FALL3D-8.0: a computational model for atmospheric transport and deposition of particles, aerosols and radionuclides – Part 2: model applications

Andrew T. Prata[1], Leonardo Mingari[1], Arnau Folch[1], Giovanni Macedonio[2], and Antonio Costa[3]

[1]CASE Department, Barcelona Supercomputing Center, Barcelona, Spain
[2]Istituto Nazionale di Geofisica e Vulcanologia, Osservatorio Vesuviano, Napoli, Italy
[3]Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Bologna, Bologna, Italy

Correspondence: Andrew Prata (andrew.prata@bsc.es)

Abstract. This manuscript presents different application cases and validation results of the latest version release of the *FALL3D-8.0* model, an open-source atmospheric transport model. The code has been redesigned from scratch to incorporate different categories of species and to overcome legacy issues that precluded its preparation towards extreme-scale computing. Validation results are shown for long-range dispersal of fine volcanic ash and SO_2 clouds, tephra fallout deposits and dispersal and ground deposition of radionuclides. The first two examples (i.e. the 2011 Puyehue-Cordón Caulle and 2019 Raikoke eruptions) make use of geostationary satellite retrievals for two purposes: first, to furnish an initial data insertion condition for the model; and second, to validate the time series of model outputs against the satellite retrievals. The metrics used to validate the model simulations of volcanic ash and SO_2 are the Structure, Amplitude and Location (SAL) metric and the Figure of Merit in Space (FMS). The other two application cases are validated with scattered ground-based observations of deposit load and local particle grain size distributions from the 23 February 2013 Mt. Etna eruption and with measurements from the Radioactivity Environmental Monitoring (REM) database during the 1986 Chernobyl nuclear accident. Simulation results indicate that *FALL3D-8.0* outperforms previous code versions both in terms of model accuracy and code performance. We also find that simulations initialised with the new data insertion scheme consistently improve agreement with satellite retrievals at all lead times out to 48 hours for both SO_2 and long-range fine ash simulations.

1 Introduction

FALL3D-8.0 is the latest major version release of *FALL3D* (Costa et al., 2006; Folch et al., 2009), an open-source code with 15+ years of track record and a growing number of users in the volcanological and atmospheric science communities. A companion paper (Folch et al., 2020) details the physics and novel numerical implementation of the code, which has been redesigned and rewritten from scratch in the framework of the EU Center of Excellence for Exascale in Solid Earth (*ChEESE*). From the point of view of model physics, a relevant improvement in the new version v8.x has been the generalisation of the code to deal with atmospheric species other than tephra such as other types of particles (e.g. mineral dust), aerosols and radionuclides (see Table 3 in Folch et al., 2020, for details). These different categories and sub-categories of species can be simulated using

independent sets of bins and dedicated parameterisations for physics, emissions (source terms) and interactions among bins
(*e.g.* aggregation, chemical reactions, radioactive decay, etc.). In terms of model performance, the new model version contains a
25 much more accurate and less diffusive solver, as well as a better memory management and parallelisation strategy that notably
outperforms the scalability and the computing times of the precedent code versions v7.x (Folch et al., 2020).

This paper complements the companion paper by Folch et al. (2020) presenting a detailed set of validation tests on several
real application cases, all included in the new benchmark suite of the code. Here we present model validation examples for
different species including dispersal and deposition of tephra, SO_2 clouds and radionuclides. The manuscript also contains
30 some novel aspects regarding satellite detection and retrievals, as well as a data insertion methodology for volcanic ash and SO_2
clouds. This data insertion scheme is a preliminary step towards model data assimilation using ensembles, a novel functionality
currently under development. Section 2 presents the methodology for ash and SO_2 detection and quantitative retrievals based on
InfraRed (IR) geostationary satellite measurements, which takes into account collocated lidar observations of cloud-top height
and thickness from the CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation) platform (Winker et al.,
35 2009). These high temporal resolution mass loading retrievals are used for quantitative model validation and to demonstrate
the new data insertion scheme in *FALL3D-8.0* for both volcanic ash and SO_2 dispersion. The data insertion scheme uses a
conservative interpolation algorithm to furnish an initial model condition from the satellite mass loading retrievals. Section 3
describes the validation metrics employed here, which include the Structure, Amplitude and Location (SAL) metric (Wernli
et al., 2008) and the Figure of Merit in Space (FMS; Galmarini et al., 2010; Wilkins et al., 2016) to quantitatively compare
40 model results with satellite retrievals and the minimization of the Root-Square-Mean-Error (RSME) and similar metrics for
validation of the ground deposit simulations of tephra and radionuclides. Section 4 describes the different validation cases
(summarised in Table 1), which include simulations of the June 2011 Puyehue-Cordón Caulle ash cloud, the June 2019 Raikoke
SO_2 cloud, the 23 February 2013 Mt. Etna ash cloud and associated tephra fallout deposit and the dispersal of radionuclides
resulting from the 1986 Chernobyl nuclear accident. Finally, Section 5 summarises the conclusions of the manuscript and
45 outlines the next steps in terms of model development and applications.

2 Satellite retrievals

2.1 Volcanic ash

Satellite detection of volcanic ash using passive IR Brightness Temperature Differences (BTD) between channels centered
around 11 and 12 μm has been widely used for more than 30 years (Prata, 1989a, b). Quantitative ash retrievals based on the
50 BTD method have also been practiced for a long time (e.g. Wen and Rose, 1994; Prata and Grant, 2001) and the uncertainties
stemming from detection (e.g. Simpson, 2000; Prata et al., 2001) and retrievals (e.g. Wen and Rose, 1994; Corradini et al.,
2008; Kylling et al., 2014; Stevenson et al., 2015; Western et al., 2015) are well-known. In order to validate *FALL3D-8.0* and
test the new volcanic ash data insertion scheme, we use measurements from the SEVIRI (Spin Enhanced Visible and Infrared
Imager; Schmetz et al., 2002) instrument onboard Meteosat-9 during the 2011 Puyehue-Cordón Caulle (hereafter Cordón
55 Caulle) eruption in Chile (described in Sec. 4.1).

2.1.1 Volcanic ash detection

Figure 1 shows SEVIRI observations of the Cordón Caulle volcanic ash plume and illustrates the steps used to detect volcanic ash in the present study. For context, Fig. 1a and 1b show a composite of MODIS true colour imagery and the SEVIRI 10.8 μm brightness temperature (T_B^{11}) respectively. Here, we propose an ash detection scheme based on applying successive masks that flag SEVIRI pixels as 'ash-affected' before attempting a subsequent quantitative ash retrieval (Sec. 2.1.2):

1. First, we apply a temperature cut-off threshold to water vapor corrected BTDs (ΔT_{ash}):

$$\Delta T_{ash} = T_B^{11} - T_B^{12} < T_{wc} \tag{1}$$

that is, only those pixels with ΔT_{ash} less than the cut-off threshold of $T_{wc} = -0.5$ K are flagged as potential ash pixels. This water vapor correction follows the semi-empirical approach of Yu et al. (2002). As illustrated in Fig. 1c, this first threshold is reasonably effective at detecting the Cordón Caulle ash cloud. However, this simple cut-off threshold may not remove false positives due to temperature inversions generated by clear land at night (Platt and Prata, 1993), ice-covered surfaces (Yamanouchi et al., 1987), cold cloud-tops (Potts and Ebert, 1996) and high satellite zenith angles (Gu et al., 2005).

2. Second, we apply a cold surface mask designed to remove false positives due to reasons mentioned above. This cut-off condition relabels potential ash pixels as 'ash free' if:

$$\Delta T_{ash} > T_{sc} \text{ and } \begin{cases} T_B^{11} > 255 \text{ K over land} \\ T_B^{11} > 240 \text{ K over ocean} \end{cases} \tag{2}$$

where $T_{sc} = -1.5$ K is the cold surface cut-off value. We note that T_B^{11} condition of this mask will preserve ash detection sensitivity for high altitude (cold) ash clouds, which is particularly well-suited for the Cordón Caulle case study. However, this condition may not be suitable for low-level ash clouds (low thermal contrast resulting in less negative BTDs in addition to warmer cloud-tops). The effect of this mask is illustrated by comparing Figs. 1c and d. Note how, for the case shown, the cold surface mask removes almost all false positives over the region covered by low-level stratiform cloud.

3. Third, we apply a mask for false positives due to an increased path length at high satellite zenith angles (Gu et al., 2005). We mask out false positives at high zenith angles imposing:

$$\Delta T_{ash} > T_{zc} \text{ and } \zeta > 80^o \tag{3}$$

where $T_{zc} = -2$ K is the zenith cut-off threshold and ζ is the satellite zenith angle. The effect of the high zenith mask can be seen by comparing Figs. 1d and e.

4. Finally, the last step in the detection process is to remove any spurious ash-labeled pixels using a noise filter that removes objects (groups of contiguous pixels) that are less than 16 pixels in size (Fig. 1f).

The MODIS true color composite shown in Fig. 1a illustrates that, even in a relatively complex scene (numerous clouds, large

85 regions of land and ocean, high mountains, ice-covered surfaces, etc.), the ash detection is robust and provides a good balance between reduced false positives and increased true positives. An interesting point to note is that negative BTDs in the vicinity of Cordón Caulle are enhanced due to the high satellite zenith angles at these locations. Gu et al. (2005) discuss the benefit of improved sensitivity to ash at high satellite zenith angles, but also show that mass loading retrievals can be overestimated in these situations. We correct for the effect of high zenith angles after retrieving the mass loading.

90 ### 2.1.2 Volcanic ash retrieval

Once pixels have been identified as being 'ash-affected' we apply a Look-up Table (LuT) approach (Prata and Grant, 2001; Prata and Prata, 2012) to retrieve volcanic ash optical depth (τ), effective radius (r_e; in μm), and column mass loading (m_l; in $\mathrm{gm^{-2}}$). The retrieval procedure is illustrated in Fig. 2a. The temperature difference model employed here is based on the forward model developed by Prata (1989b) and Wen and Rose (1994):

95 $$I_\lambda \approx e^{-\tau(\lambda)}B(T_s) + \left(1 - e^{-\tau(\lambda)}\right)B(T_c) \tag{4}$$

where I_λ is the radiance at the top of the atmosphere at wavelength (λ), τ_λ is the wavelength-dependent optical depth, $B(T_s)$ is the Planck radiance evaluated for surface temperature (T_s) below the ash cloud, and $B(T_c)$ is the Planck radiance corresponding to the temperature at the ash cloud-top (T_c). The optical depth is defined as:

$$\tau(\lambda) = \pi L \int\limits_0^\infty r^2 Q_{ext}(\lambda, r) n(r) dr \tag{5}$$

100 where L is the geometric thickness of the ash cloud, $Q_{ext}(\lambda, r)$ is the extinction efficiency factor (determined from Mie calculations), r is the particle radius and $n(r)$ represents the distribution of particles within the ash cloud. The ash mass loading is determined as:

$$m_l = \frac{4}{3}\rho \frac{r_e \tau(\lambda)}{Q_{ext}(\lambda, r_e)} \cos(\zeta), \tag{6}$$

where ρ is the ash particle density (set to $2500\ \mathrm{kgm^{-3}}$ based on field measurements reported by Dominguez et al. (2020) for

105 distal ash) and the $\cos(\zeta)$ term corrects the mass loading based on the satellite zenith angle. Uncertainties using this approach have been previously estimated to be up to 50% (Wen and Rose, 1994; Corradini et al., 2008). Our microphysical model, used to parameterize a volcanic ash cloud in the radiative transfer calculations, assumes that ash particles are spherical, composed of andesite and conform to a lognormal size distribution with a spread equal to 0.5 (geometric standard deviation $\sigma_g = 1.65$), similar to existing operational volcanic ash retrieval algorithms (e.g. Francis et al., 2012; Pavolonis et al., 2013).

110 The retrieval scheme relies on interpolating pre-computed LuTs generated by conducting radiative transfer calculations made for varying values of r_e, τ, T_s and T_c. The LuTs are generated using a new python implementation of the original FORTRAN code developed by Prata (1989b) to solve the radiative transfer equation for a single-layer ash cloud using the Discrete Ordinates Method (DOM; Stamnes et al., 1988; Laszlo et al., 2016). In the present study, we consider τ in the range from 0–9.9 in

steps of 0.1 and r_e from 1–15 μm in steps of 0.2 μm. All radiative transfer calculations use 16 radiation streams and a unique

115 LuT is generated for every combination of T_c and T_s identified from ash-affected pixels. Figure 3 shows a graphical example of a LuT generated for one combination of T_s and T_c and the range of τ and r_e considered.

To determine T_s directly from measurements it is generally recommended to find a clear-air pixel near the volcanic cloud of interest (e.g. Wen and Rose, 1994) and can sometimes be determined by finding the maximum value of T_B^{11} in the scene (Prata and Lynch, 2019). Obtaining an estimate for T_c from measurements, however, can be more difficult as the minimum value

120 of T_B^{11} may not correspond to the (semi-transparent) ash cloud of interest. Nevertheless, even if T_s and T_c can be reasonably estimated from measurements, it is often assumed that a single or mean value (and corresponding standard deviation) is representative of the entire ash cloud. Figure 1 shows that, in our case, the ash plume extends more than 60° in longitude and 20° in latitude, over land (including the Andes mountain ranges) and ocean, meaning that there is a considerable variation in cloud-top and surface temperature across ash-affected pixels. In addition, the meteorological setting within the considered spatial and

125 temporal domains is complex (significant amounts of clouds), making estimates of T_s and T_c from measurements challenging for the Cordón Caulle case study. To account for variation in T_s and T_c across space and time, we use ERA5 reanalysis data to determine T_s and T_c at every ash-affected pixel over our study period from 5–10 June 2011 (in one hour time steps).

To determine T_s from ERA5, we use the surface skin temperature (T_{skin}) and assume that the atmospheric transmittance (t_{atm}) has only a small effect on measured radiances at the top of the atmosphere for spit-window channels (*i.e.* $t_{atm} \approx 1$).

130 We also correct T_{skin} for variations in land surface emissivity using the University of Wisconsin global IR land surface emissivity database (Seemann et al., 2008). For ocean surfaces, we set the emissivity to 0.99 consistent with Western et al. (2015). Analysis comparing T_B^{11} SEVIRI measurements against the emissivity-corrected T_{skin} for clear-sky pixels on 4 June 2011 indicates average differences of ∼2 K. To determine T_c from ERA5 we require an estimate of the volcanic cloud-top height. A fortuitous CALIPSO pass early on during the eruption on 5 June 2011 reveals that the Cordón Caulle ash cloud reached as

135 high as 13–14 km above sea level (Fig. 4a) and later observations indicate heights from 10–13 km (Fig. 4b). For the retrievals presented here, we take T_c to be the ERA5 temperature at 13 km (a. s. l.) and make the simplifying assumption of constant height at all locations (and times) for every ash-affected pixel detected during 5–10 June 2011. The assumption of constant cloud-top height allows T_c to vary in time, horizontally but not vertically. However, *FALL3D-8.0* simulations indicate that the height of the Cordón Caulle ash cloud was relatively stable over the course of its dispersion in the atmosphere and so we expect

140 errors introduced by this assumption to be small. This was probably due to its injection into the stratosphere and its transport via the Southern hemisphere jet stream (height variations from 11–15 km; Klüser et al., 2013; Vernier et al., 2013; Prata et al., 2017). For our study period from 5–10 June 2011, T_c ranged from 206–226 K while T_s ranged from 230–304 K. We therefore performed radiative transfer calculations to construct unique LuTs, in steps of 2 K, for every possible combination of T_s and T_c within these ranges.

145 ## 2.2 Sulphur dioxide

Retrieval methods applicable to broadband IR satellite observations have largely focussed on exploiting SO_2 absorption features near 8.6 µm and 7.3 µm (e.g. Realmuto et al., 1994; Prata et al., 2003; Watson et al., 2004). To validate the new SO_2

scheme in *FALL3D-8.0*, we apply a three-channel technique to IR geostationary satellite measurements to retrieve total SO_2 column densities in Dobson Units (DU) (Prata et al., 2003; Prata and Kerkmann, 2007; Doutriaux-Boucher and Dubuisson, 2009). The three-channel technique exploits the strong SO_2 absorption feature near 7.3 μm and is only sensitive to upper-level SO_2 (> 4 km) due to the absorption of lower-level water vapor at this wavelength. We use the recent SO_2-rich eruption of Raikoke volcano in Russia during June 2019 as a validation case study and apply the SO_2 retrieval to observations made by the multispectral Advanced Himawari Imager (AHI) instrument aboard Himawari-8 geostationary satellite (Bessho et al., 2016).

2.2.1 SO_2 detection

The three channels used to detect SO_2 using AHI measurements are centred around 6.9, 7.3 and 11.2 μm. To determine whether there is an SO_2 signal in the data, we first construct a synthetic 7.3 μm brightness temperature by interpolating from 6.9 to 11.2 μm in the radiance space and then converting to brightness temperature via the Planck function (Prata et al., 2003). Figure 5 illustrates how the interpolation procedure works in radiance space. The resulting 'clear' brightness temperature ($T_{BC}^{7.3}$) is a good approximation of the measured value ($T_B^{7.3}$) in a SO_2-free atmosphere, so that one can identify SO_2 clouds by taking the difference between these two variables:

$$\Delta T_{SO2} = T_{BC}^{7.3} - T_B^{7.3} \tag{7}$$

In theory, ΔT_{SO_2} should be equal to zero under clear-sky conditions and increase with increasing SO_2 column density. However, in reality, high satellite zenith angles and variations in temperature and humidity can cause ΔT_{SO_2} to be positive even under clear-sky conditions (Prata et al., 2003; Doutriaux-Boucher and Dubuisson, 2009). To remove false positives due to high satellite zenith angles and high water vapour burdens, we compute two SO_2-related BTDs (ΔT_{69} and ΔT_{86}) and apply two successive temperature cut-off thresholds:

$$\Delta T_{69} = T_B^{6.9} - T_B^{7.3} > T_{69} \tag{8}$$

where only those pixels with a ΔT_{69} greater than a cut-off threshold of $T_{69} = -2.5$ K are flagged as potential SO_2. The second threshold takes advantage of the SO_2 absorption feature near 8.6 μm:

$$\Delta T_{86} = T_B^{11} - T_B^{8.6} > T_{86} \tag{9}$$

where only those pixels with a ΔT_{86} greater than a cut-off threshold of $T_{86} = 3.5$ K are flagged as potential SO_2.

In addition, the presence of meteorological clouds and embedded volcanic ash can also affect the interpolation procedure used to construct $T_{BC}^{7.3}$. Figures 6a to 6c show, respectively, $T_B^{7.3}$, $T_{BC}^{7.3}$, and ΔT_{SO_2} brightness temperatures for the SO_2-rich Raikoke cloud on 22 June 2019 at 21:00 UTC. Clearly, the interpolation procedure does a good job at removing the SO_2 signal from the measurements resulting in excellent detection sensitivly for ΔT_{SO_2}. Comparison of Figures 6d, e and f show how the ΔT_{69} and ΔT_{86} thresholds are used to remove false alarms whilst preserving legitimate SO_2-affected pixels.

Geoscientific
Model Development
Discussions

2.2.2 SO$_2$ retrieval

As mentioned above, ΔT_{SO_2} calculated via Eq. (7) is a function of the total column density of SO$_2$. The SO$_2$ retrieval is based on constructing this function from offline radiative transfer calculations. For this purpose we use the MODTRAN-

180 6.0 code (Berk et al., 2014) to compute top-of-the-atmosphere (TOA) radiances at the 7.3 μm wavelength (Fig. 2b). All radiances determined from MODTRAN-6.0 were convolved using the AHI spectral response functions. These radiances are then converted to brightness temperatures to compute BTDs between an SO$_2$-free atmosphere and atmospheres with varying column amounts of SO$_2$ at 7.3 μm (*i.e.* ΔT_{SO_2}). We are then able to generate a function representing the relationship between the SO$_2$ column density, $u(\Delta T_{SO_2})$, and ΔT_{SO_2} by interpolating between the data points generated from the radiative transfer

185 modelling (Fig. 8). In practice we generate this function using a 1D quadratic interpolation procedure implemented in the SciPy python package (Virtanen et al., 2020). Atmospheric profiles of temperature, humidity and gases were taken from the US standard atmosphere. In varying the SO$_2$ column amounts, we must specify an SO$_2$ profile. We use CALIPSO total attenuated backscatter profiles collocated with Himawari-8 observations of ΔT_{SO_2} to constrain the SO$_2$ profiles used in the radiative transfer calculations for the Raikoke case. When using the CALIPSO observations to determine the height and thickness of

190 SO$_2$ layers we make the assumption that SO$_2$ is collocated with sulphate aerosols (Carboni et al., 2016; Prata et al., 2017). Figure 7a shows a daytime CALIOP overpass intersecting SO$_2$ layers detected by Himawari-8 during the initial explosive phase of the Raikoke eruption. The vertical distribution of cloud/aerosol layers in the CALIOP observations reveal that the eastern part of the plume reached at least 12 km (a.s.l.). Later CALIOP/AHI observations reveal complex stratospheric dynamics; two distinct components are apparent in the attenuated backscatter data with thin layers (1–2 km) present at 13–15 km in the

195 northern part of the SO$_2$ cloud and ∼12 km in the southern part (Fig. 7b). Based on these initial observations, we constructed $u(\Delta T_{SO_2})$ using a uniform SO$_2$ distribution with a maximum cloud-top height of 13.5 km and thickness of 2.5 km (Fig. 8). The retrieval then proceeds by computing ΔT_{SO_2} from AHI data and evaluating $u(\Delta T_{SO_2})$ for every SO$_2$-affected pixel.

2.3 Data insertion

The data insertion scheme was recently introduced in *FALL3D-8.0* and is briefly described in Folch et al. (2020). Here we

200 describe the data insertion setup used for the Cordón Caulle and Raikoke case studies. To insert IR satellite retrievals of volcanic ash and SO$_2$ (described in Sects. 2.1.2 and 2.2.2) into FALL3D, the satellite retrievals were re-sampled (using nearest neighbour sampling) from their native projection into a regular $0.1° \times 0.1°$ latitude-longitude grid, consistent with the FALL3D grid. The vertical distribution must also be specified in the model as the satellite retrievals represent total column loadings (i.e. 2D spatial fields). For the cases presented here, CALIOP observations were used to constrain the vertical distribution of ash

205 and SO$_2$. Note that the vertical distribution is only required for the data insertion time. The data insertion times used for Cordón Caulle and Raikoke are 2011-06-05 15:00 UTC and 2019-06-22 18:00 UTC, respectively. The FALL3D configuration parameters for the simulations with and without data insertion are summarised in Table 1.

3 Validation metrics

Both qualitative and quantitative validation approaches have been used to validate previous versions of FALL3D against satel-
210 lite observations (Corradini et al., 2011; Folch et al., 2012). Here we use the SAL metric (Wernli et al., 2008) to quantitatively
compare satellite retrievals of volcanic ash (Sect. 2.1.2) and SO_2 (Sect. 2.2.2) to the corresponding simulations with and without
data insertion. The SAL metric was developed for validation of precipitation forecasts against radar and satellite data (Wernli
et al., 2008). Dacre (2011) demonstrated its use for validation air pollution simulations and Wilkins et al. (2016) employed SAL
for dispersion model validation against IR satellite volcanic ash retrievals for the 2010 Eyjafjallajökull eruption. More recently
215 SAL has been used to compare online vs. offline model simulations of volcanic ash (Marti and Folch, 2018). As in Wilkins
et al. (2016) and Marti and Folch (2018), we also use the FMS score as a complement to SAL for comparing the spatial cover-
age of observed vs. modelled fields. A detailed mathematical description of the SAL metric is presented in Wernli et al. (2008)
and so we only provide a brief description of each of the components of SAL (i.e. S, A and L) in the following subsections.
The main requirement for calculating SAL is the determination of model and observation *objects*. Objects are identified as
220 clusters of contiguous pixels whose magnitude is above some threshold corresponding to a physical quantity determined from
observations. In our case, the threshold is determined based on the detection limit of the satellite retrievals. For the SEVIRI
ash retrievals (Cordón Caulle case; Sect. 2.1.2), we use a threshold of 0.2 gm^{-2} consistent with the threshold suggested by
Prata and Prata (2012). For the Himawari-8 SO_2 retrievals (Raikoke case; Sect. 2.2.2), there is currently no commonly accepted
detection threshold. For the purposes of identifying SO_2 objects, we allowed for a threshold of 5 DU, noting that the minimum
225 detected SO_2 total column burdens at each time step in the satellite retrievals were ∼8–10 DU (after applying the threshold tests
defined in Sect. 2.2.1). The objects are determined for both observation (satellite retrievals) and model fields and are computed
as the absolute sum of S, A and L, which results in an index that varies from 0 (best agreement) to 6 (worst agreement). All
comparisons between observations and model simulations are made using a regular $0.1° \times 0.1°$ latitude-longitude grid.

3.1 Amplitude

230 The Amplitude (A) metric is the simplest of the three metrics used to construct SAL and compares the normalised difference
of the mass-averaged values of the observation and model fields. It can vary from -2 to +2 where negative (positive) values
indicate that the model is under-predicting (over-predicting) the mass when compared with observations.

3.2 Location

The Location (L) metric has two components ($L = L_1 + L_2$). L_1 is calculated as the distance between the centers of mass between
235 the model and observation fields, normalised by the maximum distance across the specified domain. It can vary from 0 to +1
and is considered a first-order indication of the accuracy of the model simulation compared with observations. However, L_1 can
equal 0 (suggesting perfect agreement) for situations where observation and model fields clearly do not agree. For example,
Wernli et al. (2008) describe the case of two objects at opposite sides of the domain having the same center of mass as a single
object in the center of the domain. L_2 was introduced to handle these situations by considering the weighted average distance

240 between the overall center of mass and the center of mass of each individual object for both model and observation fields. L_2 is computed by taking the normalised difference between the model weighted average distance and the observation weighted average distance. It is scaled such that it varies from 0 to +1 (to vary over the same range as L_1), meaning that L varies in the range from 0 to +2.

3.3 Structure

245 The Structure (S) metric is the most complex of the three metrics used to construct SAL. The general idea is to compute the normalised 'volume' of all individual objects for each dataset (i.e. the model and observation fields). The normalised volumes are computed by dividing the total (sum) mass of each object by its maximum mass. The weighted mean of the normalised volumes is then computed for the observation and model fields and S is computed by taking the difference between the weighted means. The S metric can vary from -2 to +2, where negative values indicate that modelled objects are too small or too peaked 250 (or a combination of both) compared to the observed fields.

3.4 Figure of Merit in Space

The FMS score compares the spatial coverage of observed vs. modelled fields. It is simply the area of intersection divided by the area of union between the ash mass loading observation and model fields:

$$\text{FMS} = \frac{A_{\text{mod}} \cap A_{\text{obs}}}{A_{\text{mod}} \cup A_{\text{obs}}}, \tag{10}$$

255 where A_{mod} and A_{obs} are the modelled and observed ash mass loading areas, respectively. The FMS varies from 0 (no intersection) to 1 (perfect overlap).

4 Validation cases

4.1 The 2011 Cordón Caulle fine ash cloud

The Puyehue-Cordón Caulle volcanic complex (PCCVC), located in the southern volcanic zone of the central Andes, comprises 260 a 20 km long, NW-SE oriented fissure system (Cordón Caulle) and the Puyehue stratovolcano (Elissondo et al., 2016). On 4 June at around 14:45 LT (18:45 UTC), a new vent opened at 7 km NNW from the Puyehue volcano (Collini et al., 2013), initiating a remarkable example of a long-lasting plume with complex dynamics, strongly influenced by the interplay between eruptive style, atmospheric winds and deposit erosion (Bonadonna et al., 2015). The initial explosive phase of the eruption (4-14 June) was characterised by the development of eruption columns with heights oscillating between 6-14 km above sea level 265 (a.s.l.). Plume heights progressively decreased (4-6 km a.s.l) between 15 and 30 June, and low intensity ash emission persisted for several months (Elissondo et al., 2016). Due to the predominant westerly winds, ash was transported towards Argentina and a wide area of the arid and semi-arid regions of northern Patagonia was severely affected by tephra dispersal and fallout. The eruption had multiple impacts on the ecosystem, critical infrastructures, human activities and several economic sectors (e.g.

agriculture, aviation, and tourism; Wilson et al., 2013).

270 Figures 9 and 10 compare satellite retrievals and model simulations at 24, 48, and 72 h after starting time for runs with and without data insertion, respectively. The time-series of each validation metric (for Cordón Caulle and Raikoke) are also shown in Fig. 11 and summarised in Table 2. Comparison of Figs. 9a and 10a highlight the advantage of a data insertion scheme. For the simulation without data insertion (Fig. 10a), the plume has already begun to deviate from the satellite observations with too much mass dispersing towards the south. This is reflected in both the SAL score of 1.93 and FMS score of 0.22 at this time.

275 The data insertion scheme (Fig. 9a) naturally corrects for this by taking advantage of good quality satellite observations of the vertical and horizontal structure of the Cordón Caulle ash plume at this time. For the data insertion simulations, FALL3D accurately represents the spatial structure of the satellite retrievals after 24 hours with a SAL score of 1.3 and FMS of 0.42 (Fig. 9b; see Supplementary Material for the full animation of the data insertion simulations). In addition, the accuracy of the simulations over the first 24 hours shows a marked improvement when compared to the simulations without data insertion

280 (Fig. 10b; SAL = 1.84; FMS = 0.32). The validation metric time-series show this in more detail (Figs. 11a and c). For the simulations without data insertion, the SAL score remains above 2 for most of the first 24 hours while the SAL gradually increases from 0 to 1.3 for the simulation with data insertion. Comparison of Figs. 9b and 10b shows that the data insertion simulation is better able to capture the northern portion of the plume than the simulation without data insertion at this time (an increase in the FMS by 0.1). Inspection of the time-series of the individual validation metrics for the simulations with data

285 insertion (Fig. 11a) reveals that the SAL is largely being affected by increases in the L metric (i.e. increases in the distance between the centres mass between the observation and model fields) and decreases in the A metric (model under-predicting mass compared to observations). The S metric only exhibits minor deviations when compared to the observations during the first 24 hours after data insertion. At 48 hours, the simulations with and without the data insertion are almost identical (minor differences in the modelled ash contours near 30° S, 15° W). This is because, at this time, almost all of the ash used in

290 the data insertion has exited the domain. For the simulations without data insertion, the SAL score is 1.2 and FMS is 0.14 (Fig. 9c; Table 2); however, at around 36 hours the SAL reached above 2 and then decreased sharply (Fig. 11c). The reason for the sudden reduction in SAL just after 36 hours is most likely due to the satellite retrievals being compromised by cloud interference at this time in addition to the continual input of mass at the source in the model simulations. This input of mass was included to account for ash erupted after the data insertion time. The satellite retrievals capture some of the ash plume near

295 source (Figs. 9c and 10c), but cannot be expected to accurately characterise the plume at this location due to its high opacity in the IR window. Another difference between the model and observations at this time is the large difference in the centres of mass (L = 0.32). This is due to the high mass loadings near source in the model fields and high mass loadings near the centre of the domain (43° S, 35° W) in the observed fields. The satellite is likely over-estimating mass in this part of the ash cloud because of underlying meteorological clouds that have not been accounted for in the radiative transfer modelling. After 72

300 hours (Fig. 9d), the simulations with and without data insertion are identical as all ash used in the data insertion scheme has exited the domain by this time. At this time the SAL has reached a score of 1.71 and the FMS has decreased to 0.10, reflecting the fact that the model has continued to deviate from the observations with time.

Geoscientific
Model Development
Discussions

4.2 The 2019 Raikoke SO$_2$ cloud

On 21 June 2019, a small island volcano, Raikoke (48.292° N, 153.25° E, 551 m a.s.l.), underwent a significant explosive
305 eruption disrupting major aviation flight routes across the North Pacific. Raikoke is located in the central Kuril Islands, a
remote island chain that lies south of Russia's Kamchatka peninsula. Ground-based networks are sparse in this area and so
satellite observations were crucial for tracking the volcanic ash and SO$_2$ produced by the eruption. The eruption sequence was
characterised by a series of ∼9 'pulses', injecting ash and gases into the atmosphere. The International Space Station captured
a unique view of the eruption's umbrella plume during its initial explosive phase which was reminiscent of the 2009 Sarychev
310 Peak umbrella plume (https://earthobservatory.nasa.gov/images/145226/raikoke-erupts). The eruption sequence was captured
extremely well by the Himawari-8 satellite at both IR and visible wavelengths. According to our analysis of the satellite data,
the initial explosive phase began at around 18:00 UTC on 21 June (05:00 LT on 22 June at around sunrise) and ended at
around 10:00 UTC on 22 June (21:00 LT just before sunset). The Smithsonian Institution's Global Volcanism Program (GVP)
report on the 2019 Raikoke eruption also documents less intense activity at the volcano from 23–25 June following the initial
315 explosive phase (Global Volcanism Program, 2019).

During the initial explosive phase on 21 June 2019, a significant amount of SO$_2$ was injected into the atmosphere making
the eruption an ideal case to study long-range SO$_2$ transport and dispersion. Preliminary analysis of TROPOMI SO$_2$ retrievals
indicated that ∼1.4–1.5 Tg SO$_2$ was injected into the atmosphere (Global Volcanism Program, 2019). Hyman and Pavolonis
(2020) present SO$_2$ retrievals based on Cross-track Infrared Sounder (CrIS) measurements and show a time-series of SO$_2$ total
320 mass with a peak between 1–1.1 Tg SO$_2$. The maximum total mass recorded by the Himawari-8 retrievals presented here
indicate ∼1.4 Tg SO$_2$. However, it should be noted that the retrievals presented here are preliminary and require further cross-
validation with other satellite retrievals in addition to an analysis of the uncertainty on these retrievals, which is beyond the
scope of the present study.

Figures 12 and 13 show a comparison between the satellite retrievals and model simulations with and without data insertion,
325 respectively, in addition to the SAL and FMS validation metrics. The time-series of validation metrics for the Raikoke case
study are shown in Figs. 11b and 11d and are summarised in Table 2. We selected a data insertion time of 22 June 2019 at 18:00
UTC (1 day after the beginning of the eruption) as this is a time when the SO$_2$ cloud was completely detached from source
(Fig. 12a). Note that the AHI retrievals of the SO$_2$ plume at the beginning of the eruption (Fig. 7a) were likely compromised
by interference of ice particles in the initial eruption plume (Prata et al., 2003; Doutriaux-Boucher and Dubuisson, 2009). In
330 addition, retrievals early on in the plume's dispersion may have been affected by band saturation caused by extremely high SO$_2$
column loads.

At the data insertion time (22 June 2019 at 18:00 UTC), the SAL score for the FALL3D SO$_2$ simulation without data
insertion is 2.87 and the FMS is 0.32. Therefore, applying data insertion at this time represents a significant correction of
the model simulation to the satellite observations (compare Fig. 12a and Fig. 13a). The main difference between the satellite
335 observations and simulation without data insertion is that the model indicates a portion of the SO$_2$ plume connecting back to
the volcano while this feature is not present in the observations. TROPOMI observations of the SO$_2$ cloud confirm this spatial

structure (see Fig. 13 of Global Volcanism Program, 2019). The reason for the lack of detection of SO_2 in this region in the AHI retrievals is probably due to water vapour interference, implying that this part of the plume was at lower altitudes than the main SO_2 cloud. Indeed, SO_2 height retrievals from CrIS data show that plume heights varied from $\sim$3–7 km a.s.l. in this
340 region (see Fig. 5 of Hyman and Pavolonis, 2020).

For the simulations without data insertion, at 24 hours after insertion, the validation metrics exhibit minor changes with SAL decreasing from 2.87 to 2.59 and the FMS from 0.32 to 0.23 (Fig. 13a, b). For the simulations with data insertion, SAL has steadily increased from 0 to 1.21 while the FMS has decreased from 1 to 0.29 over the first 24 hours (Fig. 12a, b; see Supplementary Material for the full animation of the data insertion simulations). Figure 11b shows that the SAL score for the
345 simulation with data insertion is largely affected by the S and A scores whereas the L score is low (0.05) indicating the FALL3D is able to track the centre of mass of SO_2 very well when initialised with satellite retrievals (Fig. 11b). In this case the A metric is negative, meaning that the model is under-predicting the mass when compared to the satellite retrievals. This is probably due to the fact that the total mass retrieved by the satellite actually increases after the data insertion time. An increase in SO_2 mass cannot be accounted for in the data insertion scheme if no new sources of SO_2 are included in the model simulations. A reason
350 for an increase in mass in the satellite retrievals, even after the SO_2 cloud has detached from source, could be due to several factors related to the detection sensitivity of the retrieval. An interesting physical reason for the increase in SO_2 could be that ice particles in the nascent plume were sequestering SO_2 initially and then releasing it later on as the plume dispersed into the atmosphere (Rose et al., 2001). Increases in the satellite-retrieved SO_2 mass can also be due to the horizontal and vertical distribution of water vapour. For example, if the SO_2 cloud is in a region of high water vapour initially and then moves into a
355 drier region it is likely that more SO_2 will be detected thus increasing the retrieved mass. This effect can also occur if the SO_2 layers are transported vertically in the atmosphere.

At 48 hours, for the simulations without data insertion (Fig. 13c), the SAL score actually improves (decreasing from 2.58 to 1.88) and the FMS largely remains the same (decreasing from 0.23 to 0.20). The improvement in the SAL score can be attributed to a steady increase in S metric and decrease in the A metric (Fig. 11d). This indicates that the structure and mass
360 (amplitude) modelled by the simulations without data insertion are converging towards that observed by the satellite over 48 hours. For the simulations with data insertion, at 48 hours after insertion (Fig. 12c), the SAL score has continued to increase (from 1.21 to 1.38) and the FMS has continued to decrease (from 0.29 to 0.25). In general, at all lead times, the validation metrics indicate that the data insertion simulations provide better agreement with observations than the simulations without data insertion (Table 2).

365 ### 4.3 The 2013 Mt. Etna tephra deposit

On 23 February 2013 at 18:15 UTC, the eruptive activity of Mt. Etna increased significantly. A buoyant plume rising up to 9 km above sea level (a.s.l.) along with incandescent lava fountains exceeding 500 m above the crater were generated during the paroxysmal phase (Poret et al., 2018), resulting in an ash plume extending towards the NE for more than 400 km away from the source and moderate ash fallout over the Italian regions of Calabria and Puglia.
370 In order to simulate this event using high-resolution wind fields, we ran first the ARW (Advanced Research WRF) core of

the WRF (Weather Research and Forecasting) model (Skamarock et al., 2008) on a single-domain configuration consisting of 700×700 grid points with $4\,\mathrm{km}$ horizontal resolution and 100 vertical levels with a maximum height of $50\,\mathrm{hPa}$. The initial and boundary conditions for WRF/ARW were extracted from hourly ERA5 reanalysis data, with a spatial resolution of around $0.25°$ and 137 vertical model levels.

375 A FALL3D run was initialised with a start time of 18:00 UTC on 23 February with a uniform source distribution (TOP-HAT option) reaching $5.5\,\mathrm{km}$ above the vent (*i.e.* $8.7\,\mathrm{km}$ a.s.l) and a thickness of $3.5\,\mathrm{km}$. The model was configured with a horizontal resolution of $0.015°$ and 60 vertical levels up to $11\,\mathrm{km}$ in a computational domain comprising all deposit sampling locations. The particle Total Grain Size Distribution (TGSD) was discretised into 32 bins (diameter d in the range $-6\Phi \leq d \leq 6\Phi$) with densities varying between 1000 and $2500\,\mathrm{kg\,m^{-3}}$ for coarser and finer bins respectively, and a constant sphericity of 0.92. We

380 considered a bi-Gaussian (in Φ) TGSD following Poret et al. (2018), who performed numerical simulations with subpopulation means of $\mu_c = -2.96$, $\mu_f = 0.49$, and standard deviations of $\sigma_c = 1.03$, $\sigma_f = 0.79$. However, as already noted by Poret et al. (2018), this distribution underestimates the fine ash fraction. In order to correct this drawback, the simulation was run with a fine-enriched bi-Gaussian TGSD given by

$$f(\Phi) = \frac{p}{\sigma_c\sqrt{2\pi}}\exp\left(-\frac{(\Phi-\mu_c)^2}{2\sigma_c^2}\right) + \frac{1-p}{\sigma_f\sqrt{2\pi}}\exp\left(-\frac{(\Phi-\mu_f)^2}{2\sigma_f^2}\right) \qquad (11)$$

385 with the mean and standard deviation for the fine subpopulation given, respectively, by $\mu_f = 2.54$ and $\sigma_f = 0.38$, and the coarse subpopulation was defined by the parameters $\mu_c = -2.96$ and $\sigma_c = 1.03$. The fractions of each subpopulation are p and $1-p$. In this work, $p = 0.7$ was used. Table 1 summarises the rest of model configuration options, and the resulting tephra ground load map is shown in Fig. 14.

In order to validate the *FALL3D-8.0* deposit we compared the simulations results to the observations reported by Poret et al.

390 (2018), which consist of deposit load and local grain-size distribution samples at 10 locations (S1-S10). Proximal sites (S1-S7) are located between 5 and $16\,\mathrm{km}$ from the vent, whereas the rest of samples (S8-S10) correspond to the locations of Messina (Sicily, S8), Cardinale (Calabria, S9) and Brindisi (Puglia, S10), the latter located at about $410\,\mathrm{km}$ from the volcano.

Figure 15a compares modelled and observed tephra loading at all sites. Note that all points lie within a factor 3 error band and, remarkably, that a perfect agreement (black solid line) is found across four orders of magnitude (from $10^{-3}\,\mathrm{kg\,m^{-2}}$ to more

395 than $10\,\mathrm{kg\,m^{-2}}$). This good agreement is also observed at a bin level in the local grain size distributions. To illustrate this, Fig. 15b compares computed and observed particle distribution modes at all sites (S01-S10). It should be noted that the model predicts unimodal distributions at all sampling sites in good agreement with field observations.

4.4 The 1986 Chernobyl nuclear accident

One of the most serious nuclear accident on the Earth occurred on 25 April 1986 at 21:23 UTC at the Chernobyl Nuclear

400 Power Plant (NPP) in Ukraine. Two explosions in the NPP dispersed the radioactive material into the atmosphere where it was transported by the winds up to distances of thousands of kilometers away from the NPP.

In order to simulate the dispersion of the radioactive material an estimation of the emission rate from the Chernobyl NPP is needed. Unfortunately, estimations of such a source term is still very uncertain and they rely on the solution of an inverse

problem aimed to reproduce the available measurements in the region of interest at the time of the accident. However, on the

405 basis of high quality deposition measurements from the Radioactivity Environmental Monitoring (REM) database (De Cort et al., 2007), Brandt et al. (2002) reconstructed the source term. Here we used the source term reported by Brandt et al. (2002) as input for the FALL3D simulations (see Code and data availability statement to access the corresponding input files).

Particle size distributions and settling velocities of radioactive material, such as ^{134}Cs, ^{137}Cs, and ^{131}I, released during the accident are uncertain and their estimations also complicated by the interaction with other atmospheric particles and aerosols

410 (Brandt et al., 2002). Effective settling velocities range from 0.0005 to 0.005 m/s for ^{137}Cs and from 0.001 to 0.02 m/s for ^{131}I (Brandt et al., 2002). Considering these ranges and discretizing velocities in 4 classes (see Table 3) we chose the effective classes and fractions through the best fit of the simulations results with the radioactivity values measured on 10 May 1986. FALL3D simulations were carried out on the computational domain shown in Fig. 16 for the period from 24 April to 10 May 1986, considering the input values reported in Table 3 and the meteorological fields obtained from ERA5 reanalysis

415 (Copernicus Climate Change Service (C3S), 2017) and accounting for atmospheric diffusion, wet deposition and radioactive decays (see Supplementary Material). The comparison of measured and simulated values for the best case (Table 3) is reported in Fig. 16. We can see that most of simulated values are within an order of magnitude of the measurements (Fig. 17). Simulations results of the radioactive cloud evolution relative to ^{137}Cs (vertically integrated radioactivity concentration in the atmosphere, expressed in Bq/m^2) from 28 April to 9 May, 1986, are reported in Fig. 18. We can clearly see that simulations

420 correctly reproduced the patterns described by Brandt et al. (2002). The evolution of the ^{137}Cs dispersal is also available as a video in the Supplementary Material, together with videos corresponding to the dispersal of ^{134}Cs and ^{131}I.

5 Conclusions

Four different examples from the new *FALL3D-8.0* benchmark suite have been presented to validate the accuracy of the lastest major version release of the *FALL3D* model and complement a companion paper (Folch et al., 2020) on model physics and

425 performance.

In the first two examples (*i.e.* far-range fine ash dispersal from the 2011 Cordón Caulle eruption and SO$_2$ cloud dispersal from the 2019 Raikoke event), collocation of geostationary Meteosat-9 (SEVIRI) and Himawari-8 (AHI) satellite observations with polar orbiting CALIPSO (CALIOP lidar) passes were used to retrieve column mass loads and characterise the vertical and horizontal structure of these volcanic clouds at selected data insertion times. Furthermore, new ash and SO$_2$ detection

430 schemes based on applying successive masks have been proposed. For the volcanic ash case study, the new scheme enhances ash detection in complex scenes (e.g. numerous clouds, ice-covered surfaces) and accounts for high satellite zenith angles (i.e. at the edges of the SEVIRI field-of-view). Simulations with and without the satellite data insertion option have been compared with 1-hourly satellite retrievals using the SAL and the FMS metrics. The time series of the validation metrics (Fig. 11) illustrate how simulations deviate from observations with time. According the SAL and FMS metrics, we find that

435 simulations initialised with data insertion consistently outperform simulations without data insertion. In general, it was found

that observations (satellite) are patchier when compared to the smoother model fields, resulting in consistently more objects being detected in the observation fields, consistent with the findings of Wilkins et al. (2016).

It is not yet clear what absolute values of SAL and FMS should represent an acceptable forecast. For the data insertion simulations presented here, SAL remained below 1 out to 18 h (Table 2) and below 2 at lead times of 24 and 48 h for both ash
440 and SO_2 simulations. It is unlikely that in an operational setting a model simulation would be relied upon beyond 48 hours. Ideally, in an operational setting, the model should be re-run with an updated data insertion time when new, good quality satellite retrievals become available. From a qualitative perspective it appears that SAL values of less than 1.5 and FMS values of $\sim$0.40 indicate good spatial agreement between the model and observation fields (see Fig. 9b). However, it's also important to consider that the satellite retrievals can be affected by cloud interference, meaning that the ash/SO_2 detection schemes
445 may miss some legitimate ash or SO_2 that the model is otherwise predicting (see Fig. 9c). A data assimilation scheme that considers the errors in the satellite retrievals in addition to errors in the model simulations (constructed based on an ensemble for example) can be used to resolve these issues (e.g. Fu et al., 2017; Pardini et al., 2020).

For the ground deposit load validation case study of the February 2013 Mt Etna eruption, we find exceptionally good agreement between field observations and FALL3D simulations of deposit loads. Acceptable ratios of model to observed ash
450 loading (between 1:3 and 3:1) were found across 4 orders of magnitude, i.e., from $10^{-3}\,\mathrm{kgm^{-2}}$ to more than $10\,\mathrm{kgm^{-2}}$. Good agreement (between model and field observations) in terms of the mode (ϕ) of distributions was also found at the majority of field sampling sites (Fig. 14b).

For the radionuclides validation, we found that the model has a very good performance in reproducing the observations of the dispersal of the radioactive cloud after the 22 April Chernobyl accident, similarly to other models in the literature (e.g.
455 Brandt et al., 2002).

From the point of view of model performance, further expected improvements in the preparation of *FALL3D* towards Exascale include memory optimisation, introduction of thread parallelism (OpenMP), code vectorisation, porting to accelerators (GPUs), performance portability, load balance, asynchronous I/O and preparation for emerging heterogeneous architectures (Exascale hardware prototypes). In terms of model utilities, next steps are to consider ensemble forecasts and to incorporate
460 data assimilation.

Code and data availability. *FALL3D-8.0* is available under the version 3 of the GNU General Public License (GPL) at https://gitlab.com/fall3d-distribution/v8.0. The satellite retrieval files for the 2011 Cordón Caulle and 2019 Raikoke case studies are also available for download on the the GitLab repository along with input files used to generate the corresponding simulations. Input files for the 2013 Etna and 1986 Chernobyl case studies are also accessible via the GitLab repository.

465 *Author contributions.* AF and AC conceived the study and planned the test cases. AF and LM have written the bulk of *FALL3D-8.0* code, with contributions of GM. ATP conducted the satellite retrievals and implemented the validation metrics. LM, AF and GM have run model executions and validations. All authors have contributed to the writing of the text.

Competing interests. The authors declare no competing interests.

References

Berk, A., Conforti, P., Kennett, R., Perkins, T., Hawes, F., and van den Bosch, J.: MODTRAN6: a major upgrade of the MODTRAN
480   radiative transfer code, in: Proceedings Volume 9088, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral
Imagery XX, edited by Velez-Reyes, M. and Kruse, F. A., p. 90880H, Baltimore, Maryland, USA, https://doi.org/10.1117/12.2050433,
http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2050433, 2014.

Bessho, K., Date, K., Hayashi, M., Ikeda, A., Imai, T., Inoue, H., Kumagai, Y., Miyakawa, T., Murata, H., Ohno, T., Okuyama, A., Oyama,
R., Sasaki, Y., Shimazu, Y., Shimoji, K., Sumida, Y., Suzuki, M., Taniguchi, H., Tsuchiyama, H., Uesawa, D., Yokota, H., and Yoshida, R.:
485   An Introduction to Himawari-8/9–Japan's New-Generation Geostationary Meteorological Satellites, Journal of the Meteorological Society
of Japan. Ser. II, 94, 151–183, https://doi.org/10.2151/jmsj.2016-009, https://www.jstage.jst.go.jp/article/jmsj/94/2/94_2016-009/_article,
2016.

Bonadonna, C., Pistolesi, M., Cioni, R., Degruyter, W., Elissondo, M., and Baumann, V.: Dynamics of wind-affected volcanic plumes: The
example of the 2011 Cordón Caulle eruption, Chile, J. Geophys. Res. Solid Earth, 120, 2242–2261, https://doi.org/10.1002/2014JB011478,
490   2015.

Brandt, J., Christensen, J. H., and Frohn, L.: Modelling transport and deposition or caesium and iodine from the Chernobyl accident using
the DREAM model, Atm. Chem. Phys., 2, 397–417, 2002.

Carboni, E., Grainger, R. G., Mather, T. A., Pyle, D. M., Thomas, G. E., Siddans, R., Smith, A. J. A., Dudhia, A., Koukouli, M. E., and
Balis, D.: The vertical distribution of volcanic SO_2 plumes measured by IASI, Atmospheric Chemistry and Physics, 16, 4343–4367,
495   https://doi.org/10.5194/acp-16-4343-2016, https://www.atmos-chem-phys.net/16/4343/2016/, 2016.

Collini, E., Osores, S., Folch, A., Viramonte, J., Villarosa, G., and Salmuni, G.: Volcanic ash forecast during the June 2011 Cordón Caulle
eruption, Natural Hazards, 66, 389–412, https://doi.org/10.1007/s11069-012-0492-y, https://doi.org/10.1007/s11069-012-0492-y, 2013.

Copernicus Climate Change Service (C3S): ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate, Copernicus
Climate Change Service Climate Data Store (CDS), Accessed 3 Feb 2020, https://cds.climate.copernicus.eu/cdsapp#!/home, 2017.

500 Corradini, S., Spinetti, C., Carboni, E., Tirelli, C., Buongiorno, M., Pugnaghi, S., and Gangale, G.: Mt. Etna tropospheric ash retrieval and
sensitivity analysis using moderate resolution imaging spectroradiometer measurements, Journal of Applied Remote Sensing, 2, 023 550,
https://doi.org/10.1117/1.3046674, http://remotesensing.spiedigitallibrary.org/article.aspx?doi=10.1117/1.3046674, 2008.

Corradini, S., Merucci, L., and Folch, A.: Volcanic Ash Cloud Properties: Comparison Between MODIS Satellite Retrievals and
FALL3D Transport Model, IEEE Geoscience and Remote Sensing Letters, 8, 248–252, https://doi.org/10.1109/LGRS.2010.2064156,
505   http://ieeexplore.ieee.org/document/5567133/, 2011.

Costa, A., Macedonio, G., and Folch, A.: A three-dimensional Eulerian model for transport and deposition of volcanic ashes, Earth and
Planetary Science Letters, 241, 634 – 647, https://doi.org/http://dx.doi.org/10.1016/j.epsl.2005.11.019, 2006.

Dacre, H.: A new method for evaluating regional air quality forecasts, Atmospheric Environment, 45, 993–1002,
https://doi.org/10.1016/j.atmosenv.2010.10.048, https://linkinghub.elsevier.com/retrieve/pii/S1352231010009349, 2011.

510 De Cort, M., Sangiorgi, M., Hernandez, C., Miguel, A., Vanzo, S., Nweke, E., Tognoli, P. V., and Tollefsen, T.: REM data bank - Years
1984-2006, [dataset], European Commission, Joint Research Centre (JRC), https://doi.org/10.2905/jrc-10117-10024, http://data.europa.
eu/89h/jrc-10117-10024, 2007.

Dominguez, L., Bonadonna, C., Forte, P., Jarvis, P. A., Cioni, R., Mingari, L., Bran, D., and Panebianco, J. E.: Aeolian Remobilisation of
the 2011-Cordón Caulle Tephra-Fallout Deposit: Example of an Important Process in the Life Cycle of Volcanic Ash, Front. Earth Sci., 7,
515   343, https://doi.org/10.3389/feart.2019.00343, 2020.

Doutriaux-Boucher, M. and Dubuisson, P.: Detection of volcanic SO2 by spaceborne infrared radiometers, Atmospheric Research, 92, 69–79,
https://doi.org/10.1016/j.atmosres.2008.08.009, 2009.

Elissondo, M., Baumann, V., Bonadonna, C., Pistolesi, M., Cioni, R., Bertagnini, A., Biasse, S., Herrero, J. C., and Gonzalez, R.: Chronology
and impact of the 2011 Cordón Caulle eruption, Chile, Nat. Hazards Earth Syst. Sci., 16, 675–704, 2016.

520   Folch, A., Costa, A., and Macedonio, G.: FALL3D: A Computational Model for Transport and Deposition of Volcanic Ash, Comput. Geosci.,
35, 1334–1342, https://doi.org/10.1016/j.cageo.2008.08.008, 2009.

Folch, A., Costa, A., and Basart, S.: Validation of the FALL3D ash dispersion model using observations of the 2010 Eyjafjallajokull vol-
canic ash clouds, Atmospheric Environment, 48, 165–183, https://doi.org/https://doi.org/10.1016/j.atmosenv.2011.06.072, http://www.
sciencedirect.com/science/article/pii/S1352231011006960, volcanic ash over Europe during the eruption of Eyjafjallajokull on Iceland,
525   April-May 2010, 2012.

Folch, A., Mingari, L., Gutierrez, N., Hanzich, M., Macedonio, G., and Costa, A.: FALL3D-8.0: a computational model for atmospheric
transport and deposition of particles, aerosols and radionuclides – Part 1: Model physics and numerics, Geoscientific Model Development,
13, 1431–1458, https://doi.org/10.5194/gmd-13-1431-2020, https://www.geosci-model-dev.net/13/1431/2020/, 2020.

Francis, P. N., Cooke, M. C., and Saunders, R. W.: Retrieval of physical properties of volcanic ash using Meteosat: A case study from
530   the 2010 Eyjafjallajökull eruption, Journal of Geophysical Research: Atmospheres, 117, https://doi.org/10.1029/2011JD016788, http:
//doi.wiley.com/10.1029/2011JD016788, 2012.

Fu, G., Prata, F., Lin, H. X., Heemink, A., Segers, A., and Lu, S.: Data assimilation for volcanic ash plumes using a satellite obser-
vational operator: a case study on the 2010 Eyjafjallajökull volcanic eruption, Atmospheric Chemistry and Physics, 17, 1187–1205,
https://doi.org/10.5194/acp-17-1187-2017, https://www.atmos-chem-phys.net/17/1187/2017/, 2017.

535   Galmarini, S., Bonnardot, F., Jones, A., Potempski, S., Robertson, L., and Martet, M.: Multi-model vs. EPS-based ensemble atmospheric
dispersion simulations: A quantitative assessment on the ETEX-1 tracer experiment case, Atmospheric Environment, 44, 3558–3567,
https://doi.org/10.1016/j.atmosenv.2010.06.003, https://linkinghub.elsevier.com/retrieve/pii/S1352231010004528, 2010.

Global Volcanism Program: Report on Raikoke (Russia), in: Bulletin of the Global Volcanism Network, vol. 44:8, Smithsonian Institution,
2019.

540   Gu, Y., Rose, W. I., Schneider, D. J., Bluth, G. J. S., and Watson, I. M.: Advantageous GOES IR results for ash mapping at high lat-
itudes: Cleveland eruptions 2001, Geophysical Research Letters, 32, https://doi.org/10.1029/2004GL021651, http://doi.wiley.com/10.
1029/2004GL021651, 2005.

Hyman, D. M. and Pavolonis, M. J.: Probabilistic retrieval of volcanic SO2 layer height and cumulative mass loading using the Cross-track
Infrared Sounder (CrIS), Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2020-41, https://www.atmos-meas-tech-discuss.net/
545   amt-2020-41/, 2020.

Klüser, L., Erbertseder, T., and Meyer-Arnek, J.: Observation of volcanic ash from Puyehue–Cordón Caulle with IASI, Atmospheric Mea-
surement Techniques, 6, 35–46, https://doi.org/10.5194/amt-6-35-2013, https://www.atmos-meas-tech.net/6/35/2013/, 2013.

Kylling, A., Kahnert, M., Lindqvist, H., and Nousiainen, T.: Volcanic ash infrared signature: porous non-spherical ash particle shapes com-
pared to homogeneous spherical ash particles, Atmospheric Measurement Techniques, 7, 919–929, https://doi.org/10.5194/amt-7-919-
550   2014, https://www.atmos-meas-tech.net/7/919/2014/, 2014.

Laszlo, I., Stamnes, K., Wiscombe, W. J., and Tsay, S.-C.: The Discrete Ordinate Algorithm, DISORT for Radiative Transfer, in: Light Scattering Reviews, Volume 11, edited by Kokhanovsky, A., pp. 3–65, Springer Berlin Heidelberg, Berlin, Heidelberg, https://doi.org/10.1007/978-3-662-49538-4_1, http://link.springer.com/10.1007/978-3-662-49538-4_1, 2016.

Marti, A. and Folch, A.: Volcanic ash modeling with the NMMB-MONARCH-ASH model: quantification of offline modeling errors, Atmo-
555 spheric Chemistry and Physics, 18, 4019–4038, https://doi.org/10.5194/acp-18-4019-2018, https://www.atmos-chem-phys.net/18/4019/2018/, 2018.

Pardini, F., Corradini, S., Costa, A., Esposti Ongaro, T., Merucci, L., Neri, A., Stelitano, D., and de' Michieli Vitturi, M.: Ensemble-Based Data Assimilation of Volcanic Ash Clouds from Satellite Observations: Application to the 24 December 2018 Mt. Etna Explosive Eruption, Atmosphere, 11, 359, https://doi.org/10.3390/atmos11040359, https://www.mdpi.com/2073-4433/11/4/359, 2020.

560 Pavolonis, M. J., Heidinger, A. K., and Sieglaff, J.: Automated retrievals of volcanic ash and dust cloud properties from upwelling infrared measurements, Journal of Geophysical Research: Atmospheres, 118, 1436–1458, https://doi.org/10.1002/jgrd.50173, http://doi.wiley.com/10.1002/jgrd.50173, 2013.

Platt, C. and Prata, A.: Nocturnal effects in the retrieval of land surface temperatures from satellite measurements, Remote Sensing of Environment, 45, 127–136, https://doi.org/10.1016/0034-4257(93)90037-X, https://linkinghub.elsevier.com/retrieve/pii/003442579390037X, 565 1993.

Poret, M., Costa, A., Andronico, D., Scollo, S., Gouhier, M., and Cristaldi, A.: Modeling Eruption Source Parameters by Integrating Field, Ground-Based, and Satellite-Based Measurements: The Case of the 23 February 2013 Etna Paroxysm, J. Geophys. Res. Solid Earth, 123, 5427–5450, https://doi.org/10.1029/2017JB015163, 2018.

Potts, R. and Ebert, E.: On the detection of volcanic ash in NOAA AVHRR infrared satellite imagery, in: 8th Australasian Remote Sensing
570 Conference, pp. 25–29, 1996.

Prata, A., O Brien, D., Rose, W., and Self, S.: Global, long-term sulphur dioxide measurements from TOVS data: A new tool for studying explosive volcanism and climate, Geophysical Monograph - American Geophysical Union, 139, 75–92, 2003.

Prata, A. J.: Observations of volcanic ash clouds in the 10–12 μm window using AVHRR/2 data, International Journal of Remote Sensing, 10, 751–761, https://doi.org/10.1080/01431168908903916, 1989a.

575 Prata, A. J.: Infrared radiative transfer calculations for volcanic ash clouds, Geophysical Research Letters, 16, 1293–1296, https://doi.org/10.1029/GL016i011p01293, 1989b.

Prata, A. J. and Grant, I. F.: Retrieval of microphysical and morphological properties of volcanic ash plumes from satellite data: Application to Mt Ruapehu, New Zealand, Quarterly Journal of the Royal Meteorological Society, 127, 2153–2179, https://doi.org/10.1002/qj.49712757615, http://doi.wiley.com/10.1002/qj.49712757615, 2001.

580 Prata, A. J. and Kerkmann, J.: Simultaneous retrieval of volcanic ash and SO_2 using MSG-SEVIRI measurements, Geophysical Research Letters, 34, https://doi.org/10.1029/2006GL028691, http://doi.wiley.com/10.1029/2006GL028691, 2007.

Prata, A. J. and Prata, A. T.: Eyjafjallajökull volcanic ash concentrations determined using Spin Enhanced Visible and Infrared Imager measurements, Journal of Geophysical Research: Atmospheres, 117, https://doi.org/10.1029/2011JD016800, http://doi.wiley.com/10.1029/2011JD016800, 2012.

585 Prata, A. T., Young, S. A., Siems, S. T., and Manton, M. J.: Lidar ratios of stratospheric volcanic ash and sulfate aerosols retrieved from CALIOP measurements, Atmospheric Chemistry and Physics, 17, 8599–8618, https://doi.org/10.5194/acp-17-8599-2017, https://www.atmos-chem-phys.net/17/8599/2017/, 2017.

Prata, F. and Lynch, M.: Passive Earth Observations of Volcanic Clouds in the Atmosphere, Atmosphere, 10, 199, https://doi.org/10.3390/atmos10040199, https://www.mdpi.com/2073-4433/10/4/199, 2019.

590 Prata, F., Bluth, G., Rose, B., Schneider, D., and Tupper, A.: Comments on "Failures in detecting volcanic ash from a satellite-based technique", Remote Sensing of Environment, 78, 341–346, https://doi.org/10.1016/S0034-4257(01)00231-0, https://linkinghub.elsevier.com/retrieve/pii/S0034425701002310, 2001.

Realmuto, V. J., Abrams, M. J., Buongiorno, M. F., and Pieri, D. C.: The use of multispectral thermal infrared image data to estimate the sulfur dioxide flux from volcanoes: A case study from Mount Etna, Sicily, July 29, 1986, Journal of Geophysical Research: Solid Earth,
595 99, 481–488, https://doi.org/10.1029/93JB02062, http://doi.wiley.com/10.1029/93JB02062, 1994.

Rose, W. I., Bluth, G. J. S., Schneider, D. J., Ernst, G. G. J., Riley, C. M., Henderson, L. J., and McGimsey, R. G.: Observations of Volcanic Clouds in Their First Few Days of Atmospheric Residence: The 1992 Eruptions of Crater Peak, Mount Spurr Volcano, Alaska, The Journal of Geology, 109, 677–694, https://doi.org/10.1086/323189, https://www.journals.uchicago.edu/doi/10.1086/323189, 2001.

Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S., and Ratier, A.: An Introduction to Meteosat Second Generation (MSG),
600 Bulletin of the American Meteorological Society, 83, 977–992, https://doi.org/10.1175/1520-0477(2002)083<0977:AITMSG>2.3.CO;2, http://journals.ametsoc.org/doi/abs/10.1175/1520-0477%282002%29083%3C0977%3AAITMSG%3E2.3.CO%3B2, 2002.

Seemann, S. W., Borbas, E. E., Knuteson, R. O., Stephenson, G. R., and Huang, H.-L.: Development of a Global Infrared Land Surface Emissivity Database for Application to Clear Sky Sounding Retrievals from Multispectral Satellite Radiance Measurements, Journal of Applied Meteorology and Climatology, 47, 108–123, https://doi.org/10.1175/2007JAMC1590.1, http://journals.ametsoc.org/doi/10.1175/
605 2007JAMC1590.1, 2008.

Simpson, J.: Failures in Detecting Volcanic Ash from a Satellite-Based Technique, Remote Sensing of Environment, 72, 191–217, https://doi.org/10.1016/S0034-4257(99)00103-0, https://linkinghub.elsevier.com/retrieve/pii/S0034425799001030, 2000.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y., Wang, W., and Powers, J. G.: A description of the Advanced Research WRF Version 3, Tech. rep., National Center for Atmospheric Research, Boulder, Colorado, USA,
610 NCAR Technical Note, NCAR/TN-475+STR, 2008.

Stamnes, K., Tsay, S.-C., Wiscombe, W., and Jayaweera, K.: Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media, Applied Optics, 27, 2502, https://doi.org/10.1364/AO.27.002502, https://www.osapublishing.org/abstract.cfm?URI=ao-27-12-2502, 1988.

Stevenson, J. A., Millington, S. C., Beckett, F. M., Swindles, G. T., and Thordarson, T.: Big grains go far: understanding the discrepancy
615 between tephrochronology and satellite infrared measurements of volcanic ash, Atmospheric Measurement Techniques, 8, 2069–2091, https://doi.org/10.5194/amt-8-2069-2015, https://www.atmos-meas-tech.net/8/2069/2015/, 2015.

Vernier, J.-P., Fairlie, T. D., Murray, J. J., Tupper, A., Trepte, C., Winker, D., Pelon, J., Garnier, A., Jumelet, J., Pavolonis, M., Omar, A. H., and Powell, K. A.: An Advanced System to Monitor the 3D Structure of Diffuse Volcanic Ash Clouds, Journal of Applied Meteorology and Climatology, 52, 2125–2138, https://doi.org/10.1175/JAMC-D-12-0279.1, http://journals.ametsoc.org/doi/abs/10.1175/JAMC-D-12-0279.1,
620 2013.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, I., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., Vijaykumar, A., Bardelli, A. P., Rothberg, A., Hilboll, A., Kloeckner,
625 A., Scopatz, A., Lee, A., Rokem, A., Woods, C. N., Fulton, C., Masson, C., Häggström, C., Fitzgerald, C., Nicholson, D. A., Hagen,

D. R., Pasechnik, D. V., Olivetti, E., Martin, E., Wieser, E., Silva, F., Lenders, F., Wilhelm, F., Young, G., Price, G. A., Ingold, G.-L., Allen, G. E., Lee, G. R., Audren, H., Probst, I., Dietrich, J. P., Silterra, J., Webber, J. T., Slaviĉ, J., Nothman, J., Buchner, J., Kulick, J., Schönberger, J. L., de Miranda Cardoso, J. V., Reimer, J., Harrington, J., Rodríguez, J. L. C., Nunez-Iglesias, J., Kuczynski, J., Tritz, K., Thoma, M., Newville, M., Kümmerer, M., Bolingbroke, M., Tartre, M., Pak, M., Smith, N. J., Nowaczyk, N., Shebanov, N., Pavlyk, O.,

630   Brodtkorb, P. A., Lee, P., McGibbon, R. T., Feldbauer, R., Lewis, S., Tygier, S., Sievert, S., Vigna, S., Peterson, S., More, S., Pudlik, T., Oshima, T., Pingel, T. J., Robitaille, T. P., Spura, T., Jones, T. R., Cera, T., Leslie, T., Zito, T., Krauss, T., Upadhyay, U., Halchenko, Y. O., Vázquez-Baeza, Y., and SciPy 1.0 Contributors: SciPy 1.0: fundamental algorithms for scientific computing in Python, Nature Methods, 17, 261–272, https://doi.org/10.1038/s41592-019-0686-2, https://doi.org/10.1038/s41592-019-0686-2, 2020.

Watson, I., Realmuto, V., Rose, W., Prata, A., Bluth, G., Gu, Y., Bader, C., and Yu, T.: Thermal infrared remote sensing of volcanic

635   emissions using the moderate resolution imaging spectroradiometer, Journal of Volcanology and Geothermal Research, 135, 75–89, https://doi.org/10.1016/j.jvolgeores.2003.12.017, https://linkinghub.elsevier.com/retrieve/pii/S0377027304000307, 2004.

Wen, S. and Rose, W. I.: Retrieval of sizes and total masses of particles in volcanic clouds using AVHRR bands 4 and 5, Journal of Geophysical Research, 99, 5421, https://doi.org/10.1029/93JD03340, http://doi.wiley.com/10.1029/93JD03340, 1994.

Wernli, H., Paulat, M., Hagen, M., and Frei, C.: SAL—A Novel Quality Measure for the Verification of Quantitative Precipitation Fore-

640   casts, Monthly Weather Review, 136, 4470–4487, https://doi.org/10.1175/2008MWR2415.1, http://journals.ametsoc.org/doi/abs/10.1175/2008MWR2415.1, 2008.

Western, L. M., Watson, M. I., and Francis, P. N.: Uncertainty in two-channel infrared remote sensing retrievals of a well-characterised volcanic ash cloud, Bulletin of Volcanology, 77, https://doi.org/10.1007/s00445-015-0950-y, http://link.springer.com/10.1007/s00445-015-0950-y, 2015.

645   Wilkins, K. L., Watson, I. M., Kristiansen, N. I., Webster, H. N., Thomson, D. J., Dacre, H. F., and Prata, A. J.: Using data insertion with the NAME model to simulate the 8 May 2010 Eyjafjallajökull volcanic ash cloud, Journal of Geophysical Research: Atmospheres, 121, 306–323, https://doi.org/10.1002/2015JD023895, http://doi.wiley.com/10.1002/2015JD023895, 2016.

Wilson, T. M., Stewart, C., Bickerton, H., Baxter, P., Outes, V., Villarosa, G., and Rovere, E.: Impacts of the June 2011 Puyehue–Cordón Caulle volcanic complex eruption on urban infrastructure, agriculture and public health, GNS Science Report 2012/20, p. 88, 2013.

650   Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K. A., Liu, Z., Hunt, W. H., and Young, S. A.: Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms, Journal of Atmospheric and Oceanic Technology, 26, 2310–2323, https://doi.org/10.1175/2009JTECHA1281.1, http://journals.ametsoc.org/doi/abs/10.1175/2009JTECHA1281.1, 2009.

Yamanouchi, T., Suzuki, K., and Kawaguchi, S.: Detection of Clouds in Antarctica from Infrared Multispectral Data of AVHRR, Journal of the Meteorological Society of Japan. Ser. II, 65, 949–962, https://doi.org/10.2151/jmsj1965.65.6_949, 1987.

655   Yu, T., Rose, W. I., and Prata, A. J.: Atmospheric correction for satellite-based volcanic ash mapping and retrievals using "split window" IR data from GOES and AVHRR, Journal of Geophysical Research, 107, https://doi.org/10.1029/2001JD000706, http://doi.wiley.com/10.1029/2001JD000706, 2002.
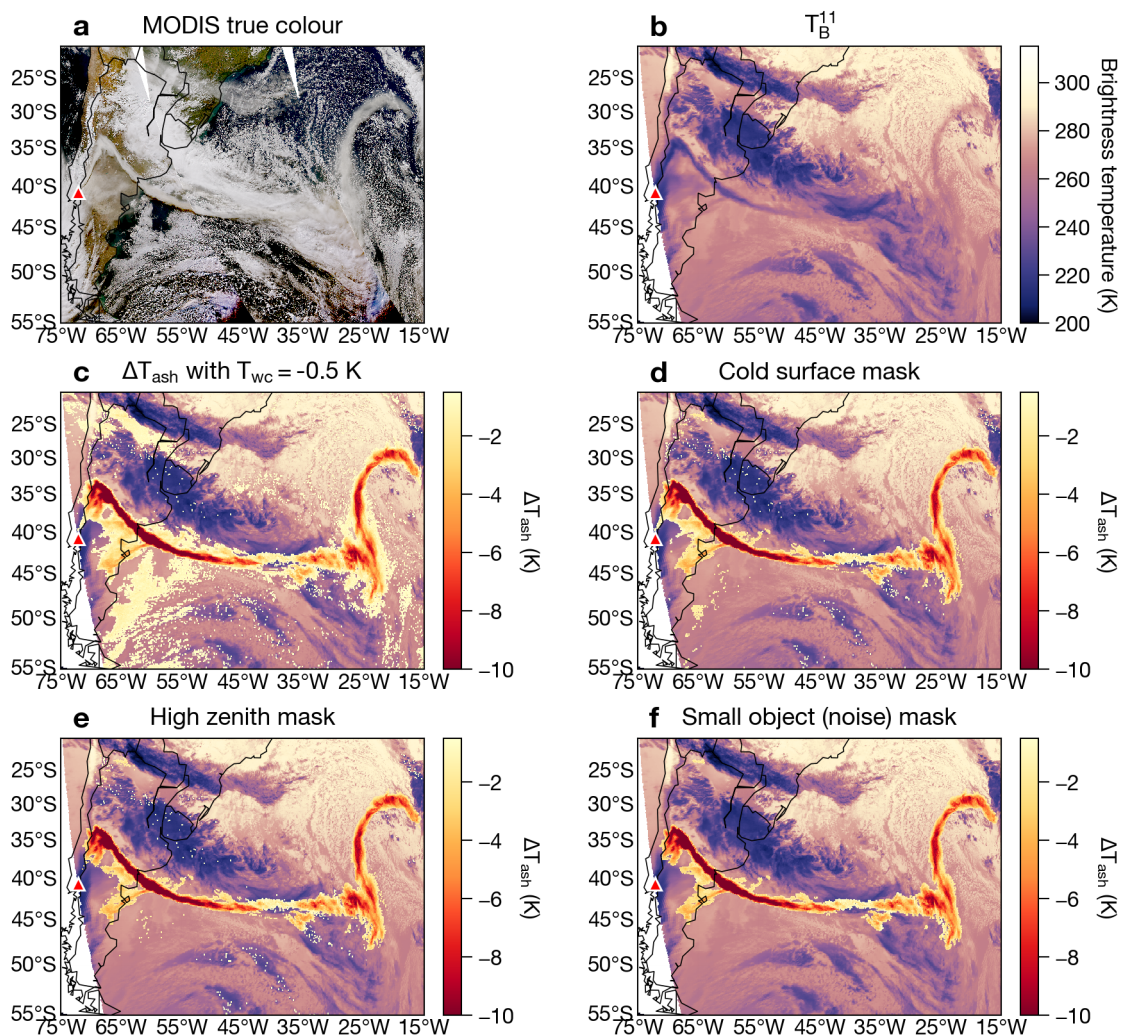
Figure 1. Volcanic ash detection scheme for the Puyehue-Cordón Caulle (indicated by the triangle on each map) eruption. (a) MODIS true color composite from 2011-06-06 at 15:15–18:40 UTC. (b) SEVIRI 10.8 μm brightness temperature (T_B^{11}) at 2011-06-06 18:45 UTC. (c) Same as (b) with water-vapor corrected BTD ($\Delta T_{ash} = T_B^{11} - T_B^{12}$) overlaid. (d), (e) and (f) are the same as (c) with cold surface, high zenith and noise masks applied, respectively.
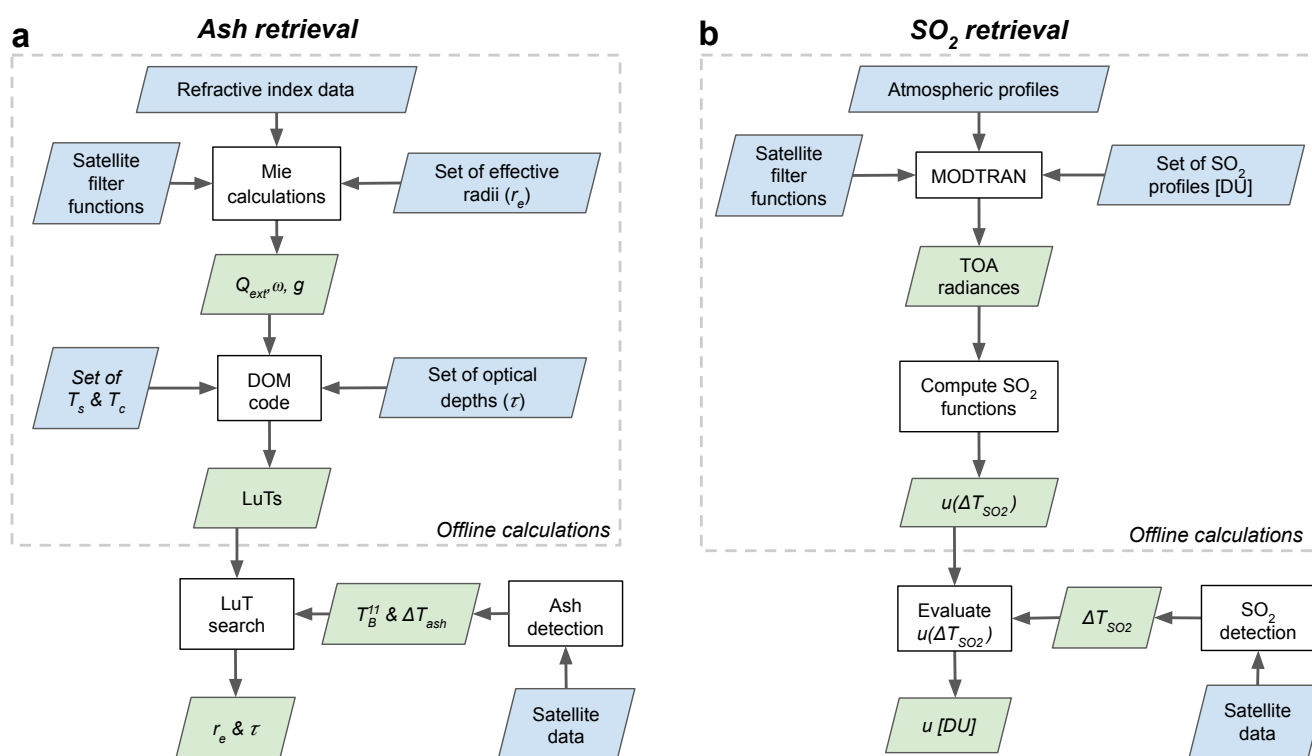
Figure 2. Flow diagram showing the (a) volcanic ash retrieval process and (b) volcanic SO$_2$ retrieval process used in the present study. Parallelograms indicate datasets (blue for inputs, green for outputs) and rectangles indicate processes (i.e. code used to implement the retrieval algorithms and perform radiative transfer calculations). Offline calculations are any computations that are pre-computed (i.e. before any observations are made by the satellite).
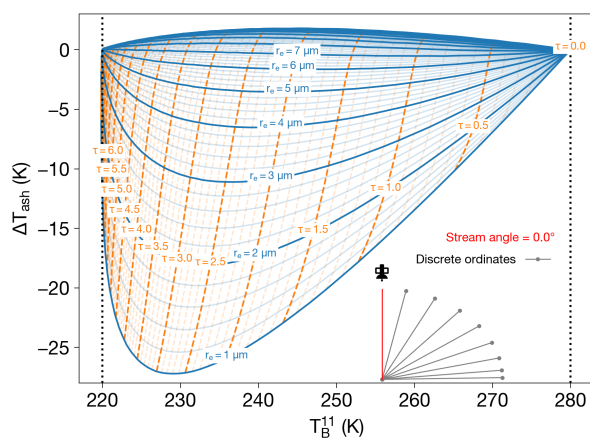
Figure 3. Graphical illustration of a volcanic ash look-up table for a surface temperature $T_s = 280$ K and cloud-top temperature, $T_c = 220$ K. Dashed near-vertical lines indicate lines of constant optical depth, τ, and solid U-shaped curves indicate lines of constant effective radius, r_e.
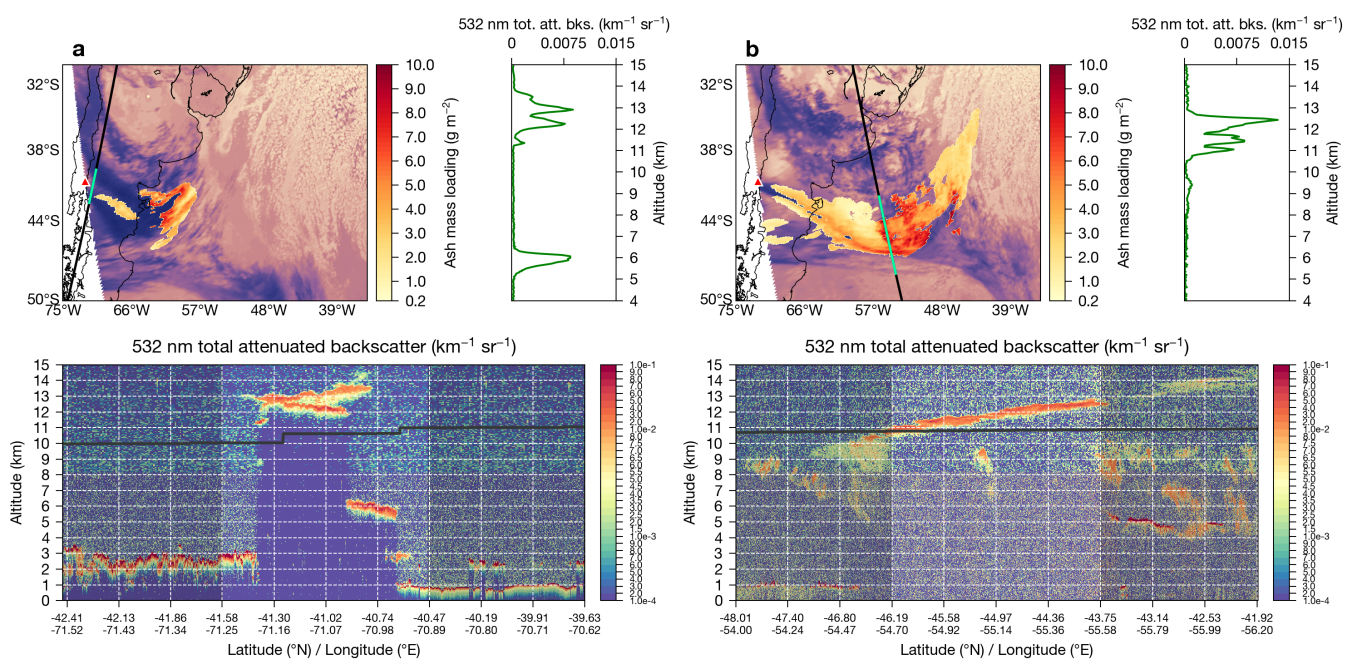
Figure 4. (a) SEVIRI ash mass loadings and CALIOP vertical profile of the Cordón Caulle ash plume on 5 June 2011 at 06:00 UTC. Top left panel: Mass loading retrievals (yellow-orange-red color scale) with brightness temperatures plotted underneath (red triangle indicates location of Cordón Caulle). Black line indicates CALIOP track and green highlight indicates full latitude/longitude range displayed on the bottom panel. Top right panel: 532 nm total attenuated backscatter profile averaged over the latitude/longitude range highlighted on bottom panel. Bottom panel: 532 nm total attenuated backscatter curtain (black line indicates tropopause). (b) Same as (a), but for 5 June 2011 at 18:00 UTC.
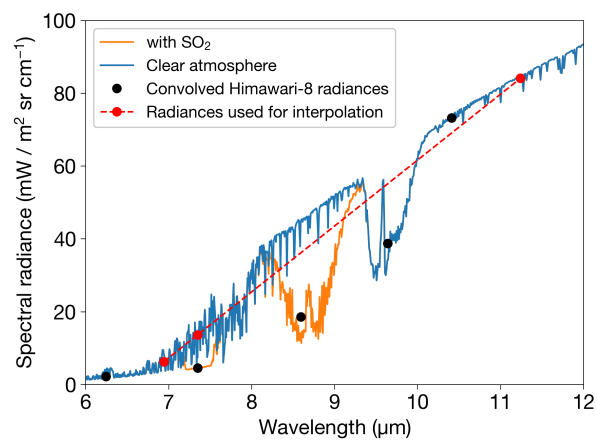
Figure 5. MODTRAN6.0 simulations for atmospheres with and without SO_2 demonstrationing how the interpolation procedure is used to estimate a clear-sky radiances from an atmosphere with SO_2. Convolved radiances were derived using the Himawari-8 spectral response functions.
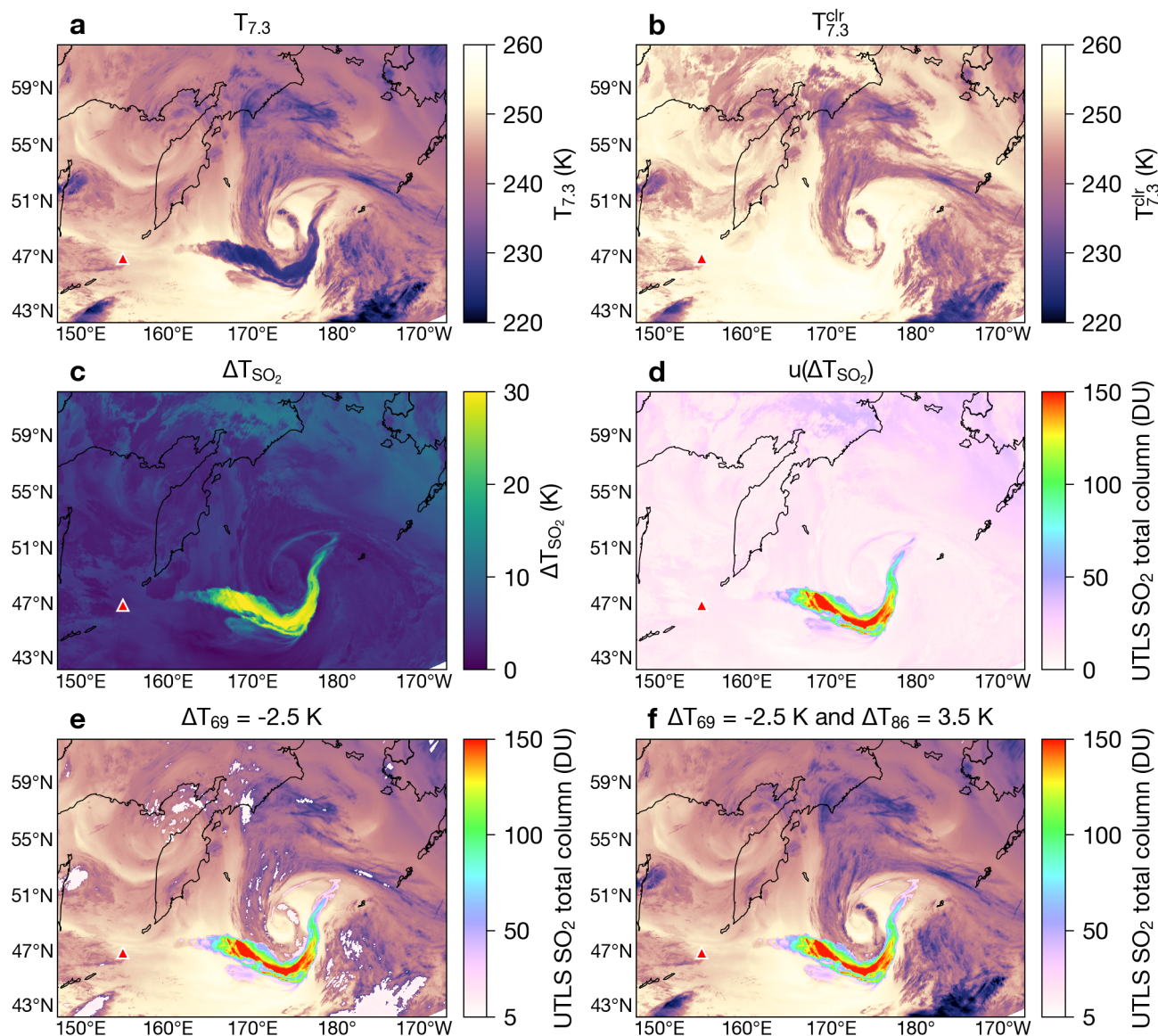
Figure 6. Volcanic SO$_2$ detection scheme applied to Himawari-8 observations of the Raikoke SO$_2$ cloud on 22 June 2019 at 21:00 UTC. Location of Raikoke volcano indicated by red triangle on each map. (a) Himawari-8 brightness temperature at 7.3 μm ($T_B^{7.3}$). (b) Synthetic 7.3 μm brightness temperature ($T_{BC}^{7.3}$) determined from interpolation procedure (see Sect. 2.2.1). (c) SO$_2$ BTD ($\Delta T_{SO_2} = T_B^{7.3} - T_{BC}^{7.3}$). (d) SO$_2$ total column loading, $u(\Delta SO_2)$, (see Sect. 2.2.2 for details). (e) $u(\Delta SO_2)$ with $\Delta T_{wv} = -2.5$ K and 5 DU thresholds. (f) $u(\Delta SO_2)$ with $\Delta T_{wv} = -2.5$ K, $\Delta T_{86} = -3.5$ K and 5 DU thresholds.
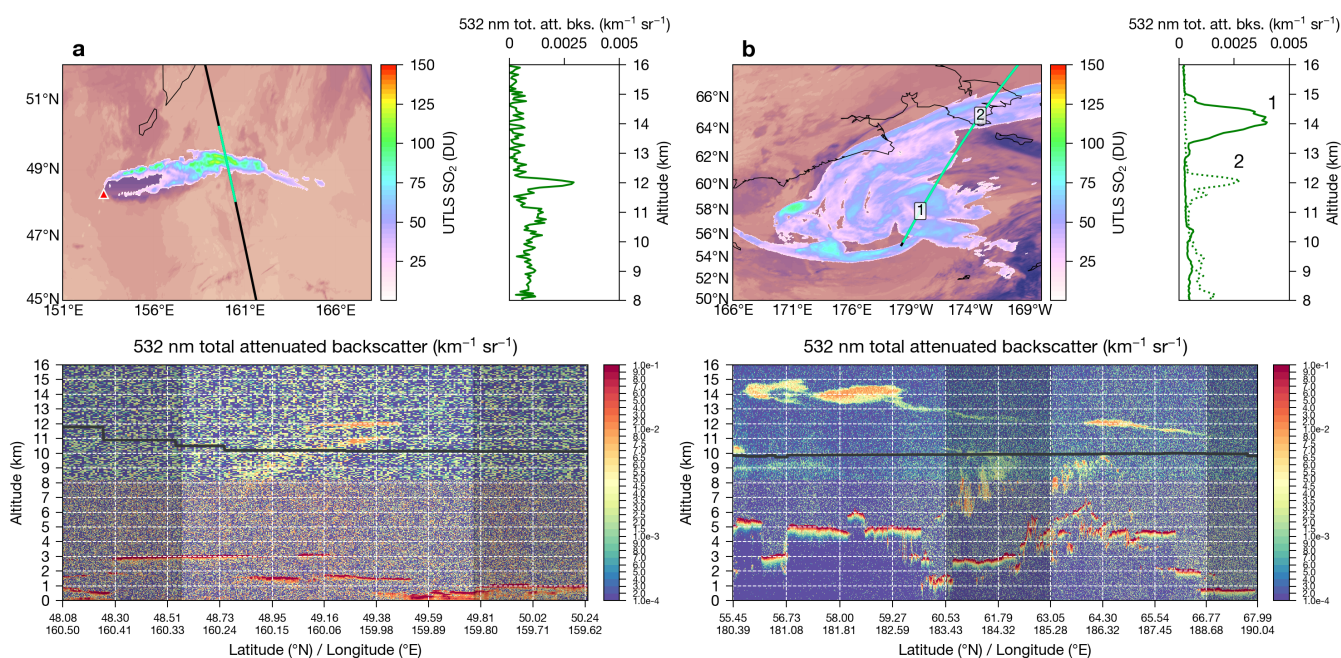
Figure 7. (a) AHI SO$_2$ upper-troposphere lower-stratosphere (UTLS) total column burdens (DU) and CALIOP vertical profile of the Raikoke SO$_2$ plume on 22 June 2019 at 02:00 UTC. Top left panel: 7.3 μm SO$_2$ total column retrievals (white-purple-green-red color scale) with 11 μm brightness temperatures plotted underneath (red triangle indicates location of Raikoke). Black line indicates CALIOP track and green highlight indicates full latitude/longitude range displayed on the bottom panel. Top right panel: 532 nm total attenuated backscatter profile averaged over the latitude/longitude range highlighted on the bottom panel. Bottom panel: 532 nm total attenuated backscatter curtain (black line indicates tropopause). (b) Same (a), but for 25 June 2019 at 14:00 UTC.
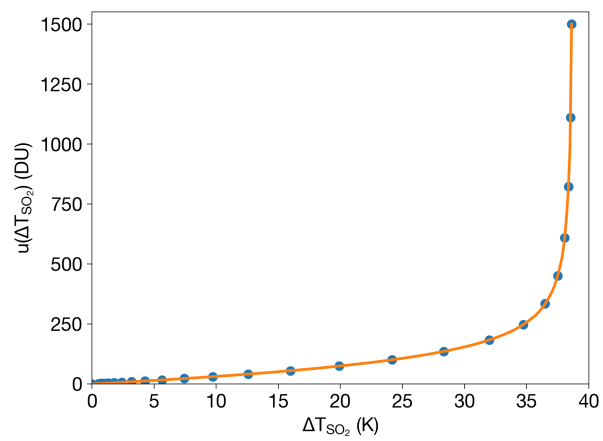
Figure 8. MODTRAN-6.0 simulation results used to derive SO_2 total column density as a function of brightness temperature difference (see Eq. 7) for a uniform distribution with cloud-top height of 13.5 km and thickness of 2.5 km.
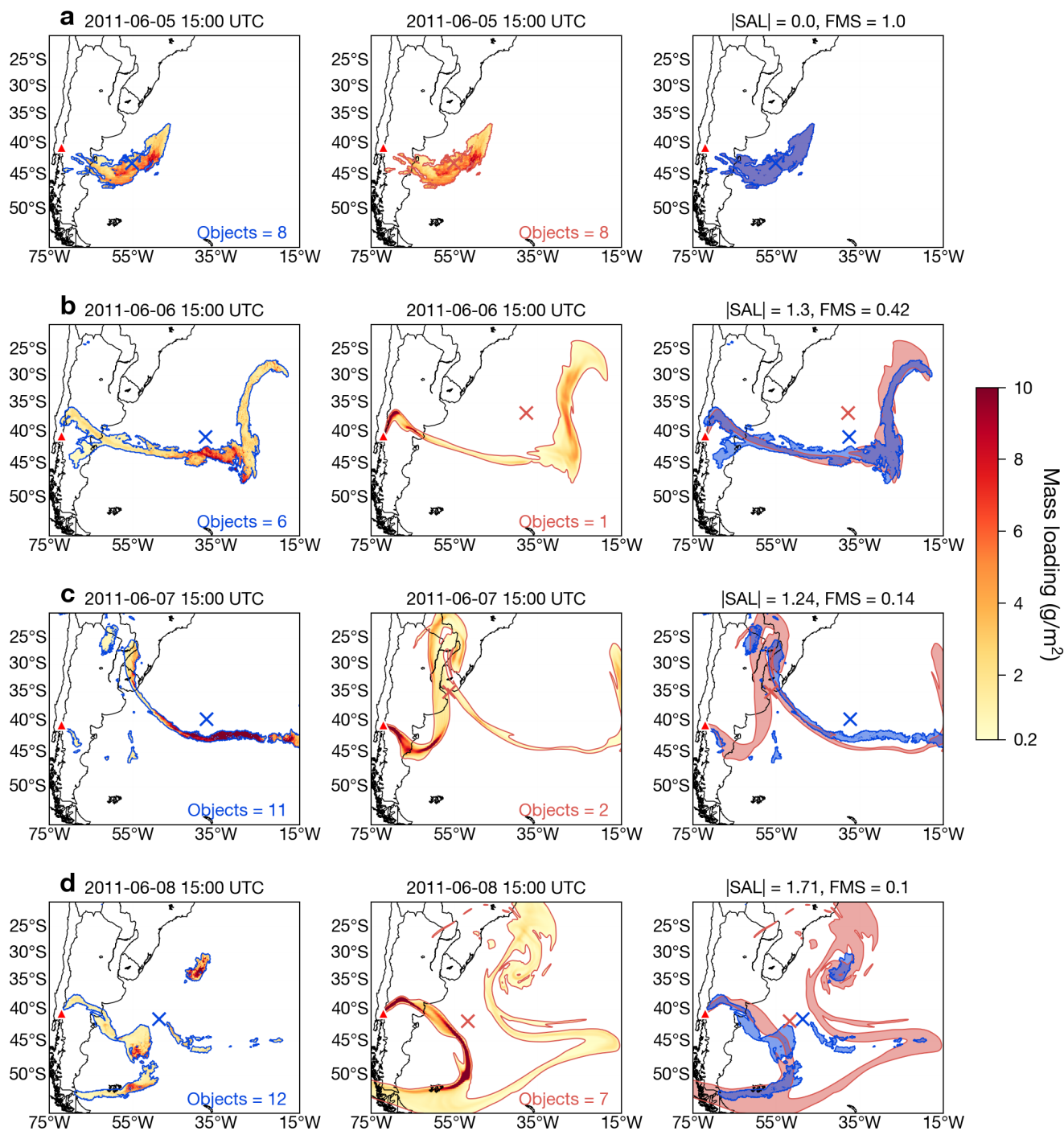
Figure 9. *FALL3D-8.0* validation of fine ash mass loading using SEVIRI mass loading retrievals on (a) 5 June 2011 at 15:00 UTC (data insertion time), (b) 6 June 2011 at 15:00 UTC, (c) 7 June 2011 at 15:00 UTC and (d) 8 June 2011 at 15:00 UTC. Left panels show satellite retrievals with 0.2 g m^{-2} contour in blue and centre of mass indicated with 'x'. Middle panels show *FALL3D-8.0* fine ash mass loading model simulation (0.2 g m^{-2} contour in red). Right panels show spatial overlap of model vs. observed fields with validation metrics annotated (see Sect. 3 for details). A full animation of the data insertion simulations is available in the Supplementary Material.
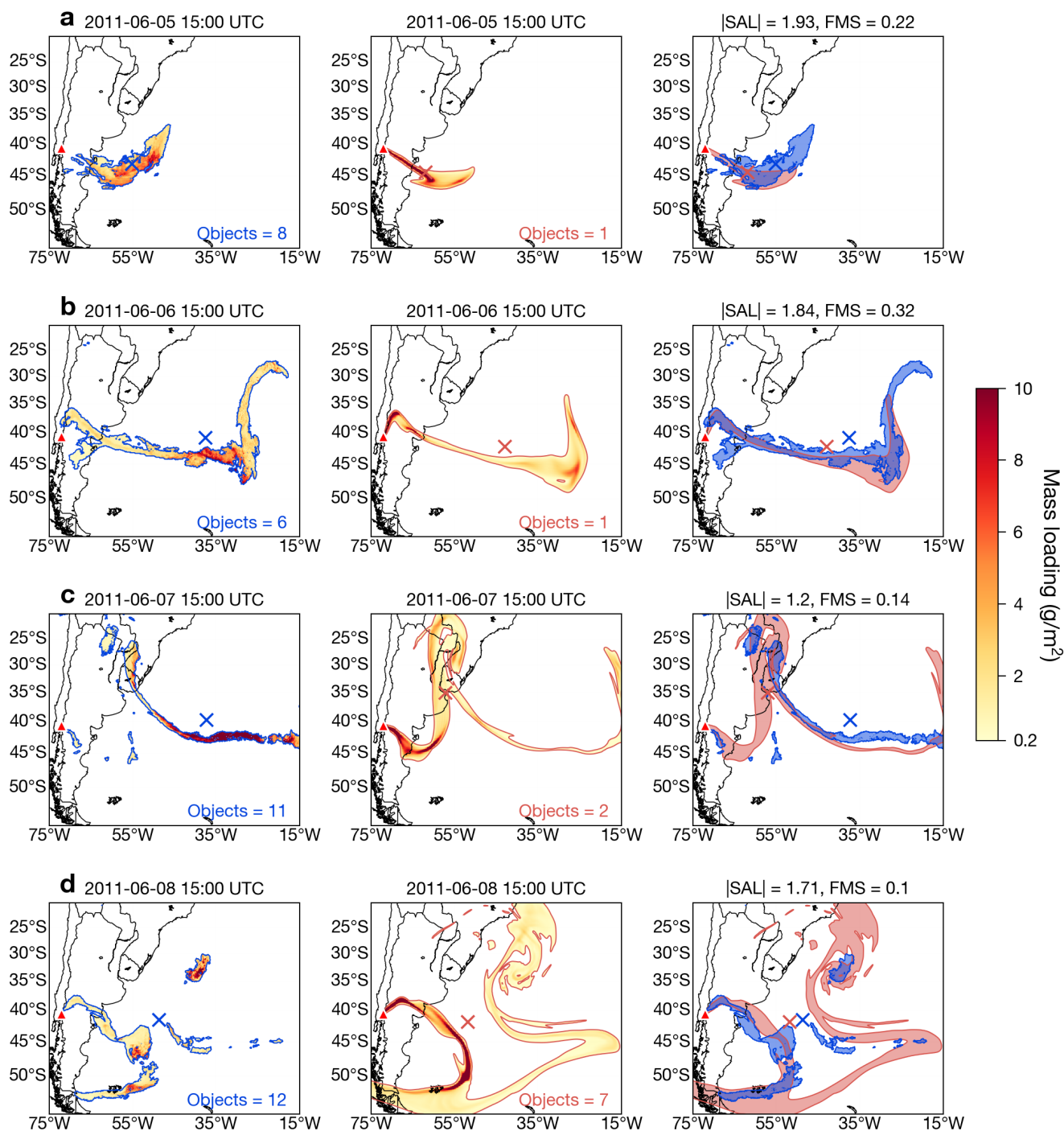
30

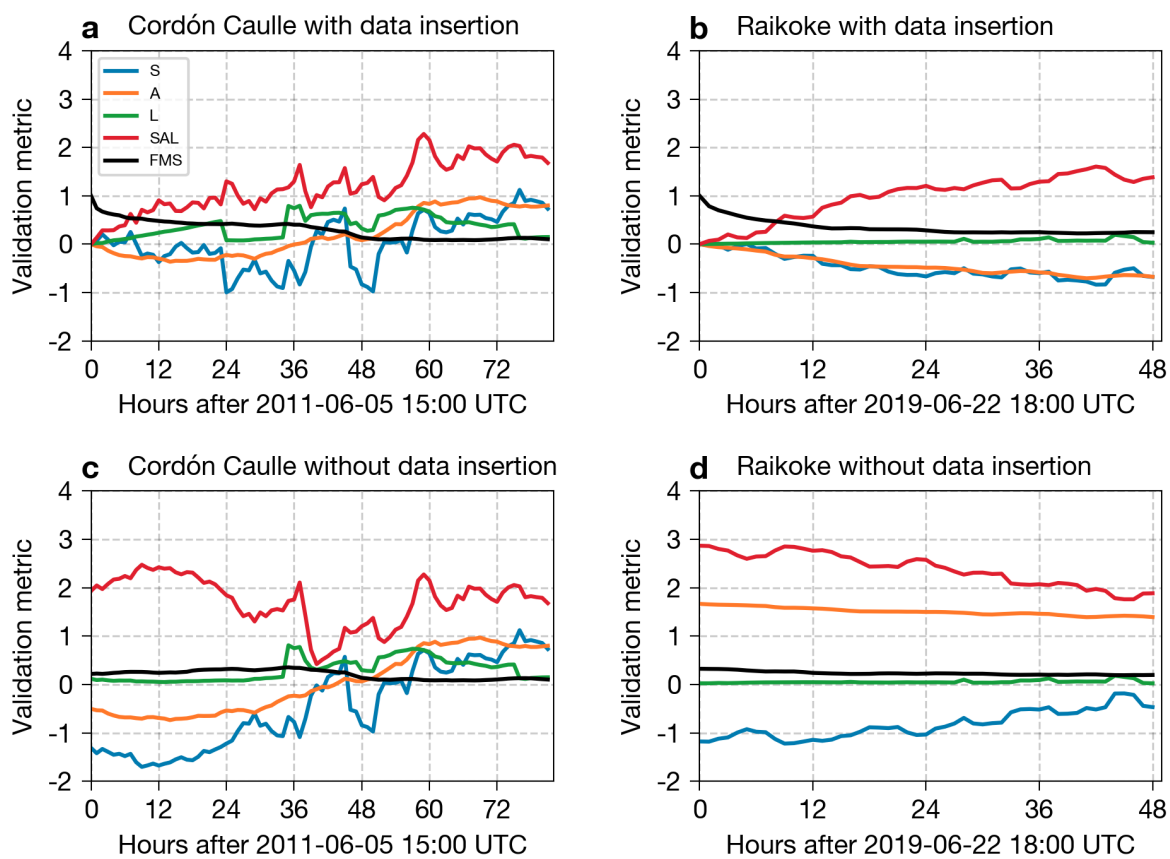**Figure 10.** Same as Fig. 9 but without data insertion.

Figure 11. Time-series of validation metrics for (a) Cordón Caulle with data insertion, (b) Raikoke with data insertion, (c) Cordón Caulle without data insertion and (d) Raikoke without data insertion.
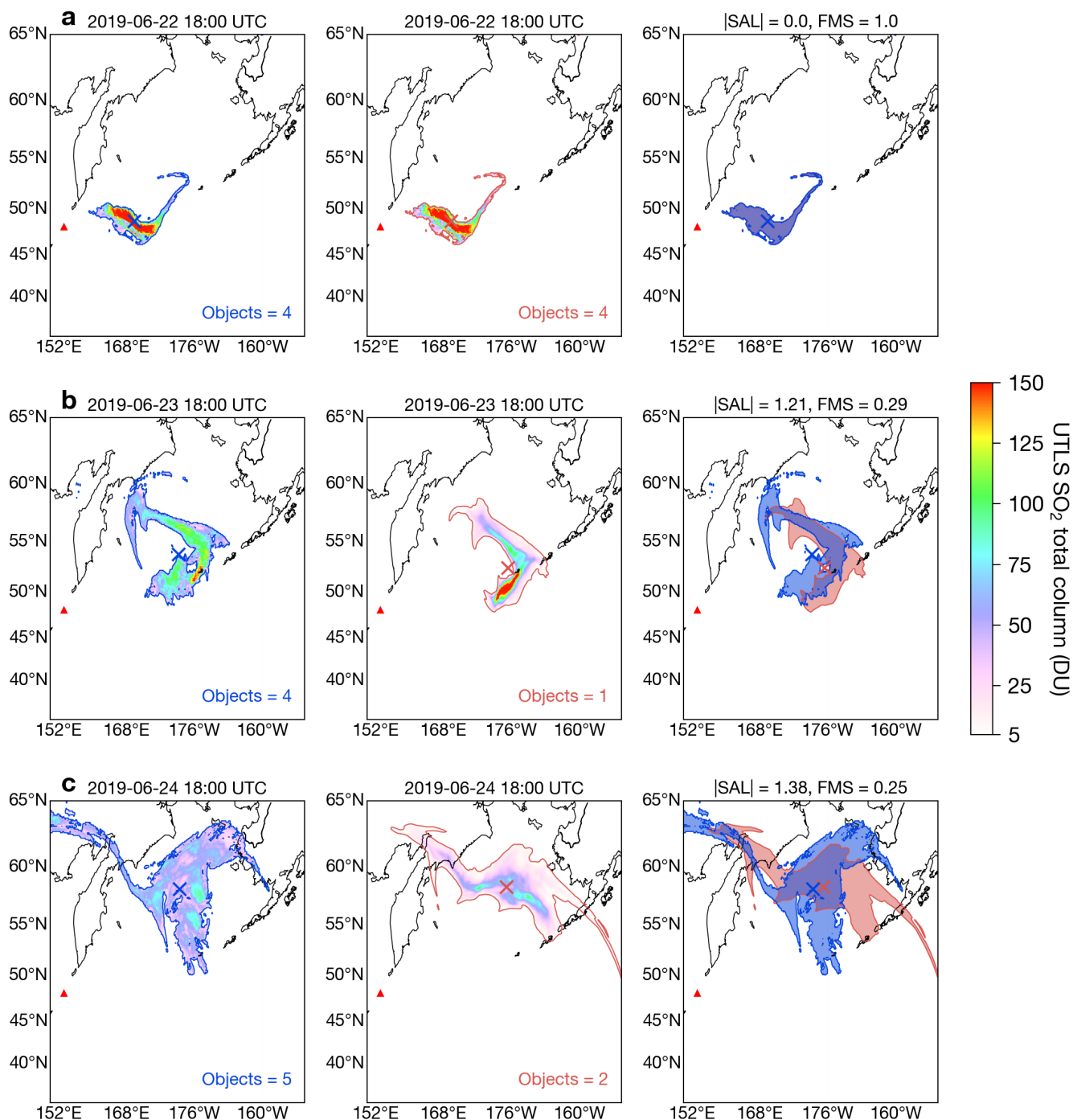
Figure 12. *FALL3D-8.0* validation of SO$_2$ simulations using AHI upper-troposphere lower-stratosphere (UTLS) total column burdens retrievals (DU) on (a) 22 June 2019 at 18:00 UTC (data insertion time), (b) 23 June 2019 at 18:00 UTC and (c) 24 June 2019 at 18:00 UTC. Left column shows satellite retrievals with 5 DU contour in blue and centre of mass indicated with 'x'. Middle column shows *FALL3D-8.0* model simulation (5 DU contour in red). Right column shows spatial overlap of model vs. observed fields. A full animation of the data insertion simulations is available in the Supplementary Material.
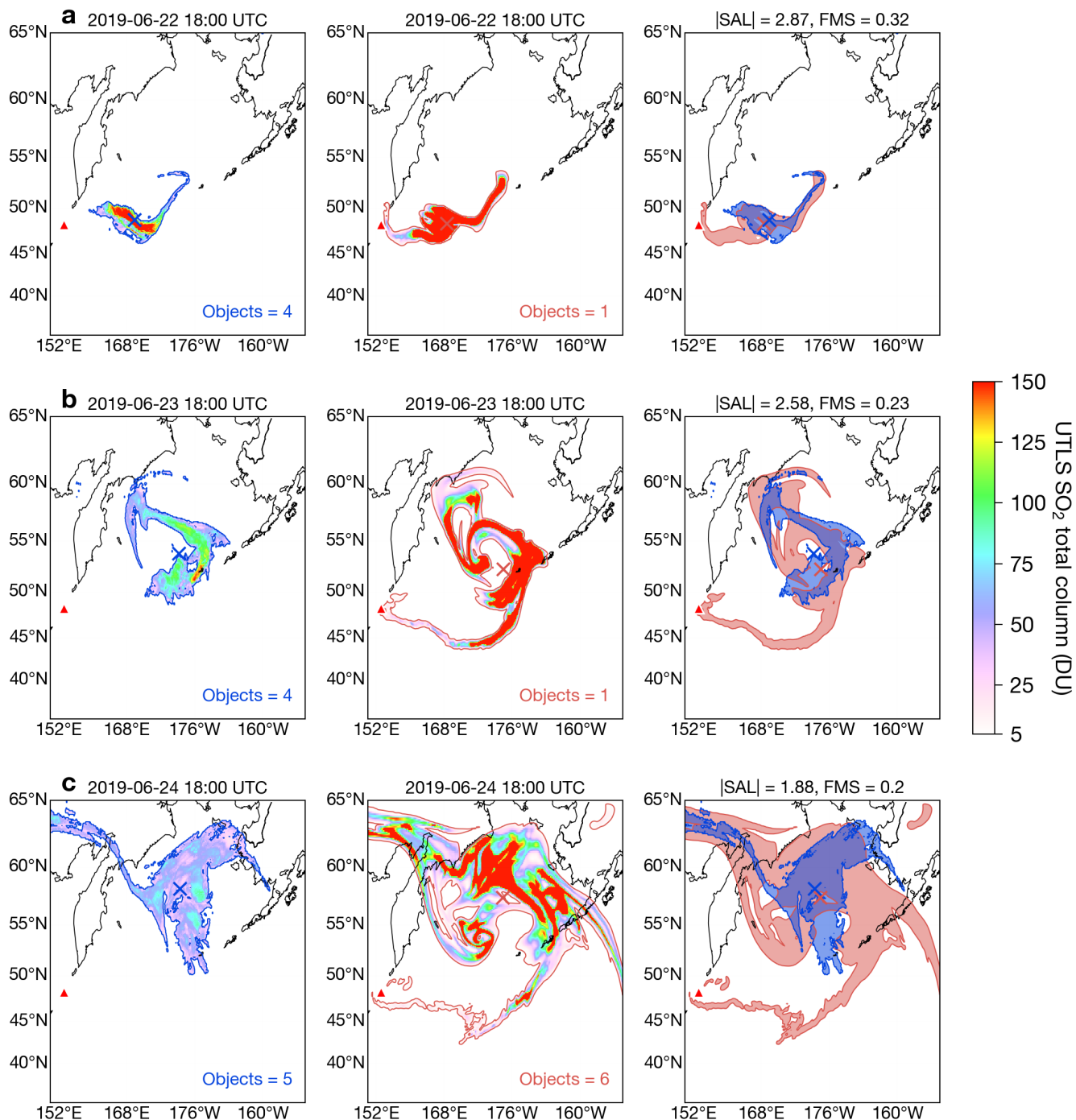
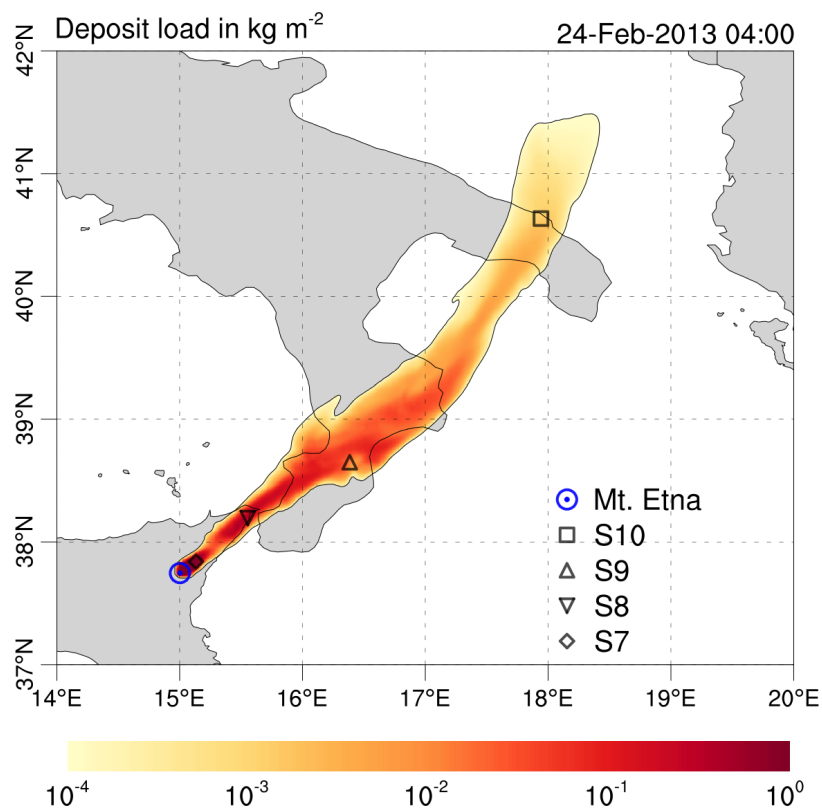**Figure 13.** Same as Fig. 12 but without data insertion.

Figure 14. Tephra fallout deposit simulated for the Mt. Etna eruptive episode on 23 February 2013. The spatial distribution of the modelled tephra loading coincides with the locations of sampling sites. Sites S7-S10 are indicated in the map by symbols. Proximal sites S1-S6 ($< 16\,km$ from the source) not shown for clarity.
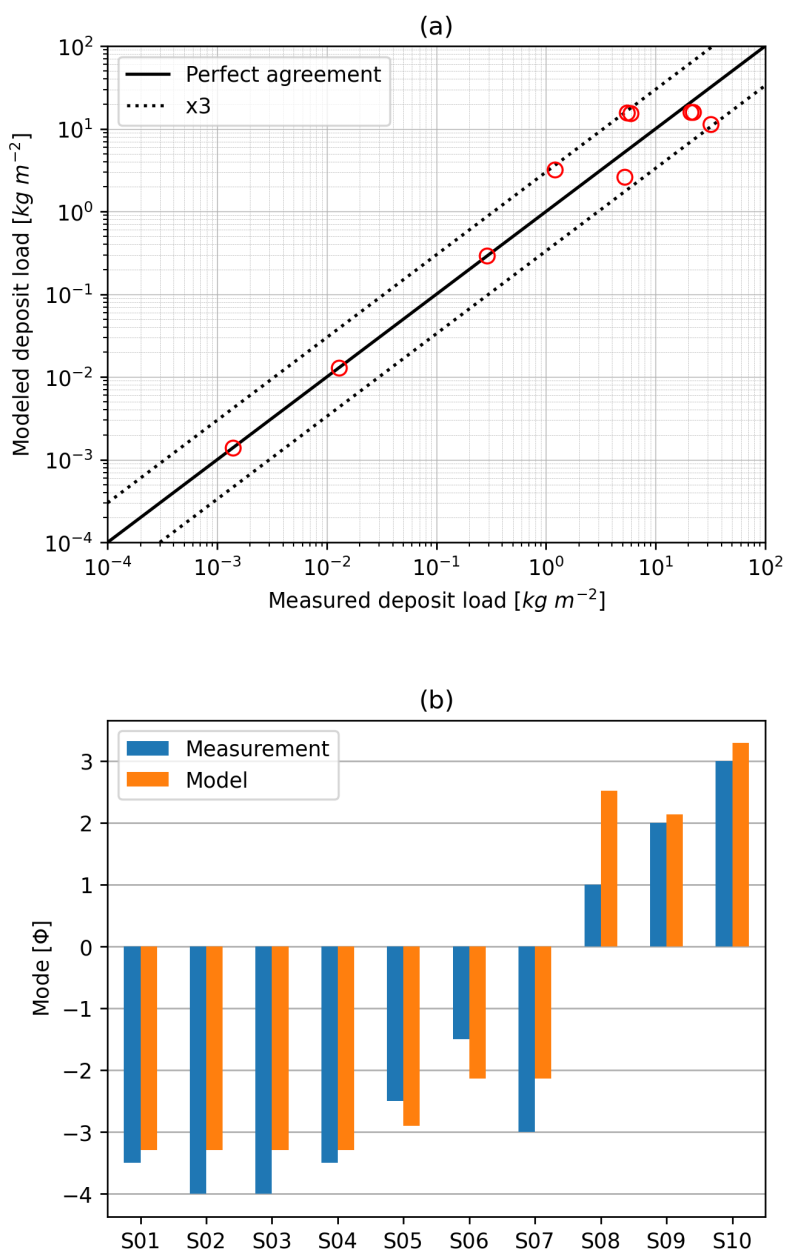
Figure 15. Comparison between field data at 10 sampling sites (S1-S10) and results of numerical simulations for the Mt. Etna eruptive episode on 23 February 2013. (a) Tephra mass loading. (b) Mode of the distributions. Field data were obtained from the sample dataset reported by Poret et al. (2018).

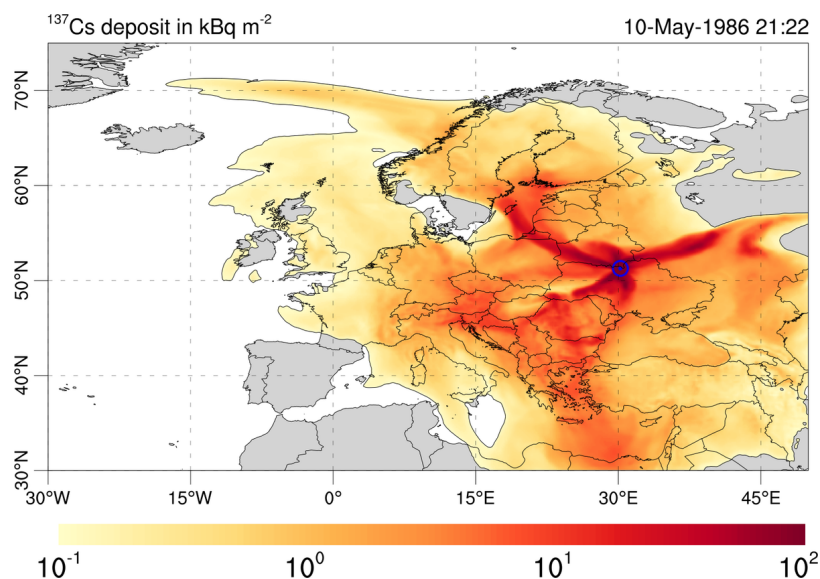**Figure 16.** Simulated accumulated total deposition of ^{137}Cs related to the Chernobyl accident, evaluated on 10 May 1986. The underlying map is reported just for reference and could contain nations borders that are under dispute.

Figure 17. Comparison between measurements and simulated deposit at different locations of ^{137}Cs, ^{134}Cs and ^{131}I, for the Chernobyl accident, evaluated on 10 May 1986.

Figure 18. Simulated vertically integrated radioactivity concentration in the atmosphere of ^{137}Cs related to the Chernobyl accident, at different times. (a) 28 April 1986, (b) 2 May 1986, (c) 5 May 1986 and (d) 9 May 1986. The underlying maps are reported just for reference and could contain nations borders that are under dispute.

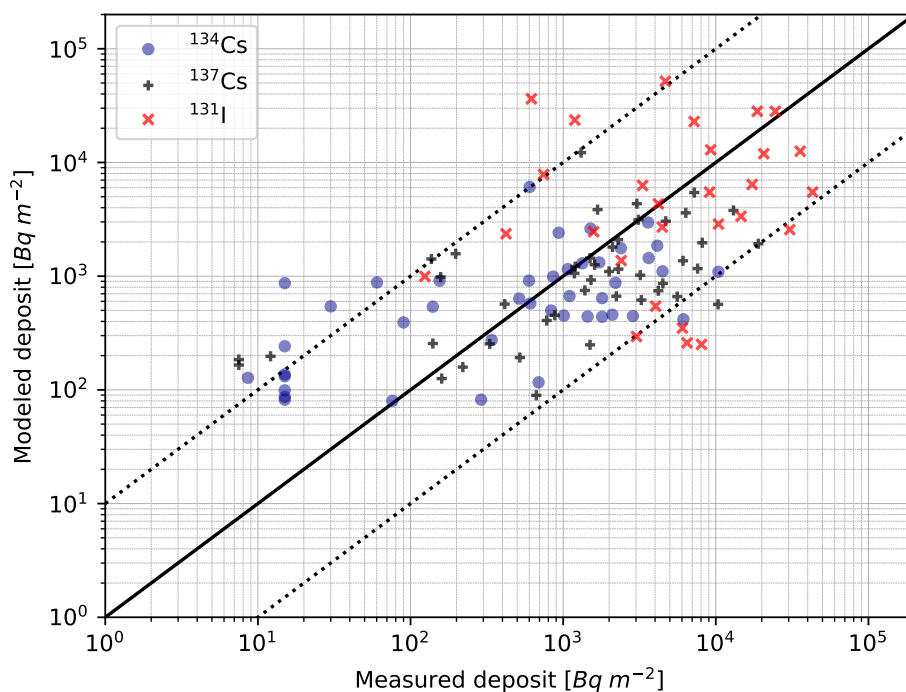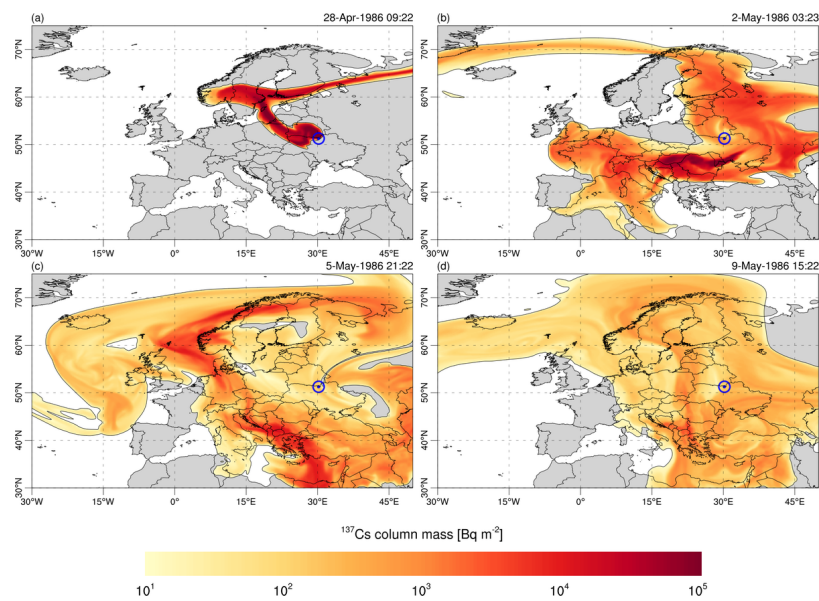**Table 1.** Summary of model setup for the validation cases.

Parameter	**Cordón Caulle** A	**Cordón Caulle** B	**Raikoke** A	**Raikoke** B	**Mt. Etna**	**Chernobyl**
Start date	2011-06-04	2011-06-05	2019-06-21	2019-06-22	2013-02-23	1986-04-25
Start time	21:00 UTC	15:00 UTC	18:00 UTC	18:00 UTC	18:00 UTC	00:00 UTC
Run period	99 h	81 h	72 h	48 h	10 h	384 h
Resolution (hor.)	$0.1°$	$0.1°$	$0.1°$	$0.1°$	$0.015°$	$0.125°$
Vertical levels	60	60	80	80	60	60
Species	Fine ash	Fine ash	SO_2	SO_2	Tephra	Radionuclides
Data insertion	No	Yes	No	Yes	No	No
Source type	Top-Hat	Top-Hat	Suzuki	No source	Top-hat	hybrid
Initial col. height	11.2 km	13 km	15.5 km (max)[†]	13.5 km	8.7 km	3.3. km
Initial col. thickness	2 km	2 km	-	2.5 km	3.5 km	-
Meteo. driver	ERA5	ERA5	GFS	GFS	WRF-ARW	ERA5
			Validation			
Validation data	SEVIRI (Meteosat-9) collocated with CALIPSO		AHI (Himawari-8) collocated with CALIPSO		10 ground points	56 ground points (REM database)
Validation metrics	SAL		SAL		point-to-point error	point-to-point error

†Variable column height between 3.5 and 15.5 km

Table 2. Summary of the SAL and FMS validation scores for the Cordón Caulle and Raikoke case studies considered in this manuscript. The 'DI' columns indicate validation scores for runs with data insertion and 'NDI' indicates scores for runs with no data insertion.

Validation metrics	S		A		L		|SAL|		FMS	
	DI	NDI	DI	NDI	DI	NDI	DI	NDI	DI	NDI
2011 Cordón Caulle										
0 h	0.0	-1.31	0.0	-0.5	0.0	0.12	0.0	1.93	1.0	0.22
6 h	0.04	-1.48	-0.24	-0.68	0.11	0.08	0.39	2.24	0.56	0.26
12 h	-0.37	-1.68	-0.3	-0.69	0.24	0.05	0.91	2.42	0.48	0.24
18 h	-0.17	-1.57	-0.33	-0.68	0.36	0.07	0.86	2.31	0.43	0.29
24 h	-1.0	-1.22	-0.22	-0.54	0.08	0.09	1.3	1.84	0.42	0.32
30 h	-0.6	-0.85	-0.24	-0.52	0.1	0.12	0.95	1.49	0.38	0.3
36 h	-0.55	-0.77	-0.01	-0.23	0.74	0.75	1.3	1.76	0.4	0.34
42 h	0.46	0.15	0.11	-0.06	0.63	0.37	1.19	0.58	0.3	0.28
48 h	-0.83	-0.84	0.08	0.06	0.32	0.3	1.24	1.2	0.14	0.14
54 h	0.04	0.03	0.41	0.43	0.71	0.67	1.16	1.13	0.11	0.1
60 h	0.64	0.64	0.83	0.83	0.67	0.67	2.15	2.15	0.09	0.09
66 h	0.41	0.41	0.91	0.91	0.45	0.45	1.76	1.76	0.09	0.09
72 h	0.46	0.46	0.89	0.89	0.36	0.36	1.71	1.71	0.1	0.1
2019 Raikoke										
0 h	0.0	-1.18	0.0	1.67	0.0	0.02	0.0	2.87	1.0	0.32
6 h	-0.09	-0.98	-0.13	1.63	0.02	0.04	0.24	2.65	0.52	0.28
12 h	-0.24	-1.14	-0.29	1.57	0.04	0.05	0.57	2.76	0.37	0.24
18 h	-0.46	-0.89	-0.46	1.51	0.04	0.04	0.96	2.44	0.31	0.22
24 h	-0.67	-1.03	-0.49	1.5	0.05	0.04	1.21	2.58	0.29	0.23
30 h	-0.62	-0.82	-0.59	1.45	0.05	0.04	1.27	2.31	0.24	0.22
36 h	-0.6	-0.52	-0.58	1.46	0.11	0.09	1.29	2.07	0.25	0.2
42 h	-0.84	-0.52	-0.69	1.4	0.07	0.06	1.6	1.97	0.23	0.21
48 h	-0.68	-0.47	-0.67	1.39	0.03	0.02	1.38	1.88	0.25	0.2

Table 3. Total radioactivity emitted in the atmosphere during the Chernobyl accident in the period 24 April - 10 May, 1986, for Caesium and Iodine isotopes, and their best fit fractions in the considered settling velocity classes

Radionuclide	Total activity (PBq)	Vs=2cm/s	Vs=3cm/s	Vs=4cm/s	Vs=6cm/s
^{134}Cs	54	0.54	0.46		
^{137}Cs	85			1.0	
^{131}I	1760				1.0