RESEARCH ARTICLE

Conducting volcanic ash cloud exercises: practising forecast evaluation procedures and the pull-through of scientific advice to the London VAAC

Frances Beckett[1] · Sara Barsotti[2] · Ralph Burton[3] · Fabio Dioguardi[4,5] · Sam Engwell[4] · Matthew Hort[1] · Nina Kristiansen[1] · Sue Loughlin[4] · Anton Muscat[1] · Martin Osborne[1] · Cameron Saint[1] · John Stevenson[4] · Declan Valters[4] · Claire Witham[1]

Abstract

The London Volcanic Ash Advisory Centre (VAAC) provides forecasts on the expected presence of volcanic ash in the atmosphere to mitigate the risk to aviation. It is fundamentally important that operational capability is regularly tested through exercises, to guarantee an effective response to an event. We have developed exercises which practise the pull-through of scientific advice into the London VAAC, the forecast evaluation process, and the decision-making procedures and discussions needed for generating the best possible forecasts under real-time conditions. London VAAC dispersion model forecasts are evaluated against observations. To test this capability in an exercise, we must create observation data for a hypothetical event. We have developed new methodologies for generating and using simulated satellite and lidar retrievals. These simulated observations enable us to practise our ability to interpret, compare, and evaluate model output and observation data under real-time conditions. Forecast evaluation can benefit from an understanding of how different choices of model setup (input parameters), model physics, and driving meteorological data impact the predicted extent and concentration of ash. Through our exercises, we have practised comparing output from model simulations generated using different models, model setups, and meteorological data, supplied by different institutions. Our exercises also practise the communication and interaction between Met Office (UK) scientists supporting the London VAAC and external experts, enabling knowledge exchange and discussions on the interpretation of model output and observations, as we strive to deliver the best response capability for the aviation industry and stakeholders. In this paper, we outline our exercise methodology, including the use of simulated satellite and lidar observations, and the development of the strategy to compare output generated from different modelling systems. We outline the lessons learnt, including the benefits and challenges of conducting exercises which practise our ability to provide scientific advice for an operational response at the London VAAC.

Keywords Exercises · Volcanic ash clouds · Hazard response · Preparedness

Introduction

Volcanic ash clouds represent a significant hazard to aviation due to the serious detrimental effect ash has on aircraft and their jet engines (Casadevall et al 1996; Clarkson et al 2016; Clarkson and Simpson 2017). To protect air traffic, a network of nine Volcanic Ash Advisory Centres (VAACs; a list of acronyms is provided at the end of this paper) issue forecasts to warn of the presence of ash in the atmosphere to the aviation industry. The Anchorage, Buenos Aires, Darwin, London, Montreal, Tokyo, Toulouse, Wellington, and Washington VAACs have areas of responsibility which cover the global airspace. National aviation regulators, airlines,

and airports use this advice to support their decisions on where it is safe for an aircraft to fly (ICAO 2020).

The London VAAC is based at the Met Office in the United Kingdom (UK) and its area of responsibility covers Iceland, the north-eastern part of the North Atlantic, Scandinavia, and the UK. Events within the London VAAC area of responsibility are infrequent; there have been five volcanic eruptions in Iceland which have produced ash clouds over the last 24 years, but when they do occur, they can have a significant impact on air travel. The eruption of Eyjafjallajökull in Iceland during 2010, for example, led to severe disruption to air traffic across European airspace, with tens of thousands of flights cancelled, and millions of passengers stranded (Harris et al. 2012).

To ensure that an effective operational response is maintained, it is important that our ability to generate and issue volcanic ash forecasts is tested (cf. Witham et al. 2020; Witham C, Kristiansen N, Gurioli L, 2023, Improving communication between volcano observatories and volcanic ash advisory centres in Europe – outcomes from a first workshop, personal communication). This requires a cooperative effort between many different international institutions, which may include, but are not limited to, the responsible volcano observatory, associated geological and geophysical institutions responsible for monitoring the site of the eruption, the responsible VAAC, volcano research institutions, and aviation authorities. To ensure it is ready to respond to the next event, the London VAAC tests their procedures through a series of regular exercises. Exercises called VOLCICE are scheduled monthly and practise the communication procedures in place between the London VAAC, the Icelandic air navigation service provider (ISAVIA), and the Icelandic Met Office (IMO), acting as the state volcano observatory. The exercises also simulate the production and issuance of volcanic ash forecast products by the London VAAC. There are two categories of VOLCICE: Category One (CAT I) exercises are shorter and typically allow for only one forecast to be issued, while CAT II exercises last longer and allow for a series of forecasts to be issued. Once a year, pan-European exercises called VOLCEX practise multi-agency response to the presence of a volcanic ash cloud in the European/North Atlantic (EUR/NAT) regions. They are organised by the International Civil Aviation Organisation (ICAO) and involve multiple organizations, including state volcano observatories, London and Toulouse VAACs, EUROCONTROL (the European Organisation for the Safety of Air Navigation), the Civil Aviation Authority (CAA) in the UK, Meteorological Watch Offices and Air Traffic Control teams across Europe, and airlines. Each year, they test a specific process, data feed, or procedure. Products are often generated ahead of time, using an eruption scenario based on historical eruptions. In addition, the UK Government Office for Science organises exercises which practise the operation of the Scientific Advisory Group for Emergencies (SAGE) for a volcanic eruption scenario. SAGE provides independent scientific and technical advice to decision-makers in the Cabinet Office during a national emergency, bringing together scientific experts within government, academia, and industry from a range of fields relevant to the nature of the emergency and the specific issues under consideration (Donovan 2021).

The VOLCICE and VOLCEX exercises are focused on ensuring systems are working as they should to deliver volcanic ash products to aviation customers, and the SAGE exercises test the flow of scientific advice to government and 'high level' decision-making. However, none of these exercises practise the scientific understanding of a volcanic ash cloud event, the interpretation of model output and observations, evaluation processes, or the ability to produce an optimal forecast. We have developed a new type of exercise, which we call SCI-VOLCICE which practises our ability to interpret and evaluate model simulations and observations, the pull-through of international scientific expertise into the London VAAC during a crisis, and the decision-making procedures and discussions needed to generate an optimal forecast.

In this paper, we describe our SCI-VOLCICE exercise methodology, which includes comparison of dispersion simulations generated using different models, initiated with different source parameters, and driven by different meteorological data generated from different institutions. We also describe the development and use of simulated observations for exercise conditions. We present three case studies for hypothetical events in Iceland, through which we assess our ability to scientifically interpret and evaluate our forecasts. We finish by outlining the lessons learnt and describe the resulting improvements made to the forecasting capability at the London VAAC.

The London VAAC forecasting process

The London VAAC is staffed by specialist operational meteorologists who provide continuous support (24 h a day, all year round). Forecasts are issued as Volcanic Ash Advisories (VAAs) and Volcanic Ash Graphics (VAGs), while ash is in the atmosphere (London VAAC advisories and graphics webpages 2023). These indicate the expected location of the ash cloud up to 18 h ahead of the issue time. The Met Office also issues supplementary ash concentration charts (London VAAC concentration chart webpages 2023). These indicate three ash contamination zones in the atmosphere, showing concentrations in the ranges 200–2000 $\mu g\ m^{-3}$, 2000–4000 $\mu g\ m^{-3}$, and greater than 4000 $\mu g\ m^{-3}$ (ICAO 2016).

The first forecast is issued on receiving a Volcano Observatory Notification to Aviation (VONA). VONAs are generated by the responding volcano observatory whenever there is a change in aviation colour code; this is an alert level

system coordinated by the ICAO that indicates whether a volcano is in a normal state, experiencing unrest or erupting, or is undergoing a significant change in the behaviour of an ongoing eruption (Lechner et al. 2017). After the initial issuance, forecasts are provided at least every 6 h, often aligned to 00:00, 06:00, 12:00, and 18:00 UTC. However, if information is received by the VAAC which contradicts the current forecast, for example, if comparison of model simulations against observations indicates that ash lies outside the current forecast area, or if the VONA indicates a significant change in the eruption behaviour, then an updated set of products can be issued.

Model simulations

The atmospheric dispersion model NAME (Numerical Atmospheric-dispersion Modelling Environment) is used by the London VAAC, via a graphical user interface, to generate forecasts of the transport and dispersion of the ash cloud in the atmosphere (Jones et al. 2007; Beckett et al. 2020). Simulations can be generated with meteorological datasets from a range of different numerical weather prediction (NWP) systems, including output from the Met Office's Unified Model (UM) (Walters et al. 2019) and the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS) model.

The NAME model runs are initialized with a set of eruption source parameters (ESPs) which describe the release of ash into the atmosphere, these include the volcano location, the source geometry, the top and bottom height of the eruption plume (the depth over which ash is released into the atmosphere), the mass eruption rate (MER), the duration of ash release, and the physical characteristics of the particles (their size range, shape, and density). Full details of the London VAAC modelling setup are provided in Beckett et al. (2020). Often, information provided in the VONA can be used to set many of the key ESPs needed for the model runs. However, information on the MER and the physical characteristics of the ash is typically not observed in real time. Instead, the default approach is to use the Mastin relationship to calculate the MER from the observed plume top height (Mastin et al. 2009). Alternatively, the MER can be calculated from a buoyant plume model (Devenish 2013) or if satellite retrievals of ash column load are available, then a Bayesian inversion tool, called InTEM, can be used to define the time-varying mass eruption rate and vertical distribution of ash at the vent (Pelley et al. 2021). It is expected that a large fraction of the total erupted mass is deposited close to the source, so as a default, the London VAAC only considers particles with diameters ≤ 100 μm and assumes that just 5% of the total calculated erupted mass survives near-source processes

and makes up the distal fine ash fraction (DFAF; Webster et al. 2012). The London VAAC has the option to use a default particle size distribution (PSD) and particle shape or choose from a set of options which represent variability in these parameters, based on observations from past events (Saxby et al. 2018, 2020; Osman et al. 2020). It should be remembered that the default MER, DFAF, and PSD can all be varied in the operational system.

Model simulations are very sensitive to the ESPs used (e.g. Scollo et al. 2008; Beckett et al. 2015; Dioguardi et al. 2020). Time-varying ESPs are applied to reflect different eruption phases, and the system allows for updates to past ESPs if further observation data become available. The London VAAC modelling system can also run multiple scenarios, in which different plume heights, MERs, DFAFs, PSDs, and particle shapes can be set, so that the sensitivity of the forecasts to any uncertainty on these ESPs can be assessed.

Forecast evaluation

During an eruption, atmospheric dispersion model simulations of the ash cloud are evaluated against observations, primarily satellite and lidar retrievals. At the London VAAC, software called the Volcanic Ash Intervention Tool (VAIT) is used, which overlays satellite imagery and dispersion model output. Satellite observations can provide information on the location and spatial extent of the ash cloud and its properties, including the ash cloud plume height and total column mass loadings in the atmosphere expressed in grams per square metre, which indicate the total mass of ash in a vertical column above the earth (Prata and Lynch 2019). At the Met Office, we use data retrieved from both geostationary and polar orbiting satellites. The geostationary Meteosat Second Generation (MSG) satellite with its Spinning Enhanced Visible and Infrared Imager (SEVIRI) provides multi-spectral data in 12 channels between 0.6 and 13.4 μm at a high temporal frequency of one image every 15 min, making it a valuable tool (Francis et al. 2012). The Met Office also maintains a network of Raman and polarisation lidars, ceilometers, and sun photometers across the UK. The network can indicate the presence of volcanic ash at sensor locations, provide height resolved estimations of ash and sulphate aerosol concentrations (expressed in g m^{-3}), and give cloud top height (Adam et al. 2018; Osborne et al. 2019, 2022).

The London VAAC is supported by a group of scientists at the Met Office who are experts in atmospheric dispersion modelling and observations, including lidar and satellite retrievals of volcanic clouds. During an eruption, these scientists meet regularly with the London VAAC operational meteorologists at 'Forecast Evaluation Meetings' to share knowledge and support decision-making. Forecasts

are evaluated by comparing dispersion model simulations generated using varying scenarios (ESPs) to the observations. These meetings are also used to identify where there is confidence in the forecast and where there are uncertainties.

Given an eruption in Iceland, there are three key institutes who provide scientific support to the London VAAC: the Icelandic Volcano Observatory based at the IMO, the British Geological Survey (BGS), and the UK's National Centre for Atmospheric Science (NCAS), although other institutes may be called upon. During an eruption, the Met Office may seek advice on the behaviour of the eruption, for example, the properties of the ash and whether the activity is expected to escalate or wane. To enable the provision of external scientific expertise into London VAAC during an event, a Science Advisory Meeting, hosted by the Met Office, may be held. This provides a platform to discuss, for example, the eruption scenario (including ESPs); compare model outputs generated by different institutions; and evaluate model simulations against observations, with the purpose of ensuring that the best possible scientific advice is used to generate the London VAAC forecasts. These meetings are time limited due to the nature of the operational response, typically occurring once a day and lasting only 30 min.

Forecasts can benefit from an assessment of how different modelling approaches, variations in model setups (including ESPs), model physics, and driving meteorological data impact the predicted extent and concentration of ash (Witham et al. 2007; Bonadonna et al. 2012; Plu et al. 2021). Through the exercises described herein, we have practised comparing results from model simulations generated by external collaborators, effectively generating a multi-model ensemble.

Exercise methodology

The SCI-VOLCICE exercises are led and coordinated by the Met Office; they have a duration of 1 day and are timed to coincide with a CATII VOLCICE exercise. Planning is led by an exercise coordinator who defines the aim of the exercise, identifies the institutes who need to be involved, and communicates the roles of each 'player'. Colleagues at the IMO design and prepare the event scenario, by choosing a target volcano, deciding on the timing of the eruption, and defining its intensity, by setting the plume height as function of time. The event scenario is designed in-line with the aims of the exercise, and we have practised responses to a range of different scenarios, including, for example:

1. Events where observations of the eruption column are limited and minimal information is provided on the ESPs
2. Events where the ESPs change with time

3. Events where model simulations and observations of the distal ash cloud disagree

During the exercise, the natural hazard specialist on shift at the IMO communicates the scenario through a series of VONAs. After each exercise, a debrief is held to identify the lessons learnt; the next exercise is then designed to ensure that any problems identified have been adequately addressed.

Roles

SCI-VOLCICE exercises are overseen and driven by an exercise coordinator; the roles played and the interactions between the players are shown in supplementary Fig. S1. On the day of the exercise, the natural hazard specialist at the IMO initiates the exercise and generates the VONAs. Scientists from the IMO and BGS provide expert advice on the event, including insight into the ESPs, the behaviour of the eruption, and the potential associated hazards. A London VAAC operational meteorologist generates and issues the forecasts, the VAG, VAA, and ash concentration charts. They are supported by the London VAAC manager and the team of scientists (atmospheric dispersion modellers and observation specialists) at the Met Office. In addition, for our exercises, scientists at the NCAS and BGS are asked to share any model simulations conducted at their institutes. During a real event, collaborators from other volcano research institutes may also be invited to provide input.

Components

Figure 1 shows the key components of our SCI-VOLCICE exercises and the sequence of tasks, which align with the response procedures outlined for an actual event. Exercises are initiated at ~08:00 UTC in the morning with the issuance of a VONA stating that an eruption is imminent in Iceland. This is followed by a VONA stating that an ash-producing eruption has started and is confirmed by monitoring data. This instigates the first model runs at the London VAAC and issuance of an initial 18-h forecast. In a real event, forecasts are, at a minimum, issued at 00:00, 06:00, 12:00, and 18:00 UTC thereafter. Because our exercises are constrained to 1 day, after the initial forecast issuance, subsequent forecast issuances are limited to 12:00 UTC and 18:00 UTC.

A key component of our SCI-VOLCICE exercises is to test the forecast evaluation procedures at the Met Office. To enable this, we use simulated observations; the generation of both simulated satellite and lidar imagery is explained in the following section. We compare the simulated observations to our model simulations in a forecast evaluation meeting. Our exercises have also practised our procedures for
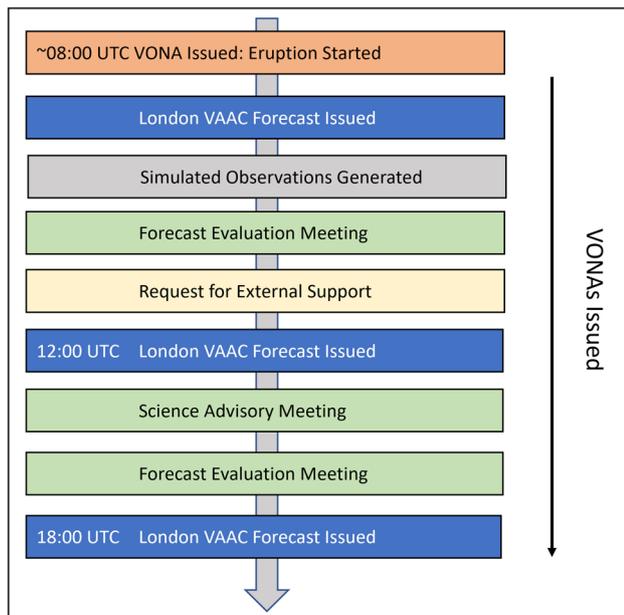
Fig. 1 The components of our exercises and the sequence in which they occur, which reflect the anticipated response for a real event. The issuance of London VAAC forecasts is indicated by the blue boxes, meetings held are indicated by the green boxes, and activities are indicated by the yellow boxes. The simulated observations are generated for exercises only and are indicated by the grey box

requesting external scientific support, including the sharing of any additional model simulations and the delivery of the Science Advisory Meeting. This tests our communication procedures; we arrange the meetings without forewarning, using agreed institute email addresses to contact our external collaborators and set up video conferencing for the Science Advisory Meeting in real time during the exercise. Following the outcomes of the Science Advisory Meeting, we practise the pull-through of scientific advice to the London VAAC. Scientists at the Met Office meet and discuss any changes that we deem necessary to the operational model configurations. These decisions must be made very quickly to feed in to the 18:00 UTC forecast (Fig. 1).

Generating simulated observations

Simulated satellite images, of volcanic ash and meteorological clouds, are created from simulated spectral radiances $(mW\ m^{-2}\ sr^{-1}\ cm^{-1})$ (Millington et al. 2012). The simulated spectral radiances are generated from a radiative transfer model which uses NAME modelled ash concentration data and numerical weather prediction data as inputs. The radiances are simulated in the infrared for SEVIRI channels centred at 8.7, 10.8, 12.0, and 13.4 μm and allow for the creation of dust RGB (red–green–blue) imagery at the resolution of the SEVIRI instrument (3 km). The simulated

spectral radiances are also used as input to a volcanic ash retrieval (Francis et al. 2012) to estimate the properties of the simulated volcanic ash cloud such as ash column loadings $(g\ m^{-2})$. For our exercises, we used simulated dust RGB composite images to identify the volcanic ash clouds, which often have a strong red signal, or a bright yellow signal if volcanic sulphur dioxide gas is also present.

We have, for the first time, made use of these simulated satellite observations in our SCI-VOLCICE exercises. The NAME output ash concentration data are transferred to the simulated satellite imagery production, which is run on a fixed processing schedule in operational mode. It takes around 45–60 min from the issuance of a VONA for the simulated satellite imagery to be generated as part of the exercise (Fig. 1). We have typically used one simulated satellite image from around 10:00 UTC during our SCI-VOLCICE exercises for comparison with the forecast issued by the London VAAC.

We have also developed new simulated lidar imagery for these exercises. These are generated using existing raw lidar data for a generally cloud free day, which are then blended with NAME simulations of ash concentration. The raw lidar data are first inverted to give vertical profiles of aerosol backscatter, extinction, and depolarisation ratio (Klett et al. 1985). The NAME ash concentration data are then used to simulate vertical profiles for ash backscatter, extinction, and depolarisation ratio. To simulate the signal-to-noise ratio of real lidar observations, random noise is added to these simulated profiles. The magnitude of the random noise depends on altitude (distance from the lidar), ash concentration, and the amount of total aerosol (real and simulated) between the lidar and the simulated ash layer. Finally, the observed and simulated profiles are combined to create simulated lidar plots. The simulated plots of range-corrected signal and volume depolarisation ratio (VDR) are useful in indicating how the Met Office lidars would observe the simulated event under realistic conditions, including background aerosol loads. In particular, the images indicate the magnitude of ash concentration and the vertical extent of ash layers that could attenuate the lidar signal and so make higher ash layers unobservable by the lidar. The imagery can be generated within a few minutes of the NAME simulations and show 24 h of simulated observations.

Unlike the satellite products, which provide synoptic scale monitoring, the lidar network provides coverage over only nine UK sites. For the purpose of the exercises, where NAME simulations show volcanic ash over at least one UK site, simulated lidar data can be created for those sites. However, where the NAME simulations do not forecast ash over a UK lidar site, then one or more 'notional lidars' can be agreed upon at other locations that are impacted by the ash cloud and simulated lidar data created for these locations.

The use of notional lidars in this way allows for the inclusion of the lidar component in exercises where the simulated ash cloud does not impact the UK.

To help promote discussion during the SCI-VOLCICE exercises, we replicate a situation where observations and model forecasts differ. To do this, the NAME model data used as input to the simulated satellite and lidar imagery need to be different to the operational NAME forecast issued by the London VAAC. During the exercises, once the VONA is received, a separate NAME model forecast is generated, with different ESPs (for example, a higher plume height) to the London VAAC forecast, and/or different meteorological data to drive NAME. The information on this setup is generated by team members supporting the exercise coordinator and, to encourage unbiased discussion in the Forecast Evaluation and Science Advisory Meetings, is not provided to the scientists involved in the exercise.

Exercise case studies

Exercise Case Study 1: 14 June 2019

This was our first SCI-VOLCICE exercise and as such practised a limited number of the components, these being the issuance of VONAs and London VAAC forecasts and the generation of dispersion model simulations by our external collaborators.

A series of VONAs were distributed by the IMO, which confirmed that an eruption had started at Katla volcano at 08:30 UTC and that ash was detected up to 18 km above sea level (asl). The VONAs also stated that the ash was spreading to the south, according to current weather conditions. Key locations in Iceland referred to in the exercises are shown in Fig. 2; all the VONAs are given in the Supplementary Material (VONAs 1–4 Exercise Case Study 1 14 June 2019).

During the exercise, collaborators at the NCAS and BGS ran the dispersion models FALL3D (Folch et al. 2020) and HYSPLIT (Stein et al. 2015), with Global Forecasting System (GFS) meteorological data, generating forecasts of the expected transport and dispersion of the ash cloud using the information provided in the VONAs. A range of products were produced including ash concentration maps, deposit loading maps, and maps of total column mass loadings which are given by way of example in Fig. 3.
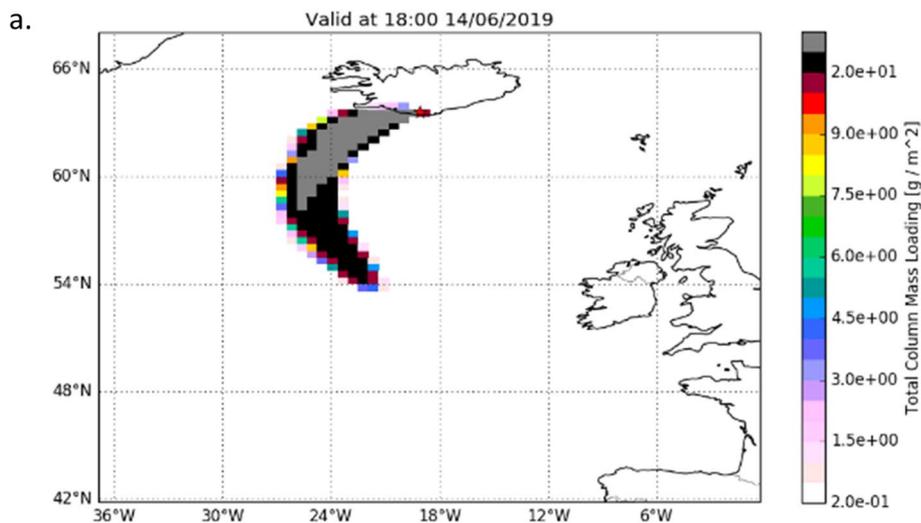
Following the exercise, we held a debrief at which we shared our model simulations and discussed our ability to make use of these during a real event. It was agreed that the variation in model setups, model physics, and meteorological data used across the groups was advantageous, as it allows us to consider how different modelling approaches impact the predicted extent and concentration of ash during an event. However, to allow meaningful comparison, the model setups used must be shared. As
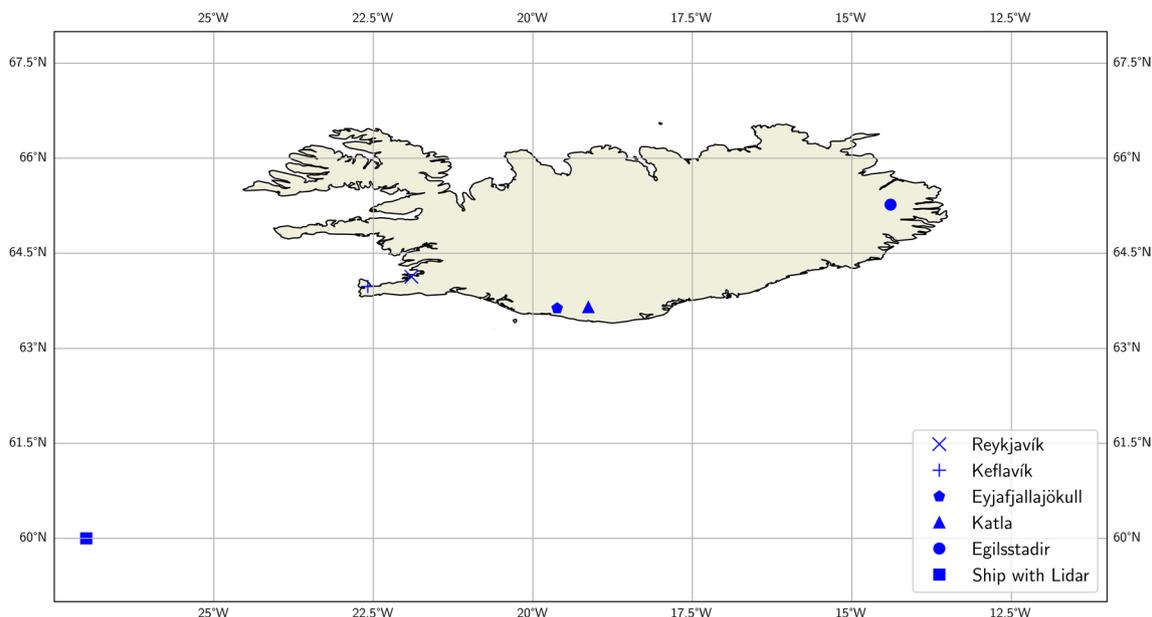

Fig. 2 To show key locations in Iceland: the capital of Reykjavik where IMO is located, Keflavik where a hypothetical lidar was placed for Exercise Case Study 3, the volcanoes used in the Exercises (Katla and Eyjafjallajökull), the radar site at Egilsstaðir used in Exercise Case Study 2, and the notional ship containing a lidar used in Exercise Case Study 3

Fig. 3 Exercise Case Study 1: forecast total column mass loadings generated using **a** NAME with UM global meteorological data and **b** HYSPLIT and **c** FALL3D with GFS data, at 18:00 UTC on the 14th of June 2019. Note that to convey a sense of the challenges faced in the exercises, all plots (including simulated observations) presented in this paper are those produced in real time and discussed during the exercise itself; they have not been modified to improve visualisation for the paper. The lack of agreement on plotting options makes it very difficult to compare the output
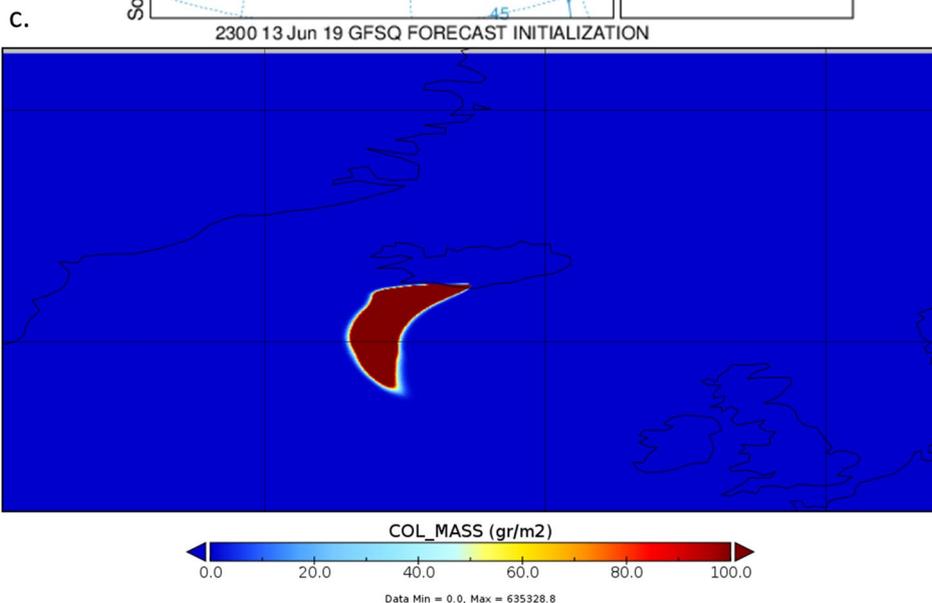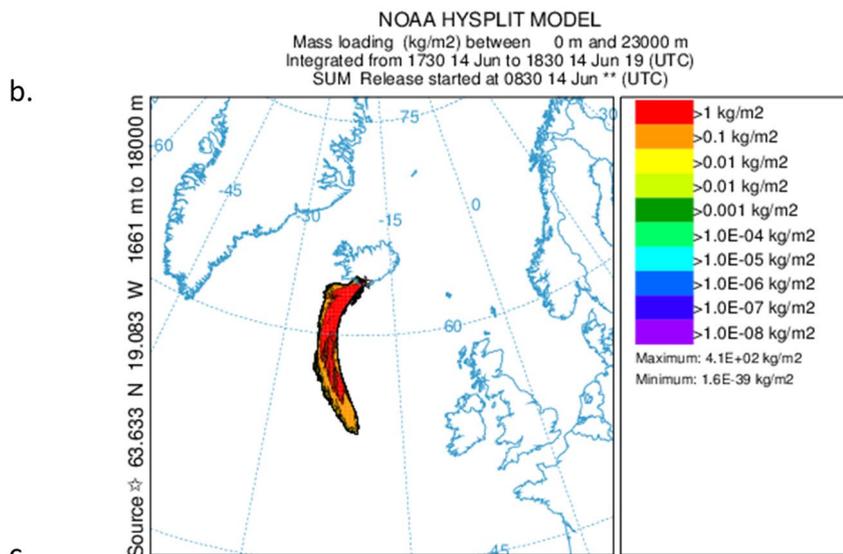
Table 1 Information on model setups which should be shared to enable meaningful comparison

Model setup information
Forecast length
Output temporal frequency and averaging period
Vertical and horizontal grid resolutions
Particle release rates (for Lagrangian models)
Source geometry
Plume height
MER
Source start and end times
PSD
Particle density
Particle shape
Driving meteorological dataset used: source, resolution, and model cycle

Table 2 Plotting setup used by all collaborators to generate comparable total column mass loading output. Note the non-linearity of the contour scale used

Projection	Plate Carrée
Output times	00:00 UTC, 06:00 UTC, 12:00 UTC, 18:00 UTC
Contour scale	0.2, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5, 9.0, 9.5, 10.0, 20
Units	g m^{-2}

such, we identified a list of key information which should be prepared and distributed during event response which is given in Table 1.

In addition, if output data are plotted using a range of different approaches and software packages, this significantly limits our ability to compare the output constructively (Witham et al. 2007). We agreed that, for future events and exercises, we would generate plots depicting total column mass loadings, as these are directly comparable to satellite retrievals and therefore important for evaluating model output. It was also agreed that plots should use the same projection, units, contouring, and colour scales, as well as similar domain extents and output times (Table 2). This allows any obvious visual differences in model output to be identified quickly and easily. To facilitate generation of comparable plots, we developed a Python package (named ash-model-plotting) to read outputs from the different models into a consistent structure. This is open-source and is shared through the GitHub repository hosted by the BGS (Stevenson et al. 2023).

Exercise Case Study 2: 11 December 2020

This SCI-VOLCICE exercise included all of the components shown in Fig. 1 and was designed to practise our response to a scenario in which there was uncertainty associated with the ESPs. This was built on the lessons learnt from previous exercises; notably, we tested our ability to generate comparable dispersion model outputs in real time using the new agreed procedures and the NCAS and BGS used ash-model-plotting.

A series of VONAs were issued by the IMO advising that an ash-rich eruption had started at Eyjafjallajökull on 10/12/2020 (dd/mm/yyyy, see supplementary material VONAs 1–3 Exercise Case Study 2 11 December 2020). These stated that the radar network in Iceland was suffering from major disruption, and as such the uncertainty on the plume height was significant.

Simulated satellite imagery was created (see Supplementary Table S2 for the NAME model setup used); the dust RGB and retrieved total column mass loadings of the ash cloud at 09:00 UTC on 11/12/2020 are shown in Fig. 4. It should be noted that simulated observations were only generated for one timestamp during the exercise. During a real event, satellite observations would be automatically processed at a high temporal frequency, approximately every 15 min, and would be much easier to align with the model output times.

Operational meteorologists met with Met Office support scientists to discuss the forecast produced by the London VAAC and the simulated observations available. Significant discrepancies were identified between the London VAAC forecast and the observations (Figs. 4 and 5a). The uncertainty on the ESPs was discussed and identified as a possible reason for the differences. It was agreed that Met Office support scientists would contact external collaborators with a request that additional model simulations be conducted and shared if possible and that a Science Advisory Meeting would be convened.

Collaborators at the NCAS and BGS ran the atmospheric dispersion models HYSPLIT and FALL3D respectively with GFS meteorological data using information provided in the VONAs, but all other model setup options were at the discretion of the modeller. As such, the model setups varied, with each partner making different decisions for the choice of forecast length, particle release rate (HYSPLIT only), the temporal and horizontal grid resolution of the output, and the ESPs, including plume height, MER, and PSD. The NCAS and BGS shared their model setups (Table 3 and Supplementary Table S1 for the PSDs used) and produced total column mass loading plots of the simulated volcanic ash cloud from Eyjafjallajökull using the ash-model-plotting Python package developed after the June 2019 exercise (Fig. 4), which used the same projection, units, contouring, and colour scales as the London VAAC output.

Fig. 4 Exercise Case Study 2: simulated satellite imagery at 09:00 UTC 11/12/2020, **a** dust RGB and **b** retrieved total column mass loadings (g m^{-2}) generated in real time during the exercise. All total column mass loadings greater than 10 g m^{-2} are indicated by the red contour. Grey areas within the ash cloud indicate locations where the satellite is unable to retrieve data due to high mass loadings
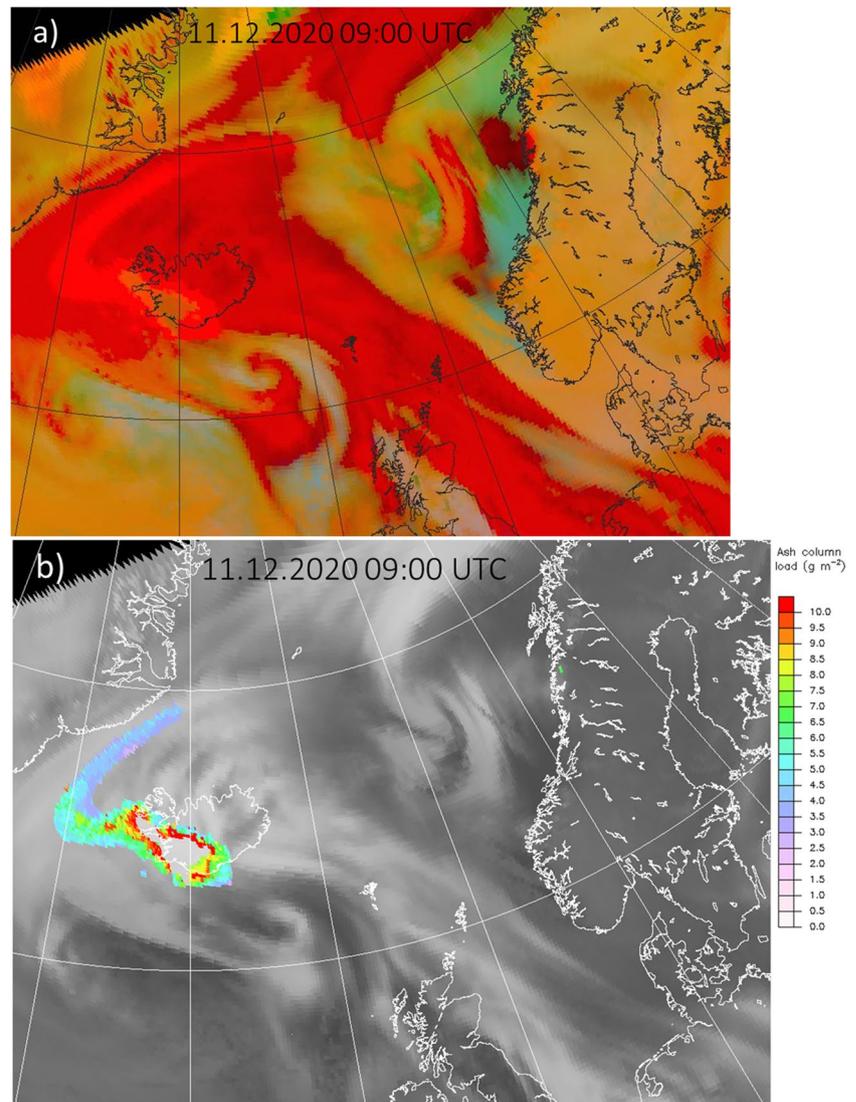
A Science Advisory Meeting was convened and attended by all exercise participants. The meeting focused on comparison of the dispersion model output to simulated satellite imagery. The modelled ash cloud generated by the London VAAC indicated ash to the north-west of Iceland, but the simulated satellite imagery suggested that the forecast was under-representing the spread of the ash cloud, particularly to the south of the eruption source and to the north along the Greenland coastline. Furthermore, there were discrepancies between the different model forecasts, with the BGS simulations suggesting that the ash cloud was splitting, causing some ash to be transported to the south, which was not observed in the simulated satellite imagery or the London VAAC and NCAS simulations. IMO then presented some additional model simulations using NAME which they generate to aid with hazard planning for the local population in Iceland (IMO Dispersion Modelling 2020). These were initiated with a plume height varying between 8 and 13 km

asl; a corresponding MER between 2.7×10^5 and 5.8×10^6 kg s^{-1}, calculated using the Mastin relationship (Mastin et al. 2009); and the total grain size distribution (TGSD) of tephra from the 2004 Grímsvötn eruption (Höskuldsson et al. 2018) with grain sizes up to 32 mm. The IMO products indicate the predicted atmospheric concentrations and deposit cumulative mass loadings over Iceland. These simulations also suggested that the ash cloud was splitting, as shown in Fig. 6. The resulting discussion led to a general agreement that differences were perhaps being caused by the uncertainty associated with the plume height, MER, and PSDs, and hence, the significant variation in values assigned to these parameters by the different players. Notably, the BGS used the plume height data provided in the VONA as input into the REFIR (Real-time Eruption source parameters FutureVolc Information and Reconnaissance system) software tool to determine average, maximum, and minimum MERs (Dürig et al. 2018; Dioguardi et al. 2020). In
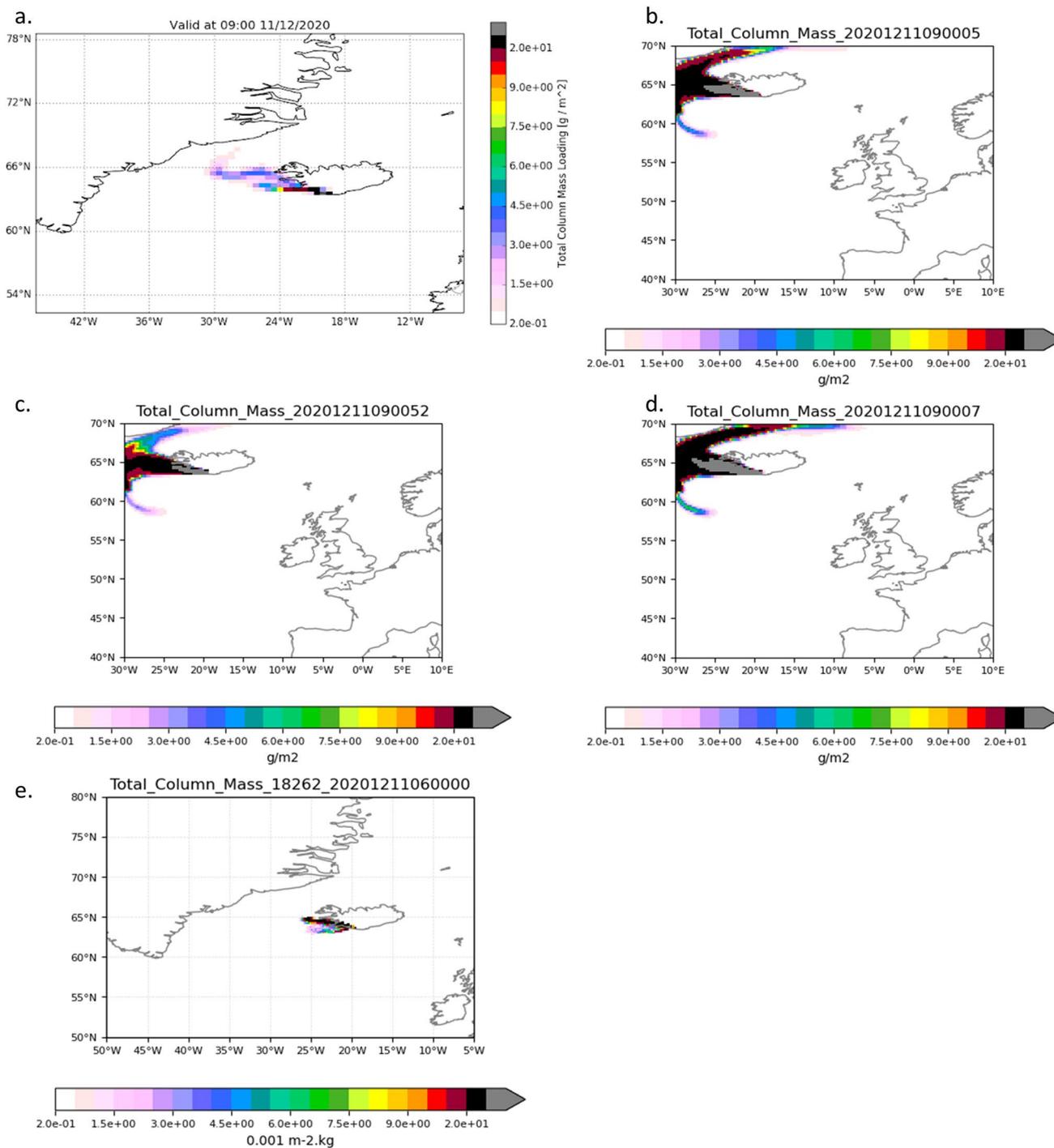
Fig. 5 Exercise Case Study 2: plots of modelled total column mass loadings generated in real time during the exercise using **a** NAME with UM (Global) meteorological data (Met Office), **b** FALL3D using GFS meteorological data (BGS) and the average MER output from REFIR, **c** using the minimum MER, **d** using the maximum MER, and **e** using HYSPLIT with GFS (NCAS). The validity times are given in the sub-title of each plot. Note the non-linearity of the contour scale used to indicate the very highest total column mass loadings. The black contour indicates column loadings of 100 g m^{-2} and all total column mass loadings greater than this are shown by the grey contour

Table 3 Exercise Case Study 2: model setups used by the Met Office, NCAS, and BGS. The PSDs used are provided in the Supplementary Material (Table S1). Plume height data generated by REFIR, as used by the BGS, are provided in the Supplementary Material (Table S3)

Partner	Met Office	BGS	NCAS
Model	NAME	FALL3D	HYSPLIT
Forecast length	24/120 h	96 h	24/48 h
Output temporal resolution	09Z, 15Z, 21Z, 03Z 1-h average	1-h time step (instantaneous)	0Z, 06Z, 12Z, 18Z 1-h average
Horizontal grid resolution (lon-lat)	0.56×0.37°	0.25×0.25°	0.2×0.2°
Vertical grid resolution	Total column	1 km (total column)	Total column
Particle release rate	15,000/hr	NA	2000/h (24-h forecast) 1000/h (48-h forecast)
Source geometry	Uniform line	Uniform line	Uniform line
Plume height	8 km 17:30 10/12/2020 13 km 08:00 11/12/2020 10 km 11:00 11/12/2020	REFIR assessed time series	4 simulations 8, 13, 18 km (1st VONA), 10 km (2nd VONA)
Mass eruption rate (MER)	Mastin 5% (8 km: 1.5E4 kg/s, 13 km: 1.67E5 kg/s, 10 km: 4.66E4 kg/s)	REFIR assessed Average: 2.76E7 kg/s Maximum: 3.49E7 kg/s Minimum: 2.05E7 kg/s	Mastin 3% (8 km: 6.7E3 kg/s, 13 km 9.2E4 kg/s, 10 km: 2.3E4 kg/s)
Source start	10/12/2020 17:30 UTC	10/12/2020 17:30 UTC	18:00 10/12/2020 UTC
Source end	Continuous release	11/12/2020 12:00 UTC	Continuous release
Particle size range	0.1–100 μm	1–500 μm	0.6–20 μm
Particle density	2300 kg m^{-3}	2000 kg m^{-3}	2500 kg m^{-3}
Particle shape	Spherical	Spherical	Spherical
Driving meteorology	Global UM	GFS	GFS
Meteorological data resolution	10 km	0.25°	0.25°

Fig. 6 Exercise Case Study 2: simulated ash concentrations in the atmosphere (g m^{-3}) and deposit mass loadings (kg m^{-2}) generated by the IMO during the exercise using NAME as an operational product
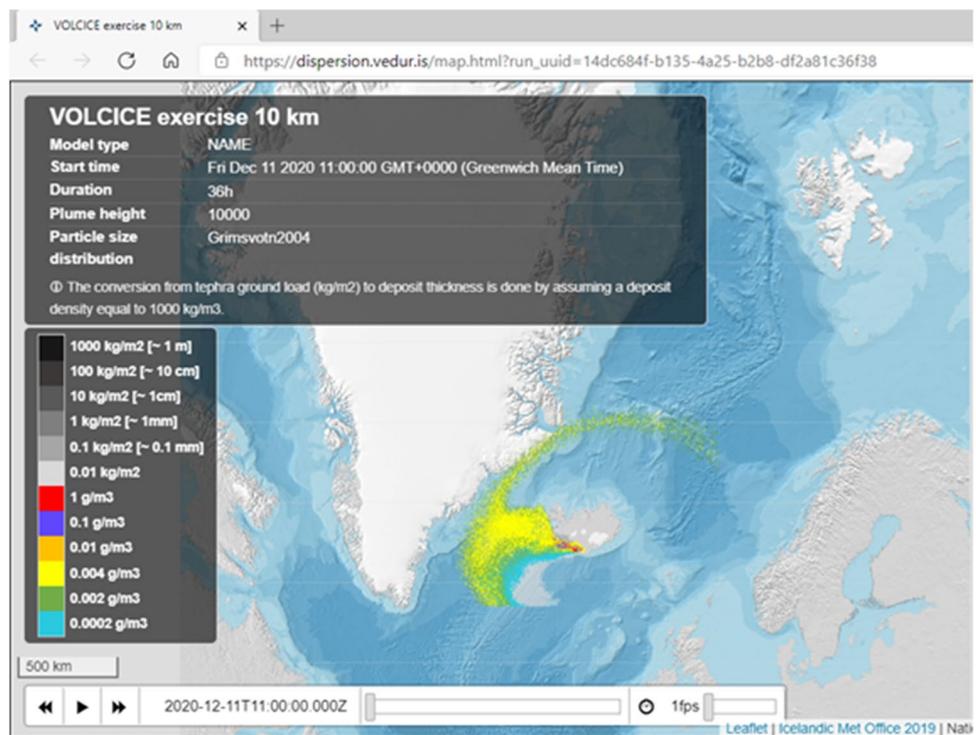
addition, the BGS and IMO used the full MER and TGSD of the tephra, with diameters of up to 500 μm and 32 mm respectively, while the London VAAC and NCAS considered particles up to diameters of 100 μm and 20 μm respectively and applied a scaled MER to represent only a fine ash fraction in their model setups (Table 3). The use of coarser particles in the IMO and BGS modelling was identified as a possible reason for the differences in modelled location of the ash cloud, with coarser particles, which settle more rapidly in the atmosphere, being advected by the northerly wind at lower levels and causing splitting of the forecast cloud. The NCAS was unable to represent temporally varying plume heights in their setup and as such performed four different simulations with plume heights of 8, 10, 13, and 18 km, fixed for the duration of the run. Output using a plume height of 10 km is shown in Fig. 5e.

Following the Science Advisory Meeting, the atmospheric dispersion scientists at the Met Office finalised the decision on the most appropriate ESPs to use. It was discussed that the buoyant plume rise scheme (Devenish 2013) could be used to constrain the MER and vertical extent of the release, and the plume height or DFAF could be adjusted, and this was communicated to the London VAAC duty meteorologist in time for setup of the 18:00 UTC forecast.

Exercise Case Study 3: 10 June 2022

This SCI-VOLCICE exercise included all components given in Fig. 1, with a focus on practising knowledge exchange of the eruption scenario and, for the first time, included simulated lidar imagery in the forecast evaluation process.

A series of VONAs were issued by the IMO advising that an ash-rich eruption was ongoing at Eyjafjallajökull (see Supplementary Material VONAs 1–3 Exercise Case Study 3 10 June 2022). In this exercise, we wished to practise our response several days into an event; as such, the eruption start date was set to be before the exercise. The initial VONA stated that the eruption had started at 08:15 UTC on 07/06/2022 and, using radar data and web cameras, the plume height had been assessed to be 15 km asl. The next VONA confirmed that the eruption was ongoing as of 10/06/2022 (the day of the exercise) and, using radar data and web camera images, the plume height was now assessed to be 13–15 km asl. London VAAC issued forecasts on 10/06/2022, using an eruption start time of 08:15 UTC on 07/06/2022 in the model simulations. Figure 7 shows the supplementary ash concentration charts that were produced for 12:00 UTC on 10/06/2022.

Simulated lidar retrievals were generated for notional lidars located at Keflavik (Iceland) and on a ship located in the North Atlantic (60° N, 27° W), between 11:00 UTC on 10/06/22 and 10:00 UTC on 11/06/22 and are shown in Fig. 8. The simulated satellite imagery, dust RGB, and retrieved total column mass loadings of the ash cloud at 10:00 UTC on 10/06/2022 are presented in Fig. 9.

Operational meteorologists and Met Office support scientists evaluated the forecast produced given the simulated observations available. The Lidar observations suggested that the base of the ash cloud over Keflavik was at ~10 to 11 km asl at 12:00 UTC (outline 1 in Fig. 8a), although it was noted that the high ash concentrations quickly attenuated the lidar signal, meaning that any higher layers would likely not be observed. The concentration charts indicated ash at all layers between FL000 and FL550 (0–~17 km asl) at this time, with concentrations > 4000 ug m^{-3} in FL200-350 (~6–11 km asl). Observations from the lidar on board the notional ship in the North Atlantic (60° N, 27° W) showed an ash plume extending from around 5 km asl to 7.5 km asl at around 12:00 UTC on 10/06/2022 (outline 3 in Fig. 8b), with a separated ash layer in the boundary layer below 2.5 km asl, as indicated by the high depolarisation values (outline 4 in Fig. 8b). Again, the range-corrected signal was quickly attenuated, and any higher layers would likely not be observed. The forecast ash concentration charts indicated that, in the atmosphere above the ship location, concentrations were > 4000 μg m^{-3} between FL000-FL200 (~ < 6 km asl) and 200–2000 μg m^{-3} at FL200-300 (~6–11 km asl).

The predicted direction of travel of the modelled ash clouds in the forecast ash concentration charts generally agreed well with the simulated satellite imagery, although the NAME simulations over-estimated the mass loadings in the atmosphere with respect to the simulated satellite mass loadings (Figs. 9 and 10). It should be noted, however, that due to choices made in the setup of the model runs, there is discrepancy between the timestamp of the simulated satellite imagery, provided at 10:00 UTC, and the model output which is at 12:00 UTC.

Following discussions at the forecast evaluation meeting, it was agreed that Met Office support scientists would run additional dispersion model scenarios, that external collaborators would be asked to provide additional model simulations if possible, and that a Science Advisory Meeting would be convened. The BGS provided information on past activity at Eyjafjallajökull and advice on the eruption scenario and possible future activity (Table 4). Additional Met Office scenario model simulations using a DFAF reduced to 2% and a coarser PSD were conducted, with the aim of considering possible reasons for the mismatch in predicted mass loadings with the satellite retrievals. Output total column mass loadings from these runs, as well as additional model output generated by NCAS, are shown in Fig. 10, and the model setup options used are provided in Table 5.

A Science Advisory Meeting was attended by all exercise participants. At the meeting, knowledge of the ongoing eruption was shared by IMO, and information on the expected future activity and controls on grain size of ash in volcanic

Fig. 7 Exercise Case Study 3: ash concentration charts issued by the Met Office during the exercise, indicating the modelled ash cloud at 12:00 UTC 10/06/2022 over three different flight levels **a** FL000-FL200, **b** FL200-FL350, and **c** FL350-FL550. Model simulations were generated using the model setup outlined in Table 5; a 5% DFAF and the default PSD were used. The locations of the notional lidars are indicated by the black circles
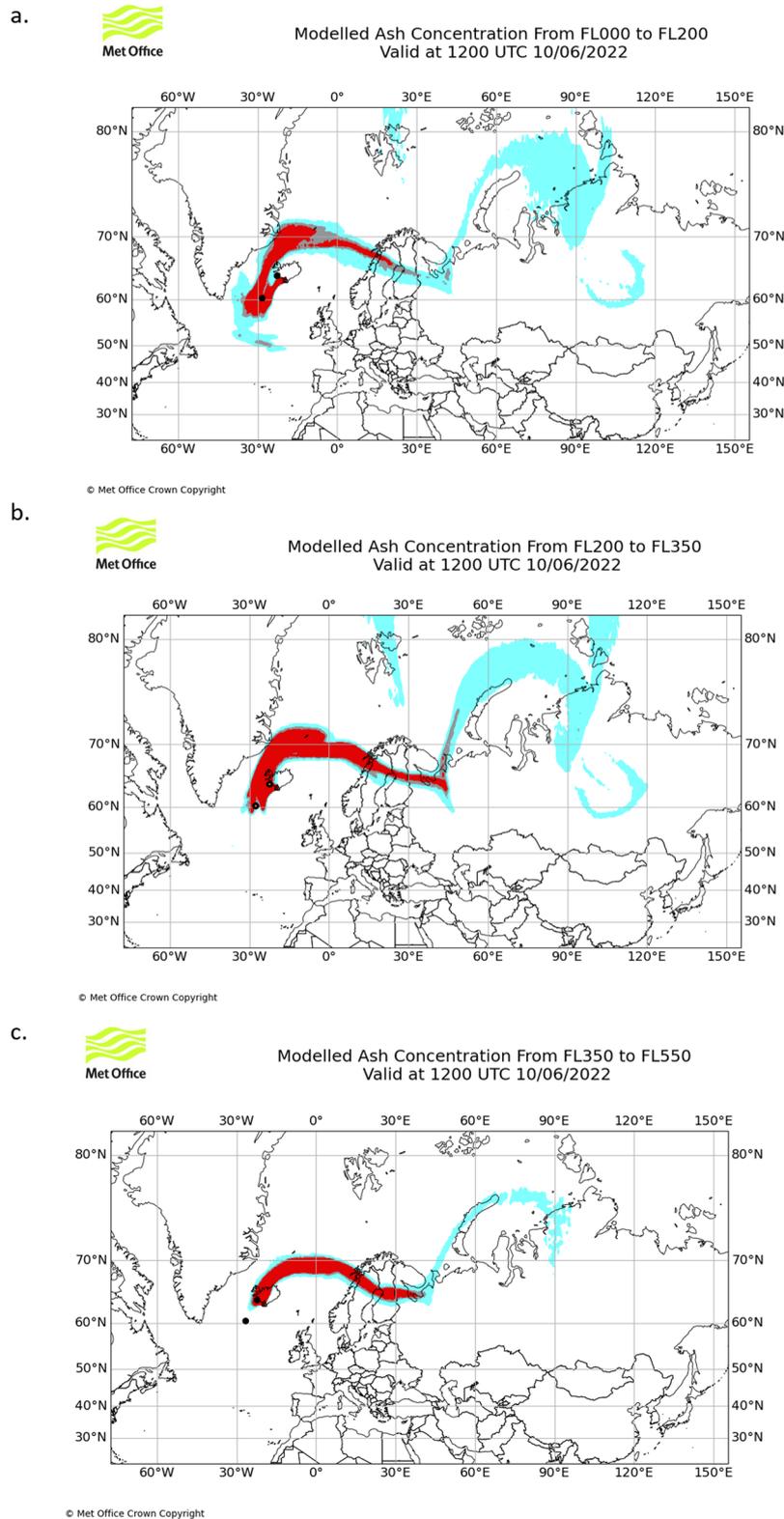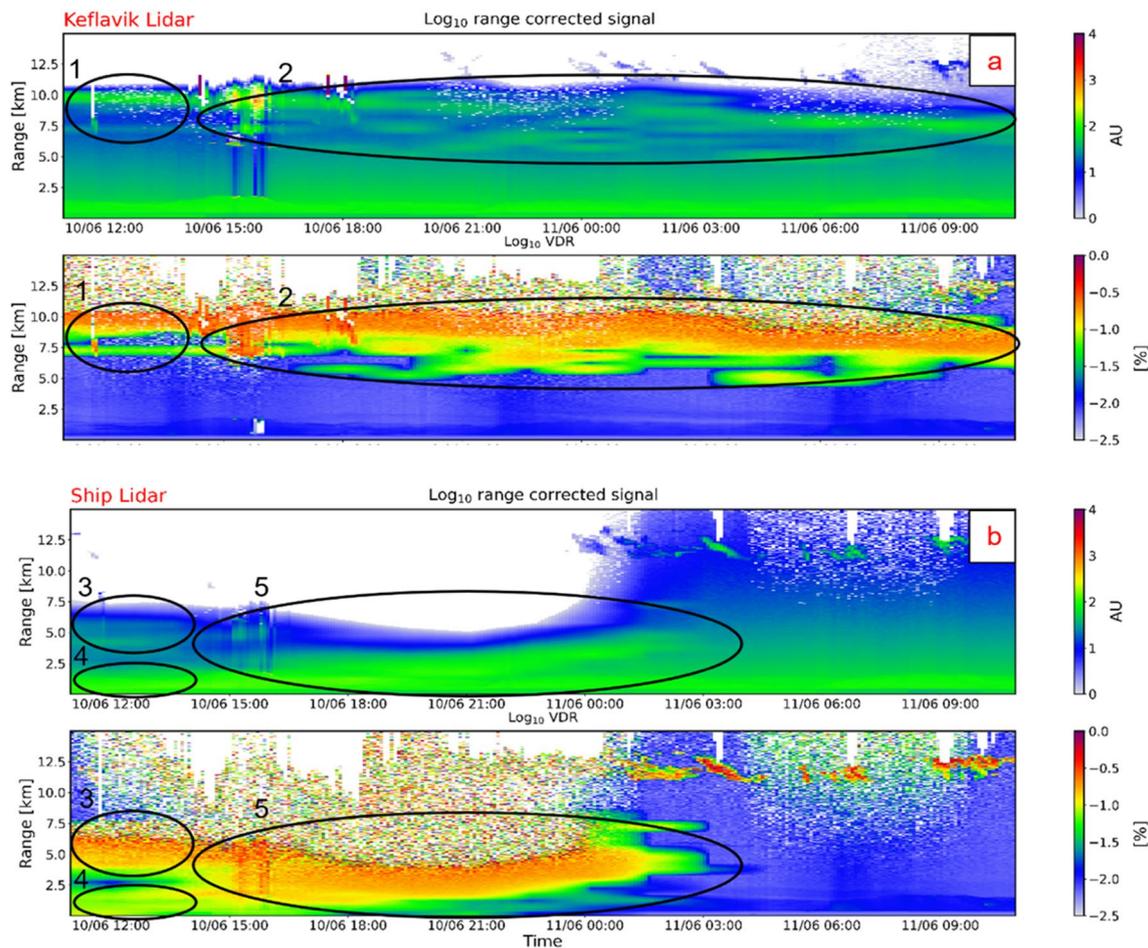
Fig. 8 Exercise Case Study 3: **a** Simulated lidar observations at Keflavik (Iceland) for range-corrected signal and volume depolarisation ratio (VDR) and **b** the same for a lidar on board a (hypothetical) ship located in the North Atlantic (60° N, 27° W), between 11:00 UTC 10/06/22 and 10:00 UTC 11/06/22. Ash plumes are outlined in black. Outlines 1, 3, and 4 show the presence of the ash layer at 12:00 UTC on the 10/06/2022 which is comparable to the output time of the ash concentration charts (Fig. 7). Outline 2 shows that the ash layer in lidar observations persisted over Keflavik throughout the simu-lated 24-h period. The VDR plots for Keflavik highlighted the high depolarisation ratio of the ash layers, which meant that it was possible to discriminate between ash and background aerosols. Outline 5 indicates that, at the ship location, the higher ash plume descended to connect with the top of the boundary layer by 15:00 UTC on 10/06/2022. The setup of the NAME model simulations used to generate these simulated observations is provided in Supplementary Material (Table S4)

clouds based on insight from global volcanic activity was provided by the BGS and discussed.

Discussion then focussed on a close-up comparison of the different dispersion model outputs to observations and the discrepancy between forecast and satellite-retrieved mass loadings in the atmosphere. There was general agreement that differences were caused by uncertainty, and hence variation, in the plume heights, MER, and PSDs used to initialize the models. The limitations of the satellite retrievals were also noted, specifically challenges in retrieving optically thick ash clouds and sensitivity to particle size.

Given the meteorology on the day, the ash cloud was transported to the north. At the Met Office, it was found that using the agreed Plate Carrée projection (for the generation of comparable plots) made it very difficult to view and hence interpret the output and a decision was made to generate additional plots using polar stereo projection instead. This, though, made it harder to compare to simulations generated by NCAS, again highlighting the challenges when comparing model output generated by different centres in real-time response. We agreed to explore the option to generate additional output in polar projections from all partners for future events.

Following the Science Advisory Meeting, the atmospheric dispersion scientists at the Met Office finalised the decision on the most appropriate ESPs. It was agreed that forecasts should use a DFAF of 2% and this was communicated to the London VAAC duty meteorologist in time for the setup of the 18:00 UTC forecast.

Fig. 9 Exercise Case Study 3: simulated satellite imagery, **a** dust RGB and **b** retrieved total column mass loadings (g m^{-2}) at 10:00 UTC 10/06/2022, generated in real time during the exercise. The red contour indicates all total column mass loadings greater than 10 g m^{-2}. Grey areas within the ash cloud indicate locations where the satellite is unable to retrieve data due to high mass loadings
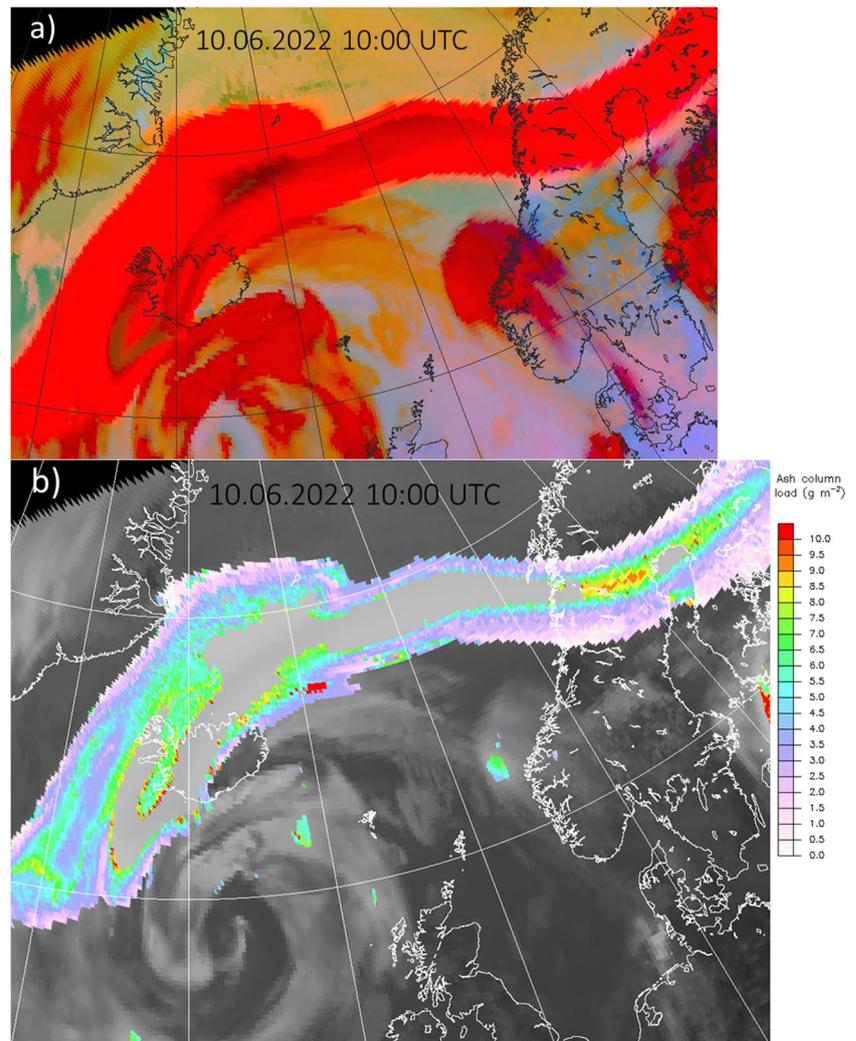

Discussion

Exercises are crucial for ensuring that we are prepared to respond effectively during a crisis. Often, these focus on practising procedure, process, and lines of communication. However, we have developed a new type of exercise for the London VAAC, called SCI-VOLCICE, which tests forecast interpretation and evaluation, with pull-through of scientific advice in real time to support operational response. Here, we discuss the key outcomes, identify the lessons learnt, and consider the implications for future development of operational VAAC forecasts.

Outcomes

Our SCI-VOLCICE exercises have provided a useful training opportunity and enabled us to improve our understanding of both model output and observations and how and why they may differ. The use of simulated observations in our exercises also allowed us to explore our decision-making processes and how we might make changes to our NAME model setup to generate an optimal forecast during a real event.

Forecast evaluation requires an awareness of the sensitivity of model simulations to the parameters used as inputs and an understanding of the possible factors, assumptions, and uncertainties that may lead to discrepancies between model outputs. Through our exercises, we have practised comparing dispersion model outputs generated using different approaches (variations in model setups, model physics, and driving meteorological data), effectively generating a multi-model ensemble. This allows us to assess how different models and modelling choices impact the forecast and its agreement with observations. Knowledge gained from this assessment can then be used to inform the modelling approach applied by the London VAAC, aiding forecasting decisions in near real time during a volcanic ash cloud event. It should be remembered though that external collaborators are not operationally required to generate model simulations;
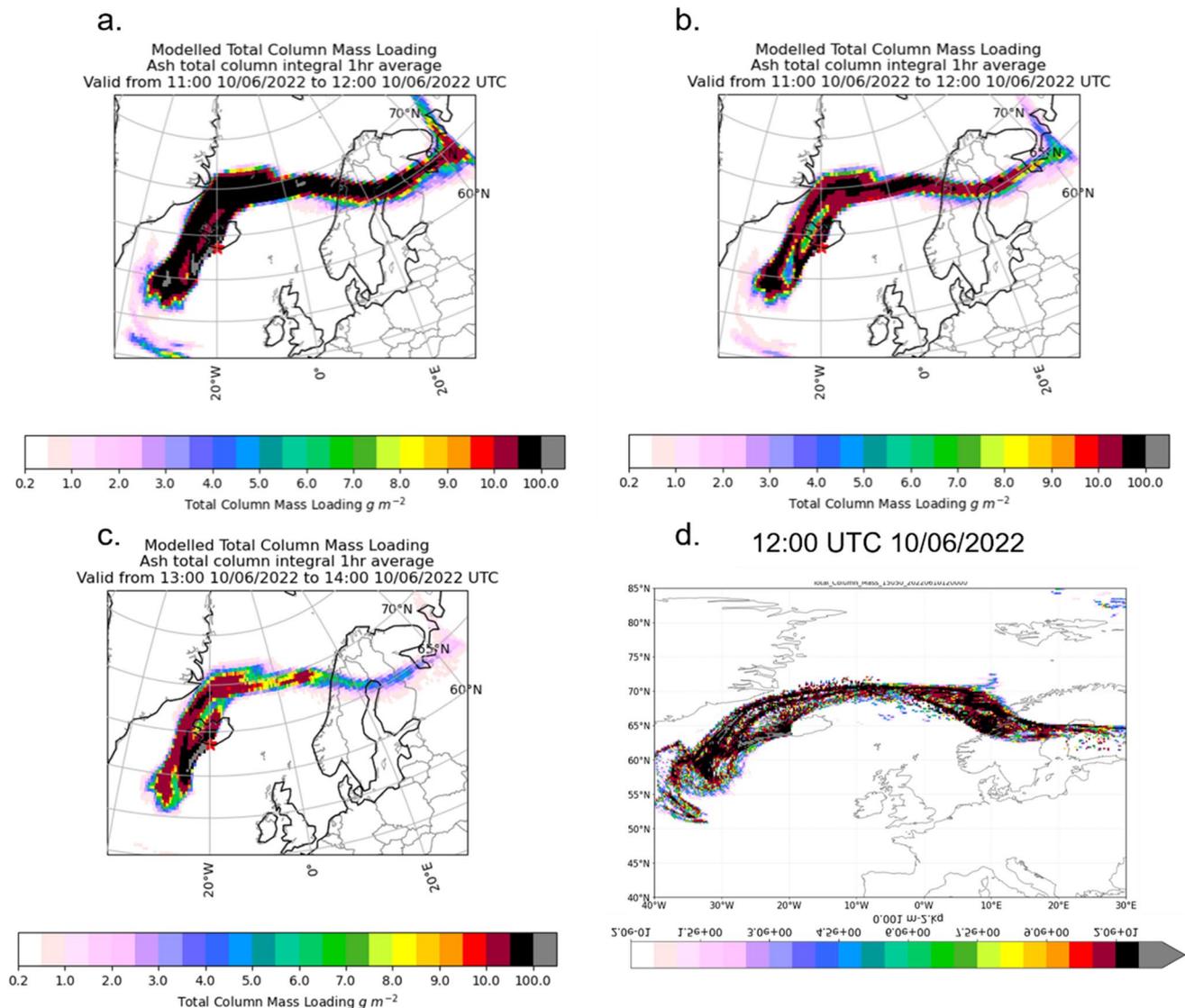
Fig. 10 Exercise Case Study 3: modelled total column mass loadings generated in real time during the exercise **a** using NAME with a DFAF of 5% (London VAAC) and additional NAME output using **b** 2% DFAF and **c** a coarse PSD; **d** the modelled total column mass loadings using HYSPLIT with GFS (NCAS). The validity times are given by the sub-title of each plot. The grey contour indicates all total column mass loadings greater than 100 g m.$^{-2}$

they work on a best-endeavour basis, and there is no formal expectation that output will be produced.

The exercises have strengthened our relationship with external collaborators, with the aim of providing the best possible scientific advice to the London VAAC. They provide an opportunity to share and develop specialist knowledge, ensure we are familiar with each other's areas of expertise and roles, and have set in place clear, formal, lines of communication. The exercises have clarified the interactions which need to take place, to ensure that scientific expertise can be pulled into the London VAAC forecasting process. Practising these interactions and the 'real-time' nature of the

exercises also improved our ability to make time-constrained decisions, under the pressure of these working conditions.

Lessons learnt

We have learnt several lessons from our SCI-VOLCICE exercises and been able to identify weakness in knowledge and understanding, all of which are pertinent for the future development of emergency response systems for all operational centres:

Lesson 1. To allow a meaningful comparison between forecasts generated using different modelling approaches, model setups must be shared. Our exercises have identified

Table 4 Exercise Case Study 3: advice provided by the BGS on the history of eruptions at Eyjafjallajökull and possible ESPs and eruption scenarios for this event, using a variety of sources including the Catalogue of Icelandic Volcanoes, Höskuldsson et al. 2018 and Bonadonna et al. 2011

Eyjafjallajökull is a central shield-like volcano which is partially covered by ice. Eruptions tend to be located at the summit area or on the volcano flanks. Characteristic eruptive activity includes both explosive and effusive eruptions. The presence of ice over the volcano means that explosive eruptions can produce large amounts fine ash. Eruptions are commonly accompanied by jökulhlaups. Most likely, plume heights based on previous activity range from 3 to 12 km asl, with a reasonable worst-case plume height of 25 km asl
Eyjafjallajökull most recently erupted in 2010, with the eruption beginning with a small eruption on the flank, followed by a moderate eruption from the summit. The 2010 eruption lasted for 2.5 months and produced plumes that reached between 5 and 10 km above vent. The total ash fraction (< 2 mm) emitted during this eruption amounted to 85%, and the fine ash fraction (< 63 microns) to 26% of the emitted ash, respectively
Scenarios for evolution of current eruptive activity include the following:
• Continued eruption at current levels over the coming hours to days
• A reduction in plume height to those observed during previous eruptions (3–12 km) with the potential for the eruption to continue at this level for weeks to months
• The eruption stops

Table 5 Exercise Case Study 3: model setups used by the collaborators. All parameters are the same as given in Table 2 except for those listed here

Partner	Met Office	NCAS
Plume height(s)	15 km	15 km
Mass eruption rate(s)	5% Mastin (London VAAC issued forecast) 2% Mastin (additional science support run)	3% Mastin
Particle size range (see Supplementary Table S1 for PSDs)	0.1–100 µm default PSD (London VAAC issued forecast) 0.1–125 µm coarse PSD (additional science support run)	0.6, 2, 6, 20 µm

that information on key parameters should be prepared and distributed during event response (Table 1). We recommend that any institution producing model simulations of the transport and dispersion of ash clouds in an event should share this information to enable robust forecast evaluation by responding agencies.

Lesson 2. Generating comparable plots when assessing outputs from different modelling systems used by different institutes is important for forecast evaluation. Although this may seem obvious, comparable plots are surprisingly hard to achieve. Following the outcomes of our exercises, we have identified and defined key plotting requirements to ensure comparable plots are generated by our external collaborators. However, our most recent exercise still highlighted challenges in achieving this. Setting the domain extent is difficult as this cannot be pre-defined, and the projection may need to be flexible depending on the direction of travel of the plume; transport over the poles may be better represented by a different projection to an ash cloud transported over the UK, Scandinavia, and Europe. Furthermore, development of methodologies to enable rigorous quantitative assessment of different model outputs would be beneficial, but would require data to be shared in real time; technical issues (e.g. data grids, formats, and sharing) would need to be addressed and suitable statistical methods identified and implemented.

Lesson 3. Practising information exchange and decision-making in a time-constrained environment is very important. Initially, we struggled to keep our discussions to time and be decisive. The exercises enabled us to consider which information is important for our forecast evaluation and Science Advisory Meetings and how we might present it. Through our exercises, we have developed clear agendas for these meetings which ensure that the key information is presented and the relevant discussions take place, within the required timeframe.

Outlook

The International Civil Aviation Organisation (ICAO) has directed that all VAACs should now be working towards providing quantitative, probabilistic, ash concentration forecasts (ICAO 2021). The results from our exercises highlight the benefits that these will bring; model simulations are sensitive to both the ESPs and meteorological data used and this uncertainty should be represented in the communication of the hazard to the aviation industry. Our results also suggest that expert knowledge will be important for assigning uncertainties to ESPs and interpreting the probabilistic forecasts generated.

The ICAO has outlined an agreed set of standards that future probabilistic datasets, generated by the VAACs, must

adhere to. When developing their forecast products, the VAACs, and indeed any centre generating model simulations during an ash cloud event, may want to consider also developing comparable graphical forecasts using agreed plotting standards. This would support forecast evaluation, particularly when an ash cloud travels through several areas of responsibility.

Data fusion algorithms that incorporate observations into model simulations, such as source inversion (Webster and Thomson 2022) or data assimilation (Mingari et al. 2022), can add considerable value for refining ESPs and improving forecasts. To date, we have not included the use of our inversion tool (InTEM) in our exercises. This is a clear next step, and we are now developing methodologies which will enable us to achieve this with simulated satellite observations.

Our operational response would benefit from better tools to evaluate forecasts, for example, software which allows direct comparison of model output and observations quickly and in near real time, e.g. statistical methods such as the structure-amplitude-location score (Wilkins et al. 2016) or tools for overlaying different data types.

When using observations for forecast evaluation, it is also important to consider associated errors and uncertainties, which can be due to limitations and assumptions in measurement techniques and retrieval methods (e.g. Ansmann et al. 2011; Prata and Prata et al. 2012; Stevenson et al. 2015; Wen and Rose 1994; Western et al. 2015). Different observation types (satellite, lidar, or radar) can detect different aspects of the ash cloud, e.g. different geographical coverage and/or ash particle size ranges. We need to continue to improve our understanding of where known discrepancies might lie between observational datasets and model simulations. Improving awareness of this across all responders will also improve our ability to generate the best possible forecast.

Operational volcanic ash cloud forecasts benefit from the support of scientific experts at volcano observatories, national/state geological or geophysical institutes, and volcano research institutions (Bonadonna et al. 2011). We should continue to strive to support these collaborations and build these necessary relationships (Witham et al. 2023, personal communication).

Conclusions

We have developed exercises, called SCI-VOLCICE, which focus on testing a multi-agency response to provide scientific support to the London VAAC for a volcanic eruption in Iceland. Our exercises have highlighted the importance of practising forecast evaluation procedures, scientific interpretation of model output and observations, and the pull-through of scientific advice into the London VAAC. These exercises are particularly important as events within our area of responsibility are infrequent.

We have developed new methodologies for generating and using simulated satellite and lidar retrievals. These proved very beneficial, as they allowed us to practise our interpretation of both model output and observation data under real-time conditions. We have also practised comparing London VAAC forecasts to model simulations generated by other institutes. We have shown that this multi-model assessment enables us to explore drivers of forecast variability. The variation in model setups (including ESPs used), model physics, and meteorological data used is advantageous, as it allows responders to consider how different modelling approaches may impact the predicted extent and concentration of ash during an event. This supports the recommendation by the ICAO that the VAACs should be developing and issuing probabilistic forecasts.

Finally, our exercises have reinforced that collaboration between experts from different institutions, who have varying roles and skillsets, is key to generating the best possible forecast. Practising a joint response to an eruption in real time enabled us to better understand each other's roles and the information that each could contribute during an event. This experience improved our group understanding of the observations being used and model forecasts generated and enabled us to improve our ability to scientifically interpret volcanic ash cloud observations and forecasts under time-limited conditions.

Acronyms BGS: British Geological Survey; CAA: Civil Aviation Authority; COBR: Cabinet Office Briefing Room (UK Government); DFAF: Distal fine ash fraction; ECMWF: European Centre for Medium-Range Weather Forecasts; ESP: Eruption source parameter; EUROCONTROL: European Organisation for the Safety of Air Navigation; FALL3D: A 3-D time-dependent Eulerian model for the transport of tephra; GFS: Global Forecasting System, National Oceanic and Atmospheric Administration (United States) NWP model; HYSPLIT: Hybrid Single-Particle Lagrangian Integrated Trajectory model; ICAO: International Civil Aviation Organisation; IFS: Integrated Forecasting System, the ECMWFs NWP; IMO: Icelandic Met Office; ISAVIA: Icelandic air navigation service provider; MER: Mass eruption rate; MSG: Meteosat Second Generation satellite system; NAME: Numerical Atmospheric-dispersion Modelling Environment; NCAS: National Centre for Atmospheric Science; NWP: Numerical weather prediction; PSD: Particle size distribution; REFIR: Real-time Eruption source parameters FutureVolc Information and Reconnaissance system; RGB: Red-green–blue composite dust satellite image; RTTOV: Radiative transfer model used at the Met Office; SAGE: Scientific Advisory Group for Emergencies (UK Government); SEVIRI: Spinning Enhanced Visible and Infrared Imager; TGSD: Total grain size distribution; UM: Unified Model, the Met Office's NWP model; VAA: Volcanic Ash Advisory; VOLCICE: Monthly exercises between the London VAAC, ISAVIA, and the IMO; VOLCEX: Pan-European exercises which practise multi-agency response

References

Adam M, Buxmann J, Freeman N, Horseman A, Salmon C, Sugier J, Bennett R (2018) The UK Lidar-sunphotometer operational volcanic ash monitoring network. EPJ Web of Conferences 176:09006

Ansmann A, Tesche M, Seifert P, Groß S, Freudenthaler V, Apituley A, Wilson KM, Serikov I, Linné H, Heinold B, Hiebsch A, Schnell F, Schmidt J, Mattis I, Wandinger U, Wiegner M (2011) Ash and fine-mode particle mass profiles from EARLINET- AERONET observations over central Europe after the eruptions of the Eyjafjallajökull volcano in 2010'. J Geophys Res 116:D00U02. https://doi.org/10.1029/2010JD015567

Beckett FM, Witham CS, Hort MC, Stevenson JA, Bonadonna C, Millington SC (2015) Sensitivity of dispersion model forecasts of volcanic ash clouds to the physical characteristics of the particles Abstract Key Points. J Geophys Res Atmos 120(22). https://doi.org/10.1002/2015JD023609

Beckett FM, Witham CS, Leadbetter SJ, Crocker R, Webster HN, Hort MC, Jones AR, Devenish BJ, Thomson DJ (2020) Atmospheric dispersion modelling at the London VAAC: a review of developments since the 2010 Eyjafjallajökull volcano ash cloud. Atmosphere 11:352. https://doi.org/10.3390/atmos11040352

Bonadonna C, Genco R, Gouhier M, Pistolesi M, Cioni R, Alfano F, Hoskuldsson A, Ripepe M (2011) Tephra sedimentation during the 2010 Eyjafjallajökull eruption (Iceland) from deposit, radar, and satellite observations. J Geophys Res 116:B12202. https://doi.org/10.1029/2011JB008462

Bonadonna C, Folch A, Loughlin S, Puempel H (2012) Future developments in modelling and monitoring of volcanic ash clouds:

outcomes from the first IAVCEI-WMO workshop on Ash Dispersal Forecast and Civil Aviation. Bull Volcanol 74:1–10

Casadevall TJ, Delos Reyes P, Schneider D (1996) The 1991 Pinatubo eruptions and their effects on aircraft operations. In Fire and mud: eruptions and lahars of Mount Pinatubo, Philippines; Newhall, C., Punongbayan, R., Eds.; University of Washington Press: Seattle, WC, USA, 625– 636.

Clarkson R, Majewicz E, Mack P (2016) A re-evaluation of the 2010 quantitative understanding of the effects volcanic ash has on gas turbine engines. J Aerosp Eng 230:2274–2291

Clarkson R, Simpson H (2017) Maximising airspace use during volcanic eruptions: matching engine durability against ash cloud occurrence. In Proceedings of the NATO STO AVT-272 Specialists Meeting on: Impact of Volcanic Ash Clouds on Military Operations, Vilnius, Lithuania.

Devenish B (2013) Using simple plume models to refine the source mass flux of volcanic eruptions according to atmospheric conditions. J Volc Geotherm Res 326:114–119

Dioguardi F, Beckett FM, Dürig T, Stevenson JA (2020) The impact of eruption source parameter uncertainties on ash dispersion forecasts during explosive volcanic eruptions. JGR: Atmos 125:e2020JD032717. https://doi.org/10.1029/2020JD032717

Donovan A (2021) Experts in emergencies: a framework for understanding scientific advice in crisis contexts. Int J Disaster Risk Reduct 56:102064. https://doi.org/10.1016/j.ijdrr.2021.102064

Dürig T, Gudmundsson M, Dioguardi F, Woodhouse M, Björnsson H, Barsotti S, Witt T, Walter T (2018) REFIR- a multi-parameter system for near real-time estimates of plume-height and mass eruption rate during explosive eruptions. J Volcanol Geotherm Res 360:61–83

Folch A, Mingari L, Gutierrez N, Hanzich M, Macedonio G, Costa A (2020) FALL3D-8.0: a computational model for atmospheric transport and deposition of particles, aerosols and radionuclides – part 1: model physics and numerics. Geosci Model Dev 13:1431–1458. https://doi.org/10.5194/gmd-13-1431-2020

Francis PN, Cooke MC, Saunders RW (2012) Retrieval of physical properties of volcanic ash using Meteosat: a case study from the 2010 Eyjafjallajökull eruption. J Geophys Res Atmos 117:D00U09. https://doi.org/10.1029/2011JD016788

Harris AJ, Gurioli L, Hughes EE, Lagreulet S (2012) Impact of the Eyjafjallajökull ash cloud: a newspaper perspective. J. Geophys. Res., 117, B00C08

Höskuldsson A, Janebo MH, Thordarson T, Andrésdóttir TB, Jónsdóttir I, Guðnason J, Schmith J, Moreland W, Magnúsdóttir AÖ (2018) Total grain size distribution in selected Icelandic eruptions, Work for the International Civil Aviation Organization (ICAO) and the Icelandic Meteorological Office, https://icelandicvolcanoes.is/data/Gosva/Total_Grain_Size_Distribution_in_Selected_Icelandic_Eruptions_01.pdf, Accessed 27/10/2023.

IMO dispersion modelling. https://dispersion.vedur.is/#. Accessed 11 December 2020.

ICAO (2016) International Civil Aviation Organization Volcanic ash contingency plan: European and North Atlantic Regions; EUR doc 019, NAT doc 006; Part II; International Civil Aviation Authority, Montreal, QC, Canada

ICAO (2020), International Civil Aviation Organisation Handbook on the International Airways Volcano Watch (IAVW), Doc9766-AN/968. https://www.icao.int/airnavigation/METP/MOG%20IAVW%20Reference%20Documents/Handbook%20on%20the%20IAVW,%20Doc%209766.pdf. Accessed 18 Sept 2023

ICAO (2021) International Civil Aviation Organization Meteorology Panel Roadmap for International Airways Volcano Watch (IAVW) in support of international air navigation. International Civil Aviation Organization. https://www.icao.int/airnavigation/METP/MOG%20IAVW%20Reference%20Documents/IAVW%

20Roadmap%20vs%205.0%20-%2001nov2021.pdf. Accessed 27 Nov 2023

Jones AR, Thomson DJ, Hort MC, Devenish B (2007) The U.K. Met Office's next-generation atmospheric dispersion model, NAME III, in Borrego C. and Norman A.-L. (Eds) Air Pollution Modelling and its Application XVII (Proceedings of the 27th NATO/CCMS International Technical Meeting on Air Pollution Modelling and its Application), Springer. 580–589.

Klett JD (1985) Lidar inversion with variable backscatter/extinction ratios. Appl Opt 24:1638–1643

Lechner P, Tupper A, Guffanti M, Loughlin S, Casadevall TJ (2017) Volcanic ash and aviation—the challenges of real-time, global communication of a natural hazard, In: Observing the volcano world: volcano crisis communication, Eds, Fearnley CJ, Bird DK, Haynes K, McGuire WJ, Jolly G.

London VAAC concentration chart webpages. https://www.metoffice.gov.uk/services/transport/aviation/regulated/vaac/advisories. Accessed 19 October 2023.

London VAAC advisory and graphics webpages. https://www.metoffice.gov.uk/services/transport/aviation/regulated/vaac/advisories. Accessed 19 October 2023.

Mastin L, Guffanti M, Servranckx R, Webley P, Barsotti S, Dean K, Durant A, Ewert J, Neri A et al (2009) A multidisciplinary effort to assign realistic source parameters to models of volcanic ash-cloud transport and dispersion during eruptions. J Volcanol Geotherm Res 186:10–21

Millington SC, Saunders RW, Francis PN, Webster HN (2012) Simulated volcanic ash imagery: a method to compare NAME ash concentration forecasts with SEVIRI imagery for the Eyjafjallajökull eruption in 2010. J Geophys Res 117:D00U17. https://doi.org/10.1029/2011JD016770

Mingari L, Folch A, Prata AT, Pardini F, Macedonio G, Costa A (2022) Data assimilation of volcanic aerosol observations using FALL3D+PDAF. Atmos Chem Phys 22:1773–1792. https://doi.org/10.5194/acp-22-1773-2022

Osborne M, Malavelle FF, Adams M, Buxmann J, Sugier J, Marenco F, Haywood J (2019) Saharan dust and biomass burning aerosols during ex-hurricane Ophelia: observations from the new UK lidar and sun-photometer network. Atmos Chem Phys 19:3557–3578

Osborne M, de Leeuw J, Witham C, Schmidt A, Beckett F, Kristiansen N, Buxmann J, Saint C, Welton EJ, Fochesatto J, Gomes AR, Bundke U, Petzold A, Marenco F, Haywood J (2022) The 2019 Raikoke volcanic eruption - part 2: particle-phase dispersion and concurrent wildfire smoke emissions. Atmos Chem Phys 22:2975–2997. https://doi.org/10.5194/acp-22-2975-2022

Osman S, Beckett FM, Rust AC, Snee E (2020) Sensitivity of volcanic ash dispersion modelling to input grain size distribution based on hydromagmatic and magmatic deposits. Atmos 11:567. https://doi.org/10.3390/atmos11060567

Pelley RE, Thomson DJ, Webster HN, Cooke MC, Manning AJ, Witham CS, Hort MC (2021) A near-real-time method for estimating volcanic ash emissions using satellite retrievals. Atmosphere 12:1573. https://doi.org/10.3390/atmos12121573

Plu M, Scherllin-Pirscher B, Arnold Arias D, Baro R, Bigeard G, Bugliaro L, Carvalho A, El Amraoui L, Eschbacher K, Hirtl M, Maurer C, Mulder MD, Piontek D, Robertson L, Rokitansky CH, Zobl F, Zopp R (2021) An ensemble of state-of-the-art ash dispersion models: towards probabilistic forecasts to increase the resilience of air traffic against volcanic eruptions. Nat Hazards Earth Syst Sci 21:2973–2992. https://doi.org/10.5194/nhess-21-2973-2021

Prata AJ, Prata AT (2012) Eyjafjallajökull volcanic ash concentrations determined using spin enhanced visible and infrared imager measurements'. J. Geophys Res 117:D00U23. https://doi.org/10.1029/2011JD016800

Prata F, Lynch M (2019) Passive earth observations of volcanic clouds in the atmosphere. Atmosphere 10:199

Saxby J, Beckett FM, Cashman KV, Rust AC, Tennant E (2018) The impact of particle shape on fall velocity: implications for volcanic ash dispersion modelling. J Volcanol Geoth Res 362:32–48. https://doi.org/10.1016/j.jvolgeores.2018.08.006

Saxby J, Rust AC, Beckett FM, Cashman KV (2020) Rodger H (2020) Estimating the 3D shape of volcanic ash to better understand sedimentation processes and improve atmospheric dispersion modelling. Earth Planet Sci Lett 534:116075

Scollo S, Folch A, Costa A (2008) A parametric and comparative study of different tephra fall out models. J Volcanol Geotherm Res 176:199–211

Stein A, Draxler RR, Rolph G, Stunder B, Cohen MD, Ngan F (2015) NOAA's HYSPLIT atmospheric transport and dispersion modelling system. Bull Am Meteor Soc 96:150504130527006. https://doi.org/10.1175/BAMS-D-14-00110.1

Western LM, Watson MI, Francis PN (2015) Uncertainty in two-channel infrared remote sensing retrievals of a well-characterised volcanic ash cloud. Bull Volcanol 77:67. https://doi.org/10.1007/s00445-015-0950-y

Stevenson JA, Millington SC, Beckett FM, Swindles GT, Thordarson T (2015) Big grains go far: understanding the discrepancy between tephrochronology and satellite infrared measurements of volcanic ash' Atmospheric. Meas Tech 8:2069–2091

Stevenson JA, Valters DA, Dioguardi, F (2023). Ash model plotting: a Python toolset for analysing results from volcanic ash dispersion model runs (v1.0.0). Zenodo. 10.5281/zenodo.7785355.

Walters D, Baran A, Boutle I, Brooks M, Earnshaw P, Edwards J, Furtado K, Hill P, Lock A, Manners J et al (2019) The Met Office Unified Model global atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations. Geosci Model Dev 12:1909–1963

Webster HN, Thomson DJ, Johnson BT, Heard IPC, Turnbull K, Marenco F, Kristiansen NI, Dorsey J, Minikin A, Weinzierl B, Schumann U, Sparks RSJ, Loughlin SC, Hort MC, Leadbetter SJ, Devenish BJ, Manning AJ, Witham CS, Haywood JM, Golding BW (2012) Operational prediction of ash concentrations in the distal volcanic cloud from the 2010 Eyjafjallajökull eruption. J Geophys Res. https://doi.org/10.1029/2011JD016790,117,D00U08

Webster HN and Thomson DJ (2022) Using ensemble meteorological datasets to treat meteorological uncertainties in a Bayesian volcanic ash inverse modelling system: a case study, Grímsvötn 2011, J Geophys Res Atmo. https://doi.org/10.1029/2022JD036469

Wen S, Rose WI (1994) Retrieval of sizes and total masses of particles in volcanic clouds using AVHRR bands 4 and 5. J Geophys Res 99(D3):5421–5431

Witham C, Barsotti S, Dumont S, Oddsson B, Sigmundsson F (2020) Practising an explosive eruption in Iceland: outcomes from a European exercise. J Appl Volcanol 9:1. https://doi.org/10.1186/s13617-019-0091-7

Witham CS, Hort MC, Potts R, Servranckx R, Husson P, Bonnardot F (2007) Comparison of VAAC atmospheric dispersion models using the 1 November 2004 Grimsvötn eruption. Meteorol Appl 14:27–38

Wilkins KL, Watson IM, Kristiansen NI, Webster HN, Thomson DJ, Dacre HF, Prata AJ (2016) Using data insertion with the NAME model to simulate the 8 May 2010 Eyjafjallajökull volcanic ash cloud. J Geophys Res-Atmos 121:306–323. https://doi.org/10.1002/2015JD023895