheric soluble Fe concentrations. A review by Mahowald et al. (2009) pointed out that human activity may have significantly modified the soluble Fe oceanic deposition flux, because anthropogenic combustion processes increased both Fe emissions and the acidity of atmospheric aerosols. Furthermore, recent studies (Meskhidze et al., 2017; Tagliabue et al., 2017) have also shown that atmospheric and oceanic organic ligands may increase the Fe solubilization in the atmosphere and in the ocean, by forming Fe complexes that further increase Fe bioavailability for the marine ecosystems. State-of-the-art global models clearly indicate a strong spatial and temporal variability of the atmospheric LFe supply to the global ocean, which can be partly attributed to atmospheric processing. The global LFe deposition flux is currently estimated to be in the range of 0.4–1.1 TgFe yr⁻¹ (Ito and Kok, 2017; Ito and Shi, 2016; Ito and Xu, 2014; Johnson and Meskhidze, 2013; Luo et al., 2008; Myriokefalitakis et al., 2015; Wang et al., 2015).

In order to constrain a global picture of the influence of the present atmospheric composition on the Fe supply to the oceans, we perform a systematic comparison between models and between models and observations. We identify possible similarities and differences among models and between models and observations. The goals of the present study are to (1) quantify the magnitude of the atmospheric TFe and LFe fluxes to the global ocean as calculated by four state-of-the-art global atmospheric aerosol models, (2) explain the differences in the simulated LFe among the participating models, and (3) to provide multi-model ensemble TFe and LFe atmospheric Fe deposition fluxes for the next generation of ocean biogeochemistry modeling studies. Overall, the importance of this work lies in an extended review and synthesis of the current knowledge of global atmospheric Fe deposition fluxes in the ocean, aiming to provide ensemble model data to the scientific community, which will be able to be used in ocean biogeochemistry models and as comparative measures for atmospheric models.

The following discussion is organized into four sections. Section 2 describes the participating models and the observations used in this study. This section aims to build a concise view of the present-day understanding on the magnitude and the distribution of the TFe and LFe simulated deposition fluxes to the global ocean; ensemble model calculations are also presented. Section 3 presents and discusses the simulated global Fe atmospheric budgets and distributions. In Sect. 4, the uncertainties in the calculated surface aerosol Fe concentrations and deposition fluxes are discussed and the potential model biases are analyzed by attributing them to their major contributors. Finally, in Sect. 5 the findings of the present study are summarized, and recommendations for future research directions are put forward.

2 Methods
2.1 Description of models
The global models participating in this study differ with respect to their spatial horizontal and vertical resolution, the meteorology, the emissions used for gas and aerosol species, and the aerosol microphysics (i.e., size distribution and refractive properties). They also differ in the gas- and aqueous-phase chemical schemes and the parameterizations of atmospheric transport and deposition processes. The main characteristics of the participating models are summarized in Table 1. Note, however, that for this intercomparison no requirements regarding specific year, meteorological conditions or emission inventories were set for the model simulations. Therefore, the data presented are mainly based on earlier published (or soon to be published) modeling experiments, which are evaluated and systematically analyzed here.

1. The Community Atmosphere Model version 4 (CAM4) is embedded within the National Center for Atmospheric Research (NCAR) Community Earth System Model version 1.0.5 (CESM 1.0.5; Hurrell et al., 2013). The CAM4 simulations are conducted with a horizontal resolution of 2.0° × 1.9° (longitude × latitude) and 56 vertical layers up to 2 hPa; it is forced by NASA’s Goddard Earth Observing System (GEOS-5) meteorology. The emission data sets for anthropogenic activities, such as fossil fuel and biofuel combustion, are taken from the Aerosol Comparison between Observations and Models (AeroCom) database (Dentener et al., 2006). Desert dust is modeled following the Dust Entrainment and Deposition (DEAD) module (Zender et al., 2003) with updates of the size fractions (Kok, 2011) and optics as described in Albani et al. (2014). The bin widths are prescribed at diameters of 0.1–1.0, 1.0–2.5, 2.5–5.0 and 5.0–10.0 µm and have fixed lognormal sub-bin distributions. Dust in CAM4 is specified into six minerals, clays (illite, kaol-
Table 1. General description of the participating models used for the atmospheric Fe simulations. For multiple year simulations, the average was used.

<table>
<thead>
<tr>
<th>Model</th>
<th>Simulated year(s)</th>
<th>Horizontal resolution (long × lat)</th>
<th>Vertical resolution (sigma levels)</th>
<th>Meteorology</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAM4</td>
<td>2007–2011</td>
<td>1.9° × 2.5°</td>
<td>56</td>
<td>GEOS-5</td>
<td>Scanza et al. (2018)</td>
</tr>
<tr>
<td>GEOS-Chem</td>
<td>1 Mar 2009–28 Feb 2010</td>
<td>2.0° × 2.5°</td>
<td>47</td>
<td>GEOS-5</td>
<td>Johnson and Meskhidze (2013)</td>
</tr>
<tr>
<td>IMPACT</td>
<td>2014</td>
<td>2.0° × 2.5°</td>
<td>59</td>
<td>GEOS-FP</td>
<td>Ito et al. (2018a)</td>
</tr>
<tr>
<td>TM4-ECPL</td>
<td>2008</td>
<td>3.0° × 2.0°</td>
<td>34</td>
<td>ERA-Interim</td>
<td>Myriokefalitakis et al. (2015, 2016)</td>
</tr>
</tbody>
</table>

2. The GEOS-Chem model is driven by assimilated meteorological fields from the Goddard Earth Observing System (GEOS-5) of the NASA Global Modeling Assimilation at a horizontal 2.0° × 2.5° (latitude × longitude) grid resolution and 47 vertical levels up to 0.01 hPa. GEOS-Chem simulates the emissions and chemical transformation of sulfur compounds, carbonaceous aerosols and sea salt, and includes H₂SO₄–HNO₃–NH₃ aerosol thermodynamics solved by the ISORROPIA II thermodynamic model (Fountoukis and Nenes, 2007) coupled to an O₃–NOₓ–hydrocarbon–aerosol chemical mechanism. GEOS-Chem combines the DEAD scheme with the source function used in the Goddard Chemistry Aerosol Radiative and Transport (GOCART) model. Once mineral dust is mobilized from the surface, the model uses four standard dust bins with diameter boundaries of 0.2–2.0, 2.0–3.6, 3.6–6.0 and 6.0–12.0 µm to simulate global dust transport and deposition, emitting 1614 Tg yr⁻¹ of mineral dust globally. Further details regarding the GEOS-Chem model used for this work can be found in Johnson and Meskhidze (2013) and references therein.

3. The Integrated Massively Parallel Atmospheric Chemical Transport (IMPACT) model (Rotman et al., 2004) is also driven by assimilated meteorological fields from the Goddard Earth Observation System – Forward Processing (GEOS–FP) of the NASA Global Modeling and Assimilation Office (Lucchesi, 2017) with a horizontal resolution of 2.0° × 2.5° and 59 vertical layers up to 0.01 hPa. The model simulates the emissions, chemistry, transport and deposition of major aerosol species (Liu et al., 2005) and their precursor gases (Ito et al., 2007). IMPACT takes emissions of primary aerosols and precursor gases of secondary aerosols such as sulfate, nitrate, ammonium and oxalate into account. The emission data sets for anthropogenic activities such as fossil fuel use and biofuel combustion are taken from the Community Emission Data System (CEDS) (Hoesly et al., 2018). Fe-containing combustion and dust aerosols are distributed among four bins in the model, with diameters of < 1.26, 1.26–2.5, 2.5–5 and 5–20 µm, respectively (Ito, 2015; Ito and Feng, 2010). The present-day emission estimates for natural sources as well as combustion aerosols from biomass burning are used together with anthropogenic emissions (Dentener et al., 2006; Ito et al., 2018a). A total dust source of 5070 Tg yr⁻¹ is dynamically calculated by a physically based dust emission scheme (Kok et al., 2014a, b) in the model for the present day (Experiments 3 in Ito and Kok, 2017). The chemical composition of mineral dust and combustion aerosols can change dynamically from that in the originally emitted aerosols due to reactions with gaseous species. Aerosol pH is calculated from the internal particle composition (H²⁺ and H₂O) for each size bin by the thermodynamic equilibrium module (Jacobson, 1999). The aerosol acidity depends on the aerosol types, mineralogy, particle size, meteorological conditions and transport pathway of aerosols (Ito and Feng, 2010; Ito and Xu, 2014; Ito, 2015). A more detailed description of the IMPACT model used for this work can be found in Ito (2015), Ito and Kok (2017), Ito and Shi (2016) and references therein.

4. The TM4-ECPL global chemistry transport model simulates the oxidant (O₃/NOₓ/H₂O/CH₄/CO) chemistry, accounting for non-methane volatile organic compounds, including isoprene, terpenes and aromatics, multiphase chemistry in clouds and aerosol water, as well as all major primary and secondary aerosol components, including sulfate, nitrate and secondary organic aerosols. TM4-ECPL is coupled with the ISORROPIA II thermodynamic model (Fountoukis and Nenes, 2007) and it uses modal size (lognormal) distributions to describe the evolution of fine and coarse aerosols in the atmosphere. Dust emissions, for the present version of the model, are calculated online based on the dust source parameterization of Tegen et al. (2002), as described in Myriokefalitakis et al. (2016); these updated dust source calculations should produce slightly higher (∼7 %) dust emissions of around
1181 Tg yr\(^{-1}\) compared to the modified AeroCom inventory (Dentener et al., 2006) that was taken into account in the previous version of the model (Myriokefalitakis et al., 2015). Dust is emitted in the fine and coarse mode with mass median radii (lognormal standard deviation) of 0.34 \(\mu\)m (1.59) and 1.75 \(\mu\)m (2.00), respectively. Furthermore, in the updated version of the model, the mineral-containing combustion aerosols are emitted with a number mode radius (lognormal standard deviation) of 0.04 \(\mu\)m (1.8) and 0.5 \(\mu\)m (2.0) for the fine and coarse modes, respectively (Dentener et al., 2006; Myriokefalitakis et al., 2016). All aerosol species in the model are subject to hygroscopic growth and removal processes that generally affect the mass median radius. The aerosol hygroscopic growth in the model is treated as a function of ambient relative humidity and the composition of soluble aerosol components and the uptake of water on aerosols change the particle size. The TM4-ECPL model is driven by the ECMWF (European Center for Medium-Range Weather Forecasts) Interim reanalysis project (ERA-Interim) meteorology and has a horizontal resolution of 3.0° × 2.0° (latitude × longitude), 34 hybrid layers from the surface up to 0.1 hPa and a model time step of 30 min. Further details on the TM4-ECPL model used for this study can be found in Myriokefalitakis et al. (2015, 2016) and references therein.

2.1.1 Iron emission parameterizations

The primary Fe sources taken into account by the models can be roughly grouped as (1) mineral dust and (2) combustion sources. Various parameterizations or simplifications of the Fe emissions are adopted by the models, with the most important in the context of this paper being the Fe content and initial Fe solubility in emissions. The mean Fe content in dust emissions, as well as the initial Fe solubility in emissions taken into account by the participating models, are presented in the Supplement (Figs. S1 and S2, respectively). More detail is also given in the following:

1. **Mineral dust emissions:** Mineral-Fe primary sources are derived from the total mineral dust emissions, the fraction of specific Fe-containing minerals in dust emissions and the Fe content of each mineral (Table S1 in the Supplement). CAM4 uses a soil mineralogy map and the Fe content in soils is estimated based on mineralogical content (Claquin et al., 1999; Scanza et al., 2015, 2018; Zhang et al., 2015). GEOS-Chem and TM4-ECPL take the global soil mineralogy data set developed by Nickovic et al. (2012) into account. GEOS-Chem prescribes an initial Fe solubility of 0.45% for the most reactive and poorly crystalline pool of Fe in desert top soils (Fig. S1), based on the synthesis of data from the Saharan and Sahel regions of northern Africa (Shi et al., 2012). The IMPACT model uses the mineralogy map and the Fe content in soils as estimated by Jour- net et al. (2014). All the Fe-containing minerals in the model (i.e., hematite, goethite, illite, smectite, kaolinite, chlorite, vermiculite and feldspars) are considered to be in the clay-sized soils (diameters < 2 \(\mu\)m), with only goethite, chlorite, and feldspars also believed to be present in the silt-sized soils (diameters between 2 and 50 \(\mu\)m; Journet et al., 2014). The Fe content averaged in size bins 1–3 (3.6%) is higher than in the size bin 4 (2.3%). IMPACT applies an initial Fe solubility of 0.1% (Ito and Shi, 2016) to the mineral dust aerosols emitted in the atmosphere (Fig. S1). In TM4-ECPL, the Fe content of the different Fe-containing minerals of dust (i.e., illite, kaolinite, smectite, goethite and hematite, and feldspars) is based on the recommenda- tions of Nickovic et al. (2013) and the assumption that it is equally distributed between clay- and silt-sized soils, while for GEOS-Chem the Fe content of mineral dust is set to the widely accepted global mean value of 3.5% (Duce and Tindale, 1991). The initial solubility of the emitted Fe-containing dust particles in TM4-ECPL is prescribed as 4.3% on kaolinite and 3% on feldspars emissions (Ito and Xu, 2014), while other minerals are considered to be emitted containing only insoluble Fe. The resulting annual global mean TFe content of emitted dust particles in TM4-ECPL is calculated to be 3.2% on average (Fig. S2).

2. **Combustion emissions:** All models but one (GEOS-Chem) include Fe-combustion emissions. These are considered to be emitted from different combustion sec- tors with various initial Fe solubilities, with the most important Fe-combustion emissions believed to be those from biomass burning, coal and oil combustion (Ta- ble S2). The CAM4 simulation includes the combustion Fe sources derived from industry, biofuels (e.g., residential heating) and fires (the sum of wildfires and anthropogenic biomass burning), as described in Luo et al. (2008), with the assumptions that 4% is soluble at emission and that atmospheric processing occurs as for dust. Shipping Fe emissions are not currently represented within CAM4. IMPACT takes Fe emissions from biomass burning, coal combustion and oil combustion into account (Ito et al., 2018a), while an initial Fe solubility (58 ± 22%) is only applied to the primary Fe emission of ship oil combustion aerosols (Ito, 2015) assuming that other Fe combustion emission sectors are insoluble. TM4-ECPL takes Fe emissions from biomass burning, coal combustion and oil combustion into account, based on the recommendations of Luo et al. (2008) for biomass burning and coal combustion and of Ito (2013) for oil combustion, assuming fixed Fe- solubilities of 12% for biomass-burning Fe emissions, 8% for coal combustion and 81% for oil combustion from shipping. Note that none of the current models
considered here take volcanic emissions into account, although they may be an important source of LFe to some regions of the ocean (e.g., Duggen et al., 2010).

2.1.2 Iron solubilization parameterizations

The conversion of insoluble-to-soluble Fe in the models can be parameterized as an aqueous-phase kinetic process that depends on (1) the proton activity (also termed acid-promoted solubilization), (2) the oxalate concentration (also termed oxalate-promoted Fe solubilization) and (3) the acetic flux (also termed photo-reductive solubilization). The simplification of the parameterizations applied differs among models; regarding the models used in this study, only IMPACT takes all three solubilization processes into account, but only in aerosol water for both dust and combustion aerosols. TM4-ECPL and GEOS-Chem only apply an acid- and oxalate-solubilization scheme for dust aerosols, in both aerosol and cloud water. However, oxalate is used in models as a proxy of all organic ligands for the ligand-promoted dissolution as (1) it is the most abundant in the atmosphere (e.g., Kawamura and Ikushima, 1993; Kawamura and Sakaguchi, 1999) originating mainly from secondary sources and only a weak contribution from combustion primary sources (e.g., Myriokefalitakis et al., 2011) and (2) it is the most effective ligand in promoting Fe solubilization (e.g., Paris et al., 2011). Nevertheless, we note that more work is required to elucidate the role of other ligands that may promote Fe dissolution in future studies.

CAM4 accounts for the atmospheric processing of both dust and combustion aerosols, based on the acid- (Meskhidze et al., 2005) and oxalate- (Paris et al., 2011) driven solubilization processes in a simplified manner appropriate for use in an Earth system model (Scanza et al., 2018). The method is described in more detail in Scanza et al. (2018), but generally the acid promoted iron dissolution depends explicitly on modeled temperature and an assumed acidity, which is either high (i.e., pH of 2) or low (i.e., pH of 7.5) based on the relative model concentrations of sulfate and calcite; furthermore, the oxalate concentrations in cloud water required for ligand promoted iron dissolution are not explicitly calculated, but are instead assumed to be proportional to the modeled organic carbon aerosol concentration. CAM4 also assumes Fe from dust to be in either a slow, medium or readily soluble state based on Shi et al. (2011) and Ito and Xu (2014), while Fe from combustion is assumed to be in a medium soluble state.

The other three models (GEOS-Chem, IMPACT and TM4-ECPL) calculate the proton-promoted solubilization rate of minerals by applying an empirical parameterization from Meskhidze et al. (2005) and Johnson and Meskhidze (2013), which takes the degree of saturation of the solution, the type of each mineral and the ambient temperature into account. The thermodynamic equilibrium modules are used to estimate the water content in the aqueous phase of hygroscopic particles (Jacobson, 1999; Fountoukis and Nenes, 2007). In addition to the mineral types, IMPACT and TM4-ECPL consider three dust-Fe pools associated with mineral source materials as measured by Ito and Shi (2016) and Shi et al. (2011), respectively, and the solubilization rates calculated by Ito and Shi (2016) and Ito and Xu (2014), respectively. Despite the different mineral databases used by the two models (see Sect. 2.1), the three Fe pools are roughly similarly characterized in the models as ferrihydrite, nano-sized Fe oxides and the heterogeneous inclusion of nano-Fe grains in aluminosilicates, respectively. For GEOS-Chem, the Fe containing mineral (i.e., hematite, goethite and illite) solubilization rate is based on the temperature-dependent equations from Meskhidze et al. (2005) and Johnson and Meskhidze (2013).

For the oxalate-promoted solubilization, CAM4, GEOS-Chem and TM4-ECPL apply a linear relationship between the solubilization rates and the oxalate concentration in the solution, based on the laboratory data from Paris et al. (2011), who measured the initial soluble Fe release rates of Fe-oxides and aluminosilicates (i.e., at a pH of 4.7, and for one hour). The TM4-ECPL model applied this oxalate-solubilization relationship for three Fe-containing minerals (hematite, goethite and illite), using illite as a proxy for all Fe-containing aluminosilicate minerals (Johnson and Meskhidze, 2013). In TM4-ECPL the formation of oxalate in cloud and aerosol water is explicitly simulated in the model (Myriokefalitakis et al., 2011), in contrast to GEOS-Chem in which the sulfate concentrations are used as a proxy for the oxalate production (Yu et al., 2005). However, in TM4-ECPL the oxalate Fe-solubilization is only applied in cloud droplets, where in GEOS-Chem it is applied in both cloud and aerosol water. IMPACT also takes an explicit scheme of oxalate formation both in cloud and aerosol water into account (Lin et al., 2014); however, the oxalate-promoted Fe-solubilization is only applied in aerosol water (Ito, 2015). The constants used to calculate these Fe solubilization rates in IMPACT are fitted to experimental data for coal fly ash (Chen and Grassian, 2013), while the rate of the photoinduced solubilization is based on the Fe-dissolution rates of coal fly ash (Chen and Grassian, 2013), scaled on the photolysis rate of H₂O₂ estimated in the model.

2.1.3 Deposition parameterizations

Dry and wet deposition are considered as loss processes for all Fe-containing aerosols in the models. For this work, the dry deposition fluxes include both the gravitational settling and the turbulent deposition and the wet deposition takes both the in-cloud nucleation scavenging and the below-cloud scavenging in all models into account. For CAM4, the dry removal of dust aerosols involves parameterizations for gravitational settling and turbulent mix out, and wet removal includes in-cloud and below-cloud scavenging (Rasch et al., 2000; Zender et al., 2003). For GEOS-Chem, the re-
The removal of mineral dust occurs through dry deposition processes such as gravitational settling (Seinfeld and Pandis, 2006) and turbulent dry transfer of particles to the surface (Zhang et al., 2001). Dust removal by wet deposition processes includes both convective updraft scavenging and rainfall/removal/washout from large-scale precipitation (Liu et al., 2001).

For IMPACT, the dry deposition of aerosol particles uses a resistance-in-series parameterization (Zhang et al., 2001). Gravitational settling is also taken into account (Rotman et al., 2004; Seinfeld and Pandis, 2006). Aerosols and soluble gases can be incorporated into cloud drops and ice crystals within cloud (rainout), collected by falling rain and snow (washout), and be entrained into wet convective updrafts (Ito et al., 2007; Ito and Kok, 2017; Liu et al., 2001; Rotman et al., 2004). The aging of dust and combustion aerosols from hydrophobic to hydrophilic enhances their dry and wet deposition. Hygroscopic growth of mineral dust and combustion aerosols in gravitational settling uses the Gerber (1991) scheme, including the particle growth due to sulfate, ammonium and nitrate associated with the particles (Liu et al., 2005; Xu and Penner, 2012). Scavenging efficiencies for mineral dust and combustion aerosols in wet deposition are calculated based on the amount of sulfate, ammonium, and nitrate coated on the particles (Liu et al., 2005; Xu and Penner, 2012). For TM4-ECPL, the dry deposition parameterizations are based on an online scheme that takes series of surface and atmospheric resistances into account (Ganzeveld and Lelieveld, 1995). The aerosol hygroscopic growth in the model is treated as a function of ambient relative humidity and the composition of soluble aerosol components (Gerber, 1985) which changes the particle size and impacts the gravitational settling of aerosols. For the wet deposition in TM4-ECPL, both the liquid and ice precipitation are taken into account, with a distinction between scavenging due to large-scale and convective precipitation. In-cloud scavenging in stratiform precipitation uses an altitude-dependent precipitation formation rate, and the scavenging efficiency is calculated taking the aerosols lognormal distributions into account. Note that in TM4-ECPL, all soluble aerosols are assumed to be completely scavenged in the convective updrafts producing rainfall rates of > 1 mm h⁻¹, and are exponentially scaled down for lower rainfall rates.

### 2.2 The ensemble model

Ensemble model calculations in this study generally aim to provide robust results of the simulated atmospheric Fe concentrations and deposition fluxes. For these calculations, all fields for TFe and LFe in mineral dust and combustion aerosols (as well as for dust aerosols) are first converted to a common 1.0° × 1.0° horizontal resolution grid, using the freely available Climate Data Operators (CDO v.1.9.5) software. CDO is a collection of operators for the standard processing of climate and forecast model data developed by the Max Planck Institute for Meteorology. For this work, these operators were applied with a bilinear interpolation to all fields, ensuring an exact mass conservation. Further details about CDO can be found online at https://code.mpimet.mpg.de/projects/cdo/embedded/cdo.pdf, last access: 5 November 2018.

The ensemble atmospheric concentrations and deposition fluxes of mineral dust Fe were calculated from four models, while for the Fe originating from combustion sources, three models were used (see Table 2). For Fe-containing combustion aerosols, we simply used the mean of the respective fields of each model to derive the ensemble model, since no considerable differences appeared among the participating models. However, since model simulations of the global dust cycle are well known to have substantial biases in the size distribution relative to in situ measurements and remote sensing observations (e.g., Huneeus et al., 2011; Kok et al., 2017; Ridley et al., 2016), we attempt to reduce these biases here by correcting the loading and deposition flux in each model’s particle bin, using state-of-the-art constraints on size-resolved dust loading, as recently derived in Kok et al. (2017).

Specifically, a correction factor c_i,j is applied to each particle bin j of model i, which equals

\[
c_{i,j} = \frac{\int_{D_{i,j-}}^{D_{i,j+}} \frac{dM_{\text{atm}}}{dD} dD}{L_{i,j}},
\]

where \(dM_{\text{atm}}/dD\) is the mass size distribution of the global atmospheric dust (see Figs. 2b and S1b in Kok et al., 2017). This mass size distribution was obtained from measurements, modeling and remote sensing constraints on the size distribution of emitted dust, the atmospheric lifetime and extinction efficiency of atmospheric dust, and the global dust aerosol optical depth (Kok et al., 2017; Ridley et al., 2016). Furthermore, \(D_{i,j-}\) and \(D_{i,j+}\) are the respective lower and upper size limits of particle bin \(j\) in model \(i\), and \(L_{i,j}\) is the (not bias corrected) simulated global dust loading in that particle bin. Since emission and deposition fluxes scale with atmospheric dust loading, we correct these fluxes as well as the atmospheric load in each particle bin of each of the contributing models by multiplying the flux by the correction factor in Eq. (1). Note, however, that the bias correction calculated in this work is expected to correct only the part of the regional bias that stems from a bias in the global deposition fluxes. Biases in the regional scales are affected by biases on the global scale, but also by other biases, such as those caused by uncertainties in the deposition scheme. The global mean bias correction factors (i.e., median, lower 95 % confidence interval and upper 95 % confidence interval) for each model and each aerosol size (bin or mode) are presented in Table 3.

For TM4-ECPL, which uses modal size distributions for dust (see Table 2), we also redistributed the fine and coarse aerosols into four bins, in order to apply the same methodology. Specifically, the new aerosol modes are recalculated using the error function and based on the characteristic (radius
Table 2. Iron representation in the models: TFeD – total Fe in mineral dust, TFeC – total Fe in combustion aerosols, LFeD – labile Fe in mineral dust, and LFeC – labile Fe in combustion aerosols.

<table>
<thead>
<tr>
<th>Model</th>
<th>Aerosol species</th>
<th>Aerosol size representation</th>
<th>Soil mineralogy</th>
<th>Atmospheric processing</th>
<th>Aqueous-phase median</th>
<th>Aqueous-phase medium</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAM4</td>
<td>TFeD, TFeC, LFeD, LFeC</td>
<td>Four bins for each Fe type with diameters of 0.1–1.0, 1.0–2.5, 2.5–5.0 and 5.0–10.0 µm</td>
<td>Claquin et al. (1999)</td>
<td>Proton- and ligand-dissolution of dust and combustion aerosols</td>
<td>Aerosol water, cloud droplets</td>
<td></td>
</tr>
<tr>
<td>GEOS-Chem</td>
<td>TFeD, LFeD</td>
<td>Four bins for TFeD with diameters of 0.2–2.0, 2.0–3.6, 3.6–6.0 and 6.0–12.0 µm one bulk for LFeD</td>
<td>Nickovic et al. (2012)</td>
<td>Proton- and ligand-dissolution of dust aerosols</td>
<td>Aerosol water, cloud droplets</td>
<td></td>
</tr>
<tr>
<td>IMPACT</td>
<td>TFeD, TFeC, LFeD, LFeC</td>
<td>Four bins for each Fe type with diameters of &lt;1.26, 1.26–2.5, 2.5–5 and 5–20 µm</td>
<td>Journet et al. (2014)</td>
<td>Proton-, ligand- and pho-toinduced dissolution of dust and combustion aerosols</td>
<td>Aerosol water, cloud droplets</td>
<td></td>
</tr>
<tr>
<td>TM4-ECPL</td>
<td>TFeD, TFeC, LFeD, LFeC</td>
<td>Two modes for each Fe type, with mass median radii (lognormal standard deviation) of 0.34 µm (1.59) and 1.75 µm (2.00) for dust aerosols and 0.04 µm (1.8) and 0.5 µm (2.00) for combustion aerosols</td>
<td>Nickovic et al. (2012)</td>
<td>Proton- and ligand-dissolution of dust aerosols</td>
<td>Aerosol water, cloud droplets</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Global mean bias correction factors (median, lower 95% confidence interval and upper 95% confidence interval) derived for each model and each aerosol size (bins or modes) based on state-of-the-art constraints on size-resolved dust loading (Kok et al., 2017), taken into account for the ensemble model calculations.

<table>
<thead>
<tr>
<th>Model</th>
<th>Aerosol size</th>
<th>Bin 1</th>
<th>Bin 2</th>
<th>Bin 3</th>
<th>Bin 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAM4</td>
<td>Accumulation mode</td>
<td>0.91, 0.89, 0.98</td>
<td>0.55, 0.57, 0.62</td>
<td>1.57, 1.32, 1.46</td>
<td>0.55, 0.57, 1.57</td>
</tr>
<tr>
<td>GEOS-Chem</td>
<td>Accumulation mode</td>
<td>1.37, 0.86</td>
<td>2.09, 0.99</td>
<td>1.44, 0.91</td>
<td>1.38, 0.91</td>
</tr>
<tr>
<td>IMPACT</td>
<td>Accumulation mode</td>
<td>0.89, 0.71</td>
<td>0.54, 0.45</td>
<td>1.38, 1.01</td>
<td>0.54, 1.01</td>
</tr>
<tr>
<td>TM4-ECPL</td>
<td>Accumulation mode</td>
<td>0.37, 0.37</td>
<td>0.22, 0.22</td>
<td>0.57, 0.57</td>
<td>0.22, 0.57</td>
</tr>
<tr>
<td></td>
<td>Coarse mode</td>
<td>0.98, 0.91</td>
<td>0.62, 0.56</td>
<td>1.46, 1.38</td>
<td>0.62, 1.38</td>
</tr>
<tr>
<td></td>
<td>Fine mode</td>
<td>1.57, 1.75</td>
<td>1.32, 1.54</td>
<td>2.44, 2.09</td>
<td>1.46, 2.09</td>
</tr>
</tbody>
</table>

The median correction factors for the corresponding bins 1–4 are 0.13, 0.69, 1.25 and 1.81, respectively. Note also that for the ensemble model calculations of this study, the TFe and LFe depositions fluxes were calculated as the sum of Fe from the corrected mineral dust and the mean combustion aerosols.
2.3 Iron atmospheric observations

To evaluate the models’ ability to reproduce the observed distributions of surface TFe and LFe aerosol concentrations over oceans, the model results were compared with available observations from Achterberg et al. (2018), Baker et al. (2006a, b, 2007, 2013), Baker and Jickells (2017), Bowie et al. (2009), Buck et al. (2006, 2010, 2013), Clifton S. Buck (personal communication, 2018), Chance et al. (2015), Gao et al. (2013), Guieu et al. (2005), Cécile Guieu (personal communication, 2018), Jickells et al. (2016), Kumar et al. (2010), Longo et al. (2016), Powell et al. (2015), Shelley et al. (2015, 2018), Sholkowitz et al. (2012), Srinivas et al. (2012), Srinivas and Sarin (2013), Wagener (2008) and Wagener et al. (2008). A total of 818 observations of TFe and 795 daily observations of LFe over the ocean performed from September 1999 to March 2015 were used for this purpose. The global distribution of the observed aerosol Fe-concentrations used for this study is presented in Fig. S3 and the respective coordinates are shown in Table S3. A bulk dust deposition flux data set compiled by Albani et al. (2014) is also used for the comparison of the Fe deposition flux after multiplying the flux by the averaged Fe content in upper crustal minerals which was 3.5 %.

TFe values were obtained from sampling aerosols in air mainly on board oceanographic cruises, and for some studies at sampling stations located on shore or on islands with no local anthropogenic influence. A variety of samplers (small or high volume) were used to collect particulate Fe, using different types of filters. TFe was either measured on whole filters or parts of filters by X-ray fluorescence or after acid digestion. LFe was obtained following several protocols, e.g., contact time, volume of media and type of media (ultrapure water or filtered surface seawater for different pH conditions and various amounts of Fe ligands). As highlighted in Baker and Croot (2010), these diverse experimental approaches used for the determination of aerosol Fe solubility cause part of the variability observed. Although there is still no consensus regarding the experimental method to determine LFe, this data set provides valuable and robust data in all oceanic regions to be compared with model outputs. Here the results from the different protocols are all combined. Overall, the results are analyzed with regard to the role of the different model complexities, providing insight into directions for future model improvements.

3 Results

3.1 Global budgets

All participating models have submitted results to enable the analysis of the TFe and LFe global budgets and atmospheric concentrations, for both dust and combustion aerosols, emissions, dry and wet deposition fluxes, atmospheric processing and atmospheric loads (see Sect. 2). Concerning the temporal resolution, daily mean spatially resolved budget terms have been submitted for IMPACT, while for GEOS-Chem, CAM4 and TM4-ECPL monthly mean fields for all budget terms are provided. However, for atmospheric concentrations all models provided daily mean fields. Concerning the aerosol size distribution (see Table 2), IMPACT and CAM4 submitted budget fields for four size bins, TM4-ECPL for two size modes and GEOS-Chem provided results as bulk aerosols. CAM4 and TM4-ECPL submitted separate fields for proton and oxalate Fe solubilization, while IMPACT and GEOS-Chem provided total fields. IMPACT and CAM4 also submitted atmospheric processing terms for dust and combustion aerosols, while TM4-ECPL only submitted terms for dust aerosols as no solubilization processes are calculated for Fe combustion aerosol in the model. GEOS-Chem does not take Fe from combustion aerosols into account.

3.1.1 Iron sources and deposition

The computed TFe and LFe emissions and the deposition fluxes for all models are presented in Table 4. Note, however, that for LFe (both from mineral dust and combustion sources) the total sources (the sum of primary and secondary sources) are discussed here rather than just the primary emissions. The models use significantly different assumptions to describe the total LFe source to the atmosphere and therefore primary (emissions) and secondary (atmospheric processing) sources cannot be accurately separated from rapid formation assumed in coarse-scale models. The computed annual TFe emissions and LFe sources (emissions and atmospheric processing) from (1) mineral dust and (2) combustion sources for each model are presented in Table 4, and the corresponding global emission/sources distributions are shown in Figs. S4 and S5, respectively.

The importance of wet versus dry deposition as removal processes for atmospheric aerosols depends on the aerosol solubility and size distribution (and the presence/amount of precipitation). However, focusing on the deposition to oceans for this study, the apportionment of the total atmospheric deposition within different oceanic regions is presented in Table 5 (for this deposition analysis in this study we use the ocean classification as provided by HTAP phase-2: available online via the HTAP Wiki). Moreover, the combined computed annual deposition flux distributions of TFe and LFe from mineral dust and combustion sources for each model are presented in Fig. S6.

Overall, the Fe sources and deposition in the models are classified here as follows:

1. Total Fe: The modeled annual mean emission fluxes of TFe from mineral dust (TFeD) are calculated to be in the range of 38–134 Tg Fe yr⁻¹. CAM4 and GEOS-Chem calculate similar annual TFeD emission fluxes (around 57 Tg Fe yr⁻¹), TM4-ECPL emissions are about 40 % lower (∼ 38 Tg Fe yr⁻¹) and IMPACT
Table 4. Annual budgets (Tg Fe yr$^{-1}$) of total (TFe) and labile (LFe) Fe for emissions (EMI), dry deposition (DRY), wet deposition (WET) and sources (SRC; i.e., the sum of emissions and atmospheric processing) for dust (TFeD, LFeD) and combustion (TFeC, LFeC) Fe-containing aerosols as calculated by the models, as well as the models’ mean (± standard deviation; SD).

<table>
<thead>
<tr>
<th>Model</th>
<th>TFeD EMI</th>
<th>TFeD DRY</th>
<th>TFeD WET</th>
<th>LFeD EMI</th>
<th>LFeD DRY</th>
<th>LFeD WET</th>
<th>TFeC</th>
<th>LFeC</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAM4</td>
<td>57.1</td>
<td>33.4</td>
<td>23.7</td>
<td>1.9</td>
<td>0.8</td>
<td>1.1</td>
<td>0.2</td>
<td>0.8</td>
</tr>
<tr>
<td>GEOS-Chem</td>
<td>56.5</td>
<td>40.2</td>
<td>16.3</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.8</td>
<td>0.6</td>
</tr>
<tr>
<td>IMPACT</td>
<td>134.1</td>
<td>67.9</td>
<td>66.2</td>
<td>2.7</td>
<td>0.7</td>
<td>1.8</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>TM4-ECPL</td>
<td>38.1</td>
<td>30.34</td>
<td>7.76</td>
<td>1.8</td>
<td>1.4</td>
<td>0.4</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Mean (±SD)</td>
<td>71.5 (±42.69)</td>
<td>43.0 (±17.13)</td>
<td>28.4 (±26.97)</td>
<td>2.1 (±0.51)</td>
<td>1.1 (±0.71)</td>
<td>0.7 (±0.20)</td>
<td>0.4 (±0.29)</td>
<td>0.1 (±0.09)</td>
</tr>
</tbody>
</table>

is about 2.4 times higher (around 134 Tg Fe yr$^{-1}$). Note that IMPACT takes the largest flux among the participating models into account (Table S1), mainly due to the largest upper size. However, dust fluxes in the same size range for the different models are comparable after the bias correction based on the analysis by Kok et al. (2017) (see Sect. 2.2). TFe emissions from combustion sources (TFeC) range between 1.8 and 2.7 Tg Fe yr$^{-1}$, with CAM4 and TM4-ECPL calculating annual mean fluxes of around 1.8 Tg Fe yr$^{-1}$ globally, and IMPACT having a 35% higher estimate. On a global scale, dry deposition is the most important removal mechanism, across all models for the TFeD; IMPACT has the highest dry deposition flux of all models ($\sim$ 68 Tg Fe yr$^{-1}$), followed by GEOS-Chem ($\sim$ 40 Tg Fe yr$^{-1}$), CAM4 ($\sim$ 33 Tg Fe yr$^{-1}$) and TM4-ECPL ($\sim$ 30 Tg Fe yr$^{-1}$). Submicron aerosols are mostly removed by wet removal, while for supermicron aerosols the gravitational settling is important (e.g., Seinfeld and Pandis, 2006). Consequently, the wet removal of TFeD across almost all models (except for IMPACT) is smaller than dry deposition flux — mainly due to the high contribution of coarse aerosol sedimentation to the dry removal processes (Table 4). The simulated wet deposition flux of TFeD ranges over about 1 order of magnitude (from about 8 to 66 Tg Fe yr$^{-1}$); IMPACT calculates the highest TFeD wet deposition flux of all models of about 66 Tg Fe yr$^{-1}$ ($\sim$ 49% of total removal), followed by CAM4 (24 Tg Fe yr$^{-1}$; 42%), GEOS-Chem (16 Tg Fe yr$^{-1}$; $\sim$ 29%) and TM4-ECPL (roughly 8 Tg Fe yr$^{-1}$; $\sim$ 20%). In contrast to TFeD, due to the similar assumptions of the size distribution and scavenging efficiency in the models, the wet deposition is the larger removal pathway for TFe from combustion processes (TFeC), (except for in TM4-ECPL, probably due to the different solubility factors in primary emissions and the different atmospheric processing parameterizations), and is responsible for about 60% of the overall total TFeC removal across models which amounts to about 1 Tg Fe yr$^{-1}$.

2. *Labile Fe*: The global annual mean LFe sources from mineral dust (LFeD) range between 0.3 and 1.0 Tg Fe yr$^{-1}$. IMPACT and GEOS-Chem calculate similar LFeD sources, close to 0.7–0.8 Tg Fe yr$^{-1}$, whereas CAM4 calculates the highest annual source and TM4-ECPL the lowest. However, these differences are mainly attributed to the secondary processes leading to LFeD production rather than the primary emissions. For example, despite the large difference between IMPACT and TM4-ECPL regarding TFeD sources, the models consider similar LFeD emissions amounts ranging between 0.12 and 0.13 Tg Fe yr$^{-1}$. In contrast, the secondary LFeD produced due to atmospheric processing is calculated to vary by a factor of 3–4 be-
Figure 1 presents the global LFe sources (positive) and oceanic deposition fluxes (negative) for all participating models and their ensemble mean (see Sect. 3.2), for the four seasons, i.e., December, January and February (DJF); March, April and May (MAM); June, July and August (JJA); and September, October and November (SON). LFe sources are mainly driven by mineral dust aerosols, although a significant fraction (6% – 62%) is due to LFe combustion aerosols, especially over the high-latitudes of the Northern Hemisphere (Ito et al., 2018b). For LFe sources, despite the different assumptions applied in the models (i.e., atmospheric processing and direct LFe emissions), maximum sources are calculated for MAM and JJA due to intense dust emissions and biomass burning, respectively. The models with the highest LFe sources, also exhibit the highest deposition fluxes to the ocean. However, significant differences in the magnitude of the deposition fluxes are calculated between models (Fig. 1). A seasonal maximum in the deposition fluxes is calculated by CAM4 and GEOS-Chem during MAM, attributed to Saharan mineral dust aerosols, while IMPACT and TM4-ECPL present a seasonal maximum during JJA.

Figure S7 (Supplement) further presents the zonal mean seasonal variability of the LFe global sources and oceanic deposition fluxes. Most of LFe emissions are calculated to occur over the midlatitudes of the Northern Hemisphere (NH) for all seasons, with a maximum during MAM and JJA and minima during SON. In DJF, and to a lesser extent in JJA, two zonal maxima are shown near the Equator and around 30° N. However, the equatorial maximum in DJF is shifted to the Northern Hemisphere in JJA following the intertropical convergence zone (ITCZ) migration and the subsequent geographic change in the location of biomass burning emissions. Again, all models appear to have similar LFe seasonality, with the highest LFe oceanic deposition fluxes across all models calculated by CAM4 (Table 5).

### Table 5. Annual deposition fluxes (Tg Fe yr$^{-1}$) to different ocean basins of total (TFe) and labile (LFe) Fe, as calculated by the contributing models and the derived ensemble model.

<table>
<thead>
<tr>
<th>Ocean basin</th>
<th>CAM4</th>
<th>GEOS-Chem$^a$</th>
<th>IMPACT</th>
<th>TM4-ECPL</th>
<th>Ensemble</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TFe</td>
<td>LFe</td>
<td>TFe</td>
<td>LFe</td>
<td>TFe</td>
</tr>
<tr>
<td>North Atlantic</td>
<td>4.914</td>
<td>0.202</td>
<td>6.407</td>
<td>0.109</td>
<td>7.541</td>
</tr>
<tr>
<td>South Atlantic</td>
<td>0.681</td>
<td>0.043</td>
<td>2.067</td>
<td>0.017</td>
<td>4.414</td>
</tr>
<tr>
<td>North Pacific</td>
<td>1.084</td>
<td>0.057</td>
<td>1.653</td>
<td>0.041</td>
<td>0.972</td>
</tr>
<tr>
<td>South Pacific</td>
<td>0.111</td>
<td>0.007</td>
<td>1.114</td>
<td>0.01</td>
<td>2.337</td>
</tr>
<tr>
<td>Indian</td>
<td>3.658</td>
<td>0.076</td>
<td>3.818</td>
<td>0.049</td>
<td>6.911</td>
</tr>
<tr>
<td>Mediterranean</td>
<td>1.596</td>
<td>0.019</td>
<td>2.664</td>
<td>0.043</td>
<td>4.005</td>
</tr>
<tr>
<td>Baltic Sea</td>
<td>0.011</td>
<td>0.001</td>
<td>0.011</td>
<td>&lt; 0.001</td>
<td>0.017</td>
</tr>
<tr>
<td>Black &amp; Caspian seas</td>
<td>0.342</td>
<td>0.006</td>
<td>0.398</td>
<td>0.005</td>
<td>0.966</td>
</tr>
<tr>
<td>Hudson Bay</td>
<td>0.003</td>
<td>&lt; 0.001</td>
<td>0.006</td>
<td>&lt; 0.001</td>
<td>0.003</td>
</tr>
<tr>
<td>Arctic (&gt; 66°C N)</td>
<td>0.127</td>
<td>0.008</td>
<td>0.142</td>
<td>0.001</td>
<td>0.063</td>
</tr>
<tr>
<td>Southern (&gt; 60°C S)</td>
<td>0.010</td>
<td>0.001</td>
<td>0.030</td>
<td>&lt; 0.001</td>
<td>0.058</td>
</tr>
<tr>
<td>Global ocean</td>
<td>12.538</td>
<td>0.419</td>
<td>18.31</td>
<td>0.275</td>
<td>27.287</td>
</tr>
</tbody>
</table>

$^a$ In GEOS-Chem, only the Fe from mineral dust is considered.
Figure 1. Seasonal LFe sources (positive bars) and oceanic deposition fluxes (negative bars/pale colors) in Tg Fe season$^{-1}$ for December, January and February (DJF); March, April and May (MAM); June, July and August (JJA); and September, October and November (SON), as calculated by each model (CAM4, magenta; GEOS-Chem, red; IMPACT, green; and TM4-ECPL, blue), as well as, the ensemble model (yellow). The hatched areas correspond to the combustion aerosols and the error bars correspond to the standard deviation of the respective season.

Figure 2. Ensemble model results for annual mean (a) surface TFe concentration (µg m$^{-3}$), (b) the percentage contribution of Fe-containing combustion aerosols, (c) the Fe solubility (%) in surface TFe concentration and (d) the initial solubility (%) in Fe-containing dust emissions.
3.2 Ensemble model calculations

3.2.1 Iron surface concentrations

The annual mean surface TFe aerosol concentrations for the ensemble model exceed 100 µg Fe m$^{-3}$ over the major dust source regions such as the Sahara Desert, where mineral dust particles dominate the atmospheric Fe burden (Fig. 2a). Relatively high TFe concentrations (e.g., up to 10 µg Fe m$^{-3}$ over the tropical Atlantic Ocean) are calculated for ocean regions at the outflow from these source regions. High TFe concentrations of around 6 µg Fe m$^{-3}$ are also calculated over heavily polluted areas like China, while secondary maxima up to 2–5 µg Fe m$^{-3}$ are calculated over central Africa, Asia and Indonesia, where Fe-containing aerosols are associated with biomass burning emissions (Fig. 2b).

Model Fe solubility calculations (Fig. 2c) clearly suggest the impact of atmospheric processing on the derived LFe ensemble surface concentrations, with high Fe solubilities calculated far from source regions over the remote tropical oceans, corresponding to low TFe concentrations. Ensemble annual mean LFe concentrations of around 0.5 µg Fe m$^{-3}$ occur downwind of the Sahara and of around 0.1 µg Fe m$^{-3}$ downwind of the Arabian and Gobi deserts. At the outflow of these regions, the Fe solubility over the global ocean is calculated to be about 1%–1.5%, with the highest Fe solubilities (4%–5%) over the tropical Atlantic Ocean (Fig. 2c). Additionally, LFe concentrations over polluted regions may range up to 0.05 µg Fe m$^{-3}$, indicating a significant anthropogenic contribution via direct combustion emissions and atmospheric processing (Fig. 2b). Over central South America, Asia and Indonesia, LFe concentrations of about 0.03–0.05 µg Fe m$^{-3}$ (corresponding to high Fe solubilities up to 5%) are found due to both direct biomass-burning emissions and due to ligand-promoted dissolution. The latter process is enhanced in these areas by the respective enhanced oxalate production upon the oxidation of emitted biogenic VOCs precursors, such as isoprene, under cloudy conditions (Lin et al., 2014; Myriokefalitakis et al., 2011).

3.2.2 Iron deposition fluxes

Model calculations indicate that about 71.5 (±43) Tg Fe yr$^{-1}$ of TFe from mineral dust are deposited to the Earth’s surface (Table 4), with ensemble deposition fluxes of around 5000–8000 mg Fe m$^{-2}$ yr$^{-1}$ calculated downwind of the main desert source regions (Fig. 3a). However, within the northern Atlantic Ocean in the outflow of the Sahara, the model mean indicates deposition fluxes up to 2400 mg Fe m$^{-2}$ yr$^{-1}$, while within the northern Pacific Ocean in the outflow of Gobi Desert and within the Southern Ocean downwind of the Patagonia Desert the ensemble model shows annual mean fluxes of ∼ 34 and ∼ 10 mg Fe m$^{-2}$ yr$^{-1}$, respectively (Fig. 3a). The TFe annual mean global deposition flux from combustion sources (Table 4) is calculated to be about 2.2 (±0.5) Tg Fe yr$^{-1}$, with two main regions where TFe concentrations exceed 2500 mg Fe m$^{-2}$ yr$^{-1}$, one near biomass burning regions (e.g., southern Africa, South America and Southeast Asia) with up to ∼ 3000 mg Fe m$^{-2}$ yr$^{-1}$, and a second near highly populated regions with Fe released from coal and oil combustion processes (India and China) with up to ∼ 3500 mg Fe m$^{-2}$ yr$^{-1}$ (Fig. 3b).

A global mean LFe deposition flux of 0.7 (±0.2) Tg Fe yr$^{-1}$ is derived from all models (Table 4), with about one-third (∼ 0.24 Tg Fe yr$^{-1}$) calculated to be deposited to the global ocean for the ensemble model (Table 5). The highest annual mean LFe deposition fluxes (up to 36 mg Fe m$^{-2}$ yr$^{-1}$) are simulated within dust source regions (Fig. 3c), mainly owing to the LFe content of the emissions (e.g., see Fig. 2d). The global model-mean LFe deposition fluxes from combustion sources are calculated at about 0.2 (±0.04) Tg Fe yr$^{-1}$ (Table 4), with maximum global deposition rates of 4–5 mg Fe m$^{-2}$ yr$^{-1}$ (Fig. 3c) simulated in the outflow of tropical biomass burning regions (i.e., South America, Africa and Indonesia), clearly reflecting the contribution of combustion processes. Focusing on the marine environment, annual mean LFe deposition rates of 15 mg Fe m$^{-2}$ yr$^{-1}$ are calculated for the tropical Atlantic Ocean and for the Indian Ocean (up to 16 mg Fe m$^{-2}$ yr$^{-1}$) under the influence of the Arabian and Indian peninsulas, but up to ∼ 29 mg Fe m$^{-2}$ yr$^{-1}$ for the Mediterranean Sea downwind of the Sahara Desert. Deposition rates around 1 mg Fe m$^{-2}$ yr$^{-1}$ are calculated to occur within the northern Pacific in the outflow from the Gobi Desert as well as within the Southern Hemisphere downwind of Patagonia to the Southern Ocean (up to 0.1 mg Fe m$^{-2}$ yr$^{-1}$) and downwind of the dust source regions of Australia and South America (up to ∼ 4 mg Fe m$^{-2}$ yr$^{-1}$). The LFe deposition rates to the Southern Ocean are mainly associated with the Patagonian, southern African and Australian deserts, with a smaller contribution in the subtropical ocean from biomass burning sources (Fig. 3d). Note, that the largest fluxes of LFe deposited to the HNLC region of the Southern Ocean (e.g., south of the Antarctic Circumpolar Current) are simulated to be originating from a Patagonian mineral dust source, with rates reaching 0.1 mg Fe m$^{-2}$ yr$^{-1}$. The ensemble annual mean deposition fluxes to various oceanic regions are further presented in Table 5.

4 Uncertainties

4.1 Comparison with measurements

The TFe concentrations, Fe solubility and LFe concentrations from the models are compared with the measurements and presented in Fig. 4. We use the monthly mean of the model output to compare with the measurements. The normalized bias (NB) at a given grid box is calculated as fol-
Figure 3. Ensemble model results for annual deposition fluxes (mg m\(^{-2}\) yr\(^{-1}\)) for (a) TFe and for (c) LFe and their respective percentage contribution of combustion aerosols (b, d).

The normalized bias (NB), a measure of the discrepancy between modeled and observed aerosol concentrations, is defined as:

\[
NB_i = \frac{C_{\text{model},i} - C_{\text{obs},i}}{C_{\text{obs},i}},
\]

where \(C_{\text{model},i}\) is the modeled aerosol concentration in grid box \(i\) and \(C_{\text{obs},i}\) is the measured aerosol concentration in the same grid box. When discussing the multi-model results we use the mean of all models, while we also analyze the mean normalized bias (MNB) of the models against measurements (a perfect comparison would have an MNB of 0 and correlation, \(R\), of 1). A model’s MNB is derived as the arithmetic mean of all NB\(_i\) values; thus, overestimates are weighted more than equivalent underestimates.

All models captured a tendency for higher Fe concentrations near and downwind of the major dust source regions. The MNB for the Northern Hemisphere (14 for ensemble model) is larger than that for the Southern Hemisphere (2.4 for ensemble model). This reflects that the models generally overestimate TFe surface mass concentrations. However, from Fig. 4 we can see that this overestimate is higher for the highest TFe concentrations near the dust source regions and tends to become an underestimate for the lowest concentrations observed over remote oceans. Overall, bias correction for the ensemble model improves the agreement of the ensemble model against measurements (Fig. S8). However, we note that matching the atmospheric concentrations may cause a high bias in the simulated Fe depositions at low values in the Southern Hemisphere (Albani et al., 2014; Huneeus et al., 2011) (Fig. S9). The computed correlation coefficients of the ensemble model against measurements at the surface are 0.13 for TFe, 0.05 for Fe solubility and 0.25 for LFe, respectively, which are much smaller than those between the participating models: 0.57–0.90 for TFe (GEOS-Chem vs. IMPACT)–(CAM4 vs. TM4-ECPL), 0.05–0.56 for Fe solubility (CAM4 vs. TM4-ECPL)–(CAM4 vs. IMPACT), and 0.40–0.75 for LFe (GEOS-Chem vs. IMPACT)–(CAM4 vs. GEOS-Chem). This indicates a linear dependence of the model results and that the models have similar behavior accounting for the same key processes that affect Fe deposition. The small positive correlation between models and observations indicates that the models miss – or do not accurately represent – important processes that drive the variability in the observations. Indeed, when comparing the computed and observed solubilities of Fe, all models overestimate the lowest solubilities (< 0.1%) observed close to the source regions (14 samples in the Arabian Sea and 2 samples in the tropical Atlantic). This is primarily due to the assumed solubility of the dust aerosols at emissions, and the subsequent enhancement of Fe solubility estimated from the simulated amount of atmo-
Figure 4. Comparison of simulated and observed TFe concentrations (ng m$^{-3}$), Fe solubility (%) and LFe concentrations (ng m$^{-3}$) in the Northern Hemisphere (blue circles) and Southern Hemisphere (red squares) for (a, b, c) CAM4, (d, e, f) GEOS-Chem, (g, h, i) IMPACT, (j, k, l) TM4-ECPL, and (m, n, o) the ensemble model. The mean normalized biases (MNB) between the models and observations are presented in parentheses. The solid line represents a 1 : 1 correspondence and the dashed lines show the 10 : 1 and 1 : 10 relationships, respectively. The bias correction in the mineral dust size distribution is applied for the comparison with field data (Kok et al., 2017).

Figure 5. Fe solubility versus atmospheric concentrations of aerosol Fe (ng m$^{-3}$) in the Northern Hemisphere (blue circles) and Southern Hemisphere (red squares) for (a) the measurements, (b) CAM4, (c) GEOS-Chem, (d) IMPACT, (e) TM4-ECPL and (f) the ensemble model. The bias correction in the mineral dust size distribution is applied for the comparison with field data (Kok et al., 2017).

spheric processing during transport in the models. However, it is noted that the lowest solubilities in the measurements are outliers in the negative slope of Fe concentrations versus Fe solubility in the Northern Hemisphere (Fig. 5). Our calculations show that the models have difficulties simulating the 4 orders of magnitude variability from 0.02 % to 98 % in the Fe solubility observed in the atmosphere (Fig. 5a). IMPACT simulates almost 3 orders of magnitude variability in Fe solubility. In the other models, including the ensemble model, Fe solubility is less variable (only 1–2 orders of magnitude). In particular, low solubilities (high concentrations near sources) are overestimated and high solubilities (low concentrations at remote locations) are underestimated. This may indicate that the primary LFe in the models is overestimated and that models are missing solubilization processes during transport or that those considered in the models are not sufficiently effective.

All models underestimate the high end of the observed values (> 10 %) in the Southern Ocean, which are mainly associated with transported and aged aerosols; this may be a significant shortcoming because the Southern Ocean is an oceanic HNLC region where the atmospheric Fe supply has a potentially important impact on ocean productivity in past and future climate. In GEOS-Chem, IMPACT and TM4-ECPL, Fe dissolution over the Southern Ocean is mainly suppressed due to the lack of anthropogenic emissions and the subsequent acidification of the aerosols. CAM4 is relatively insensitive to the acidity, as most labile Fe is formed via in-cloud processes. Thus, the model results from CAM4 can in part test whether in-cloud processing can realistically describe the observed pattern of solubility over the Southern Ocean.
Ocean. CAM4 shows higher Fe solubility than field data for most samples in the Southern Ocean (69 % in CAM4, 7 % in GEOS-Chem, 55 % in IMPACT and 5 % in TM4-ECPL; Figs. 4 and 5). Thus, the wide range of observed Fe solubility cannot be explained by excluding the effect of modeled aerosol acidity over the Southern Ocean. It is worth mentioning that IMPACT, which has the highest complexity in simulating labile Fe, reproduces the widest range of observed Fe solubility. However, it should be noted that the comparison of monthly mean model results with the shorter-term (e.g., daily) observations during different sampling periods introduces inaccuracies due to the episodic nature of high Fe solubility. A more detailed comparison of Fe solubility between models and observations is presented in a separate companion paper to this work (Ito et al., 2018b).

4.2 Model-to-model comparison

Model budget analysis and model evaluation indicate that even though the models are able to reproduce surface Fe measurements to some extent, large differences exist among modes in processes such as emissions, transport and deposition. A large diversity is documented here between models in terms of LFe primary sources (i.e., emissions) and secondary processes (i.e., atmospheric processing) which introduce uncertainties in the estimated oceanic deposition. However, there are many intrinsic reasons for this diversity in the Fe simulations among models: besides the emitted Fe mass in the atmosphere (especially from dust aerosols), the aerosol size distribution, the soil mineralogy, the strength of combustion aerosol sources, as well as the parameterizations used to calculate the pH of the aerosol water and the oxalate production are large sources of uncertainty in model simulations. The aerosol size and solubility are important factors driving the atmospheric cycle of Fe, as they both control the removal processes from the atmosphere via the dry deposition (including gravitational settling) and the wet scavenging (e.g., Albani et al., 2014). It is well documented that the lifetime of Fe-containing aerosols ranges between less than 1 day for the coarse mode (with diameters larger than 1 µm) particles to weeks for the fine mode (with diameters less than 1 µm) (e.g., Ginoux et al., 2012; Luo et al., 2008; Mahowald et al., 2009; Tegen and Fung, 1994), with the overall lifetime of dust aerosols usually ranging between 1.6 and 7.1 days in the models (Huneeus et al., 2011). Fine particles have longer lifetimes and thus experience more atmospheric processing. The conversion of insoluble mineral content to soluble forms as the result of aging during atmospheric transport increases aerosols’ solubility. Lifetime calculations can thus provide a valuable tool to determine the Fe persistence in the atmosphere, overall integrating sources, transport and deposition differences among models, especially over remote areas such as the open ocean.

Figure 6 presents the spatial distribution of the TFe lifetime over the ocean (i.e., atmospheric concentrations divided by total sinks), as calculated for the ensemble model. For TFe originating from dust sources, the calculated global mean lifetime over the oceans is 6 ± 4 days (Fig. 6a). The lifetime of combustion TFe over oceans is longer, at around 14 ± 9 days (Fig. 6c), due to the generally smaller size of the combustion Fe aerosols compared to that of mineral dust (affecting their sedimentation processes and their horizontal and vertical transport in the atmosphere), but also because of the low precipitation over part of the regions in which mineral dust and combustion aerosols are transported. The ensemble model indicates long TFe lifetimes over remote oceanic regions, such as in the outflow of South America, in the outflow of South Africa and the outflow of Australia (Fig. 6c). Over these so-called “ocean deserts”, where precipitation is low and thus the wet deposition rates are low, the ensemble model generally results in longer lifetimes for Fe-containing aerosols, although the LFe atmospheric concentrations can be extremely low.

To further analyze the differences among the models, the standard deviation (SD) of the TFe lifetime is calculated for each grid for dust aerosols (Fig. 6b), combustion aerosols (Fig. 6d) and the combined TFe lifetime (Fig. 6f). Over remote oceanic regions, the high SD can be related to the different assumptions used by the models to parameterize (1) the long-range transport of Fe-containing aerosols of different parameterizations of the sizes, (2) the wet and dry deposition parameterizations and (3) the soluble fraction of Fe-containing combustion aerosols (e.g., the differences in ship oil combustion and biomass burning emissions). From the models that include Fe from combustion processes, IMPACT assumes that only ship oil combustion emissions have an initial fraction of soluble Fe; thus, all the LFe from continental sources is produced due to atmospheric processing. CAM4 does not take ship emissions into account but includes both primary and secondary sources for LFe from other combustion sources. TM4-ECPL, in comparison, includes continental and ship oil combustion emissions for both TFe and LFe, but it does not take any dissolution processes for combustion aerosols into account, although continental and ship oil combustion emissions both for TFe and LFe are considered. Differences in the precipitation patterns or parameterization for wet or dry deposition in the models can also partially explain the models’ diversity. In addition, less constrained parameters like dissolution, long-range transport and Fe removal (which affect Fe solubility) can further increase the model diversity, and in turn the SD, over these remote areas.
5 Conclusions and future directions

Here we present the first model intercomparison study of the atmospheric Fe-cycle by assessing aerosol simulations of total and labile Fe using four state-of-the-art global aerosol models (CAM4, IMPACT, GEOS-Chem and TM4-ECPL). The TFe emissions from dust sources in the models range from $\sim 38$ to $\sim 134$ Tg Fe yr$^{-1}$, with a mean value of 71.5 ($\pm 43$) Tg Fe yr$^{-1}$. The models simulate the secondary formation of soluble Fe in the atmosphere, as a result of mineral Fe atmospheric processing by acids, organic ligands and photochemistry, but the absolute amount of the simulated LFe remains highly uncertain. The simulated LFe deposition fluxes from mineral dust span from 0.3 to 1.0 Tg Fe yr$^{-1}$, with a mean value of around 0.7 ($\pm 0.3$) Tg Fe yr$^{-1}$. All models capture the main features of the distribution of TFe and LFe, i.e., the large deposition rates to the Sahara and Gobi deserts and regions downwind of strong dust sources. Models also show significant LFe deposition to the oceans downwind from the Middle East and the continents of South America, Africa and Australia in the Southern Hemisphere; the Middle East has large dust sources, while South America, Africa and Australia experience strong biomass burning emissions.
The models are able to simulate the main features of TFe and LFe atmospheric concentrations and deposition fluxes. On average, the ensemble model computes a roughly $50\%$ higher lifetime with respect to the deposition flux for combustion aerosol ($\sim 14$ days) than for dust aerosol, reflecting differences in the size distribution and the location of the emitted aerosols. Ensemble model calculations present an overestimate of the observed TFe surface mass concentrations near the dust source regions and underestimate the Fe concentrations over the Southern Ocean compared with cruise measurements; similar to what was pointed out in Albani et al. (2014) and seen in the dust model intercomparison study of Huneeus et al. (2011). Note that the latter is important because of the key role of Fe in the biogeochemistry of these ocean waters. For the ensemble model mean, the MNB for the Northern Hemisphere (about 14) is larger than for the Southern Hemisphere (about 2.4). Note, however, that the evaluation of monthly mean model results by comparison with the shorter-term (e.g., daily) observations during different sampling periods introduces uncertainties, due to the short sampling frequencies; a comparison of long-term measurements with a multi-year modeling would allow for the assessment of the model performance regarding the capture of labile Fe concentrations under specific events.

The model intercomparison and model–observation comparison revealed the Fe size distribution and the relative contribution of the dust and combustion sources to be two critical issues for LFe simulations that now require further study. Thus, the diversity of how the models represent Fe emissions as well as of deposition fluxes among the models can be large, especially over source regions. The model diversity over remote oceans reflects uncertainty in the Fe content parameterizations of dust emissions (e.g., soil mineralogy and the initial Fe soluble content in primary sources) and combustion aerosols, and/or in the parameterizations of the size distribution of the transported aerosol Fe, and in turn the representation of deposition fluxes – which generally control the atmospheric lifetime of Fe. Conversely, there are many other intrinsic reasons for this diversity, especially for the LFe aerosol fraction, as it involves complex atmospheric chemical processes driven by atmospheric acidity. For example, detailed chemical mechanisms need to be invoked to simulate a multi-phase, multi-component solution system, since such a system may not be accurately solved using a thermodynamic equilibrium approach for the entire grid box due to sub-grid processes, e.g., when the dust plume is not well mixed with surrounding pollutants. Consequently, a reasonable aerosol pH simulation further depends on the representation of soluble acidic and basic compounds, as well as the water content of hygroscopic particles.

In this respect, new field observations are needed to improve the understanding of the Fe solubilization process, and how this process alters in the presence of anthropogenic pollution. Modeling studies and their evaluation based on a greater number of atmospheric observations, especially over the remote ocean regions, are deemed necessary in order to reduce the uncertainty associated with the model performance in simulating the atmospheric Fe deposition. For example, the participating models in this study predict that wet deposition processes are important for the LFe atmospheric cycle; however, this is rather hard to test well, due to the lack of respective field data. Moreover, although the models do well in higher dust and pollution regions, from an oceanographer’s perspective, the regions with the lowest Fe supply are of greatest interest, due to the creation of HNLC situations in the water column. However, model evaluation can also be difficult due to the lack of standardization in the protocols used to determine the soluble Fe fraction in marine aerosol samples (e.g., Baker and Croot, 2010). Protocols that involve different solutes, aerosol–solution contact times and filter pore sizes, among other differences, are used by different investigators and these presumably introduce some, as yet unquantified, uncertainties into the available databases of aerosol soluble Fe concentrations and Fe solubility. Model developments related to the atmospheric Fe cycle must be performed in parallel with an extensive model evaluation in order to better understand the underlying mechanisms and to provide, overall, realistic labile Fe deposition fluxes for the next generation of ocean biogeochemistry modeling studies. Although the calculation of the Fe concentrations in the ocean is outside the scope of this paper, we expect that the Fe deposition fluxes here provided will be used in oceanic models.

Data availability. Global $1.0 \times 1.0$ fields for TFe and LFe deposition fluxes (i.e., from combustion and dust origin and their sum) as well as the Fe solubility ($S_{Fe}$) derived from the ensemble model calculations are available online at https://ecpl.chemistry.uoc.gr/GESAMP/, last access: 5 November 2018 (contact: Stelios Myriokefalitakis, steliosm@noa.gr; Maria Kanakidou, mariak@uoc.gr).

Author contributions. This paper resulted from the deliberations of United Nations GESAMP Working Group 38, “The Atmospheric Input of Chemicals to the Ocean” (SM, AL, MK, AN, NMM, ARB, TDJ, MMS, RS, MMGP, CG, RAD), February 2017, University of East Anglia, Norwich, UK. SM manipulated all model fields, analyzed the data and prepared the respective figures and fields of this work. ARB performed the model evaluation and the respective statistical analysis. JFK provided the analysis for dust correction factors for model data. SM, AL, MK, MCK, AN, NMM, RAS, DSH, NM and MSJ, provided the model products. ARB, TDJ, MMS, SB, RS, and CG contributed to the measured data. All authors contributed to the preparation of the paper.
Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. Stelios Myriokefalitakis acknowledges financial support for this research from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 705652 – ODEON. Support for this research was provided to Akinori Ito by JSPS KAKENHI grant no. JP16K00530 and the Integrated Research Program for Advancing Climate Models (MEXT). Natalie M. Mahowald, Rachel A. Scanza and Douglas S. Hamilton acknowledge support from a Department of Energy grant (DE-SC0006791CE14). Andrew Bowie and Morgane M. G. Perron were funded by an Australian Research Council Future Fellowship (FT130100037). We are grateful to Thibaut Wagener for kindly providing the field data set. This work is prepared in the framework of the Joint Group of Experts on the Scientific Aspects of Marine Environmental Protection (GESAMP; http://www.gesamp.org/, last access: 5 November 2018), Working Group 38, the Atmospheric Input of Chemicals to the Ocean. We thank the Global Atmosphere Watch and the World Weather Research Programme of the World Meteorological Organization, the International Maritime Organization, the U.S. National Science Foundation, the ICSU Scientific Committee on Oceanic Research (SCOR) and the University of East Anglia for their financial support. We also thank SOLAS for their sponsorship. Data post-processing was carried out on the Dutch national e-infrastructure with the support of the SURF Cooperative. The publication of this work was financed by Utrecht University and by the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 705652 – ODEON.

Edited by: Maria Tzortziou
Reviewed by: two anonymous referees

References


Heimburger, A., Losno, R., and Triquet, S.: Solubility of iron and other trace elements in rainwater collected on the Kerguelen


Moore, J. K., Lindsay, K., Doney, S. C., Long, M. C., and Misumi, K.: Marine Ecosystem Dynamics and Biogeochemical Cycling in the Community Earth System Model [CESM1(BGC)]: Comparison of the 1990s with the 2090s under the RCP4.5 and RCP8.5 Scenarios, J. Climate, 26, 9291–9312, https://doi.org/10.1175/JCLI-D-12-00566.1, 2013.


