Operational prediction of ash concentrations in the distal volcanic cloud from the 2010 Eyjafjallajökull eruption

H. N. Webster,[1] D. J. Thomson,[1] B. T. Johnson,[1] I. P. C. Heard,[1] K. Turnbull,[1] F. Marenco,[1] N. I. Kristiansen,[2] J. Dorsey,[3] A. Minikin,[4] B. Weinzierl,[4] U. Schumann,[4] R. S. J. Sparks,[5] S. C. Loughlin,[6] M. C. Hort,[1] S. J. Leadbetter,[1] B. J. Devenish,[1] A. J. Manning,[1] C. S. Witham,[1] J. M. Haywood,[1] and B. W. Golding[1]

Received 26 August 2011; revised 14 November 2011; accepted 16 November 2011; published 27 January 2012.

[1] During the 2010 eruption of Eyjafjallajökull, improvements were made to the modeling procedure at the Met Office, UK, enabling peak ash concentrations within the volcanic cloud to be estimated. In this paper we describe the ash concentration forecasting method, its rationale and how it evolved over time in response to new information and user requirements. The change from solely forecasting regions of ash to also estimating peak ash concentrations required consideration of volcanic ash emission rates, the fraction of ash surviving near-source fall-out, and the relationship between predicted mean and local peak ash concentrations unresolved by the model. To validate the modeling procedure, predicted peak ash concentrations are compared against observations obtained by ground-based and research aircraft instrumentation. This comparison between modeled and observed peak concentrations highlights the many sources of error and the uncertainties involved. Despite the challenges of predicting ash concentrations, the ash forecasting method employed here is found to give useful guidance on likely ash concentrations. Predicted peak ash concentrations lie within about one and a half orders of magnitude of the observed peak concentrations. A significant improvement in the agreement between modeled and observed values is seen if a buffer zone, accounting for positional errors in the predicted ash cloud, is used. Sensitivity of the predicted ash concentrations to the source properties (e.g., the plume height and the vertical distribution of ash at the source) is assessed and in some cases, seemingly minor uncertainties in the source specification have a large effect on predicted ash concentrations.

Citation: Webster, H. N., et al. (2012), Operational prediction of ash concentrations in the distal volcanic cloud from the 2010 Eyjafjallajökull eruption, *J. Geophys. Res.*, *117*, D00U08, doi:10.1029/2011JD016790.

1. Introduction

[2] The danger posed to aviation from volcanic ash is well known and has resulted in loss of power to all engines in the most serious aircraft encounters, most notably in aircraft encounters with volcanic ash clouds from Galunggung in 1982 [*Hanstrum and Watson*, 1983] and from Mount Redoubt in 1989 [*Casadevall*, 1994]. The need to be able to warn the aviation community of volcanic activity and to forecast the atmospheric transport of the volcanic ash cloud was recognized in the establishment of the Volcanic Ash Advisory Centres (VAACs) in the 1990s by the International Civil Aviation Organization (ICAO).

[3] During the eruption of the Icelandic volcano Eyjafjallajökull (63.63°N, 19.62°W) in 2010, the airline industry suffered major disruption as restrictions on European airspace were enforced over a prolonged period. These restrictions were initially based on the accepted guidance that volcanic ash should be completely avoided by aircraft [*International Civil Aviation Organization*, 2002]. During this time the London VAAC, which is part of the Met Office, UK, issued regular volcanic ash advisories forecasting the transport and dispersion of the ash cloud using the Met Office's atmospheric dispersion model, NAME (Numerical Atmospheric-dispersion Modelling Environment, *Jones et al.* [2007]). These advisories simply indicated the regions where ash was forecast to be present. However, as the eruption continued, the Met Office and the aviation industry agreed that attempts be made to forecast ash concentrations in order that a revised procedure could be introduced that allowed aircraft to fly within the predicted volcanic ash cloud in regions with low levels of ash. These predictions of ash

[1]Met Office, Exeter, UK.
[2]NILU, Norwegian Institute for Air Research, Kjeller, Norway.
[3]NCAS, University of Manchester, Manchester, UK.
[4]Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany.
[5]Department of Earth Sciences, University of Bristol, Bristol, UK.
[6]British Geological Survey, Edinburgh, UK.

0148-0227/12/2011JD016790

concentration were issued as supplementary guidance to the official VAAC product. The decision to move from predicting just ash presence to also forecasting actual ash concentrations within the volcanic ash cloud brought with it new challenges which needed to be addressed quickly.

[4] In this paper we describe the changes to the volcanic ash cloud forecasts made by the Met Office during the 2010 Eyjafjallajökull eruption enabling ash concentration predictions to be issued. Some subsequent modifications to the forecasting method are also described. The accuracy of predicted ash concentrations is assessed against both ground-based observations and measurements from instrumentation onboard research aircraft. (For comparisons of ash predictions against satellite observations, see B. J. Devenish et al. (Sensitivity analysis of dispersion modeling of volcanic ash from Eyjafjallajökull in May 2010, submitted to *Journal of Geophysical Research*, 2012) and *Millington et al.* [2012].) A variety of modeling approaches are assessed and discussed including the current operational approach.

2. NAME Modeling of Volcanic Ash

[5] NAME is the operational dispersion model of the London VAAC and was used during the 2010 eruption of Eyjafjallajökull. It was initially developed to predict the dispersion of radionuclides following the Chernobyl nuclear accident, but now has a wide range of applications. These applications include emergency response work predicting the dispersion of hazardous substances (e.g., chemical, biological and nuclear material), predicting the spread of airborne diseases, air quality forecasting, dust modeling, and source apportionment studies, in addition to volcanic ash modeling. The performance of NAME has been validated in a range of situations against a number of dispersion experiments (e.g., ETEX, Kincaid) and against other dispersion models [*Galmarini et al.*, 2010; *Potempski et al.*, 2008; *Ryall and Maryon*, 1998; *Thomson and Jones*, 2011; *Webster and Thomson*, 2002; *Witham et al.*, 2007].

[6] NAME is a Lagrangian model in which large numbers of model 'particles' are released into and tracked through the computational atmosphere. The model particles are moved by the resolved wind and by a random motion which represents the effects of unresolved motions. The random motion is independent from particle to particle and is uncorrelated from time step to time step (i.e., it is what is often referred to as a "random displacement model" or "Markov displacement model" [*Boughton et al.*, 1987] which corresponds, via the Fokker-Planck equation, to an advection–diffusion model). The mass of dispersing material is divided among the model particles and, when the material is a particulate, the mass of one model particle generally corresponds to many actual particles.

[7] Gravitational settling of heavy particles is represented in NAME with the fall velocity calculated from the particle density and diameter using the Reynolds number dependent drag coefficient for spherical particles given by *Maryon et al.* [1999] with the Cunningham correction applied for small particle sizes [*Pruppacher and Klett*, 1999]. NAME can represent a distribution of particle sizes by giving different model particles different sizes and hence different sedimentation velocities. NAME also includes parameterizations of wet and dry deposition processes (here dry deposition refers to the non-sedimentation-driven dry deposition associated with surface absorption of gases or impaction of particles on surface roughness elements and turbulent diffusion to the surface). The dry deposition parameterization uses a deposition velocity calculated by the resistance analogy and is combined with gravitational settling as described by *Webster and Thomson* [2012] using the approach of *Underwood* [1999]. For particulates, including volcanic ash, the surface resistance is set to be zero since particles encountering the surface are assumed to deposit. The wet deposition parameterization uses scavenging coefficients and accounts for rain out and wash out by precipitation as described by *Maryon et al.* [1999].

[8] In modeling volcanic ash, NAME does not model any of the complex near-source processes associated with the rising eruption column or with any umbrella cloud. This includes the near-source fall-out processes associated with large grain sizes, aggregation of grains, accretion of ice or water onto ash, and possibly coherent downdrafts. Aggregation [*Costa et al.*, 2010; *Durant et al.*, 2009; *Folch et al.*, 2010; *Sparks et al.*, 1997] is caused by collision of ash particles (because of turbulent plume motions, differences in settling velocities, electrostatic forces and possibly Brownian motion) and their ability to adhere (for example, due to the surface tension of condensed water or other liquids, ice crystal growth, mineral crystal growth associated with the interaction between ash and condensed acids, or electrostatic forces), and can result in efficient removal of ash due to the larger sedimentation velocity of the aggregates. Instead NAME uses an 'effective source' which aims to represent the net effect of the near-source processes in a way appropriate for estimating the long range dispersion. After ash is released from this effective source, it is assumed to disperse passively (in the sense that the ash does not alter the flow). The term 'distal' is used here to refer to the passive regime which can be represented by the effective source, whereas the term 'near-source' is used to refer to the regime prior to this. For operational predictions and for many of the simulations presented here, the effective source is taken to be a uniform line source from the volcano vent height to the observed plume rise height. This is unlikely to be an accurate representation of the vertical distribution of the ash source. However, it is chosen to try to avoid predicting no ash in regions where ash may be present. This approach is supported by *Webley et al.* [2009], who modeled dispersion from an eruption of Mount Spurr and found that using a source confined to the umbrella cloud region gave worse agreement with the satellite-inferred ash distribution than using either a source with a uniform distribution in the vertical or a more complex source which, however, still had some material emitted at all heights in the eruption column.

[9] Volcanic ash is released in NAME with the particle size distribution shown by the dotted line in Figure 1 which is based on measurements from explosive eruptions of Mount Redoubt, St Augustine and Mount St Helens presented by *Hobbs et al.* [1991]. This particle size distribution is intended to be representative of ash particles which survive the near-source fall-out, in keeping with the idea of using an effective source. Ash particles (either single unaggregated ash grains or aggregates) larger than 100 μm in diameter are assumed to fall out near to the source and are therefore of minimal interest here; in NAME, particles with a diameter of 100 μm

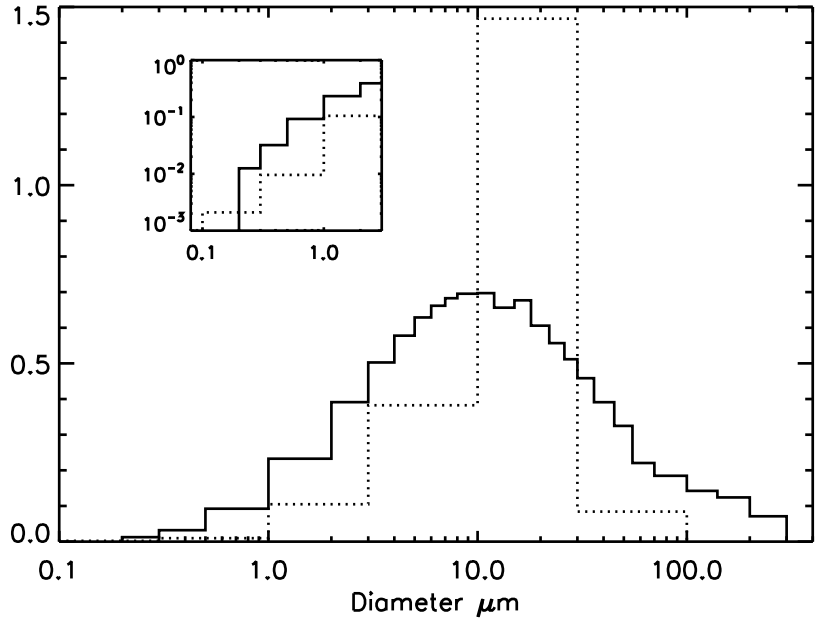

Figure 1. The default particle size distribution used in NAME for volcanic ash (dotted line), and the particle size distribution given by *Stohl et al.* [2011] for the 2010 Eyjafjallajökull eruption and used in section 4.3 (solid line). The area on the plot between any two diameters is proportional to the mass in this size range. Inset: the range 0.08 to 3.0 μm in more detail.

have fall speeds of approximately 0.55 m/s. Aggregation processes are not explicitly modeled in NAME. This is in keeping with the use of an effective source because most aggregation is thought to occur near the source where the ash particle number density is largest and there may be condensed water or ice and other chemicals. A particle density of 2300 kg/m^3 (a value toward the low end of the possible values for rock) is assumed. Although this may be too large for some aggregates, the evidence needed to refine the density estimate is not usually available, especially for real-time prediction. Similarly we have no information on how the suspended distal ash mass is divided between individual grains and aggregates of grains. One can regard the density and particle size as the effective values needed to determine the sedimentation velocity and so any errors in density can be regarded as absorbed into errors in the particle size distribution.

[10] Old volcanic ash is removed from NAME forecasts for reasons of computational efficiency which are important considerations in an operational emergency response setting. During the 2010 Eyjafjallajökull eruption, and for all model predictions presented here, ash particles were removed from the model six days after their release into the computational domain. For a wind speed of 10 m/s, this equates to a travel distance of about 5,000 km. In justification, we note that the vast majority of damaging ash encounters with aircraft have occurred within ash clouds that are less than a day or so old [*Guffanti et al.*, 2010]. Furthermore, it is envisaged that much of the ash will have been removed from the atmosphere by natural processes within 6 days and that only widely dispersed and diluted ash in the fine size range will remain airborne after this time. The decision to remove aged ash from the modeling may, however, need to be revisited if evidence suggests otherwise.

[11] For most applications, NAME is driven by meteorological data from the Met Office's weather forecast model, MetUM (the Met Office Unified Model). Meteorological data from the global configuration of the MetUM with a temporal resolution of 3 hours, a horizontal resolution of approximately 25 km in mid latitudes and a forecast out to six days was used during the 2010 Eyjafjallajökull eruption. Meteorological data is interpolated in both space and time within NAME.

2.1. Evolution of Ash Modeling During the 2010 Eyjafjallajökull Eruption

[12] Initial modeling by the London VAAC during the 2010 Eyjafjallajökull eruption predicted regions of significant ash based on nominal emission rates and the table of 'visual ash' threshold concentrations developed by the National Oceanic and Atmospheric Administration (NOAA) for the VAFTAD (Volcanic Ash Forecast Transport and Dispersion) model [*Heffter and Stunder*, 1993]. This table of threshold concentrations is described by *Dacre et al.* [2011], *Leadbetter and Hort* [2011] and *Witham et al.* [2007]. The table of threshold concentrations (in units/m^3) was designed to be used with a model releasing a nominal 1 unit of ash over a relatively short (<24 hour) eruption in order to predict the extent of visual ash. The table gives discrete thresholds which vary in factors of 10 as a function of plume rise height and vent height, with the dependence of the threshold on the plume rise height reflecting the correlation between the actual emission rate and plume rise [e.g., *Sparks et al.*, 1997, section 5.2]. The thresholds are intended for use with deep-layer averaged model concentrations (over depths of order 20,000 ft) and for use as default values, to be modified in the light of satellite evidence (B. J. B. Stunder, NOAA, personal communication, 2011). At the London VAAC they were regarded as relating to an emission rate of 1 unit / 6 hours and

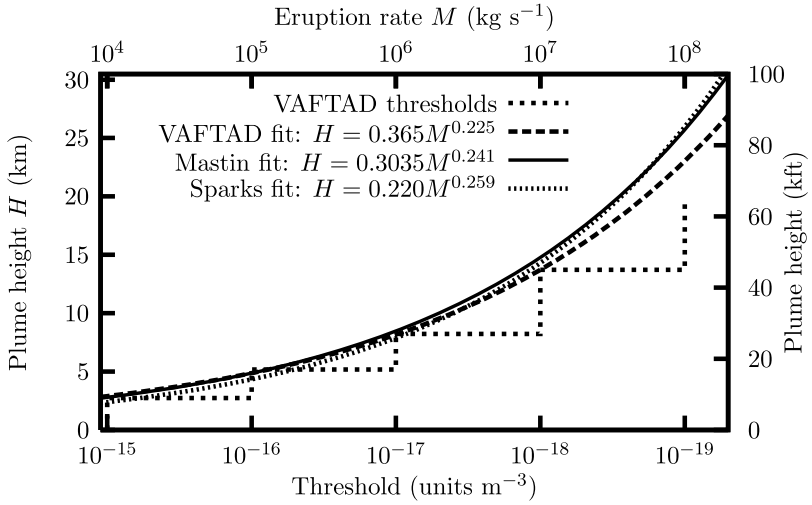

Figure 2. A comparison of estimates of eruption rate M as a function of the plume rise height H above the vent. The estimates made by *Sparks et al.* [1997, section 5.2] and by *Mastin et al.* [2009] are shown, together with the estimates obtained using the calibrated VAFTAD thresholds and using a power law fit to the calibrated thresholds. *Mastin et al.* [2009] assume a magma density of 2500 kg/m^3 in relating the volumetric flow rate to the mass emission rate. (This magma density is not related to the ash particle density used in NAME and on which the sedimentation rate depends.) The same density has been assumed for the *Sparks et al.* [1997] data. For the VAFTAD thresholds, the lower axis shows the VAFTAD 'visual ash' threshold for a 1 unit release while the upper axis shows the eruption rate. The VAFTAD values are for a vent height of 5000 feet, the midpoint of the VAFTAD range within which the Eyjafjallajökull volcano summit lies.

were used to produce forecasts based on model concentrations averaged over six hours, over horizontal grid-boxes of roughly 40 km by 40 km and over deep layers (FL000 - FL200, FL200 - FL350 and FL350 - FL550, where FL indicates the flight level which is approximately equivalent to height in hundreds of feet). Although there was no systematic use of satellite data in the way envisaged by NOAA, the model results were adjusted by forecasters on the basis of all available evidence. Within a few days of the high ash emissions starting on 14 April, the approach based on the VAFTAD table was modified so that the emission rate varied instead of the threshold in order to accommodate variations in the eruption strength. This is possible without changing the final results in periods with constant plume rise height since the concentration predicted by the model is proportional to the emission rate for an unchanged vertical distribution at the source.

[13] The decision to forecast actual ash concentrations within the ash cloud required a number of modifications to the modeling procedure. First, real (as opposed to nominal) emission rates needed to be modeled. Second, most of the emitted ash falls out near to the source and NAME only represents those ash particles which survive the near-source fall-out. Consequently, the effective source used in the NAME modeling only needed to account for a fraction of the erupted mass. Last, forecasts of ash concentrations produced with NAME are mean ash concentrations over some space-time volume, both as a result of explicit averaging (there is always some explicit averaging in a Lagrangian model) and because of other resolution limitations (e.g., due to the uniform vertical profile of the effective source, the time resolution of variations in the effective source, the resolution of the driving meteorology and the sub-grid scale parameterizations). The aim agreed around 19 April with the aviation authorities (following discussions over the previous few days) was, however, to forecast likely local peak ash concentrations. Hence, some account of regions of higher concentrations unresolved by the NAME modeling was necessary.

[14] Real emission rates were estimated from the observed plume rise height, provided by the Icelandic Meteorological Office (IMO), using an empirical relationship between these two quantities. A number of such relationships exist within the literature [e.g., *Mastin et al.*, 2009; *Sparks et al.*, 1997]. In addition the VAFTAD thresholds can be interpreted as being inversely proportional to the actual emission rate and so can be regarded as implying a particular shape for the relationship between emission rate and rise height. Initially this implied shape was adopted and calibrated using the relationship given by *Mastin et al.* [2009]. Later on during the eruption it was replaced by a smooth power law fitted through the data. Both in calibrating the VAFTAD thresholds and in fitting the power law, the VAFTAD thresholds were regarded as most reliable at the top of each discrete plume rise height range, and so only these values were used. This is because these values correspond to the most vigorous plumes within each height range and are the most critical; at the bottom of each height range just above a transition there is an extra factor of 10 safety relative to that just below the transition. Figure 2 shows the discrete and continuous relationships estimated using the NOAA VAFTAD thresholds together with the empirical relationships of *Mastin et al.* [2009] and *Sparks et al.* [1997].

[15] Note that, in these relationships, emission rate is a strong function of plume rise height, so a 20% increase in

plume rise height means a doubling of the source strength. Consequently, uncertainties in the estimated plume rise height, which can be significant at times, can translate into large errors in the calculated mass emission rate. The issue of what measure of plume rise height is the correct value to use (e.g., should one use the maximum plume-top height or the maximum height of dispersing ash, and should one use the instantaneous value or the maximum or mean over some period for a fluctuating eruption) is an interesting one. *Mastin et al.* [2009] note that, for the observations used in deriving their relationship, different measures of plume rise height are employed. In addition, for a given rise height, the emission rate will vary with the type of eruption and with the meteorology, in particular with the wind speed, stratification and humidity [*Sparks et al.*, 1997]. The size of the errors in these empirical relationships, characterized by the range within which 50% of cases fall, is estimated as plus or minus a factor of 3 to 4 [see *Mastin et al.*, 2009, Figure 1]. However, substantially larger plume rise heights for the mass emitted are possible. For example, deep moist convection, either triggered by the eruption or already active in the area, may loft a volcanic plume from low altitudes to aircraft cruising levels [*Mastin*, 2007; *Tupper et al.*, 2005, 2006]. Furthermore, emission rates from eruptions in wet environments, such as glaciers, may deviate from that suggested by the empirical plume rise height relationships due to latent heat effects [*Mastin*, 2007].

[16] The effective source strength used in NAME should reflect the fraction of the mass emission that survives the near-source fall-out processes. We call this the 'distal fine ash fraction' and the total emission rate of the volcano should be scaled by this factor for use in NAME. Estimates of the distal fine ash fraction have been made in a limited number of case studies, for example, during the recent Eyjafjallajökull eruption [*Dacre et al.*, 2011; *Devenish et al.*, 2012, also submitted manuscript, 2012] and for some previous eruptions [*Rose et al.*, 2000]. These studies give values of the distal fine ash fraction in the range between about 0.05% and 10%. There are suggestions that the fraction may tend to decrease with increasing eruption strength [*Rose et al.*, 2000] as well as in wet eruptions in which aggregation processes are more important near the source [*Sparks et al.*, 1997]. The near-source fall-out rate and the uncertainties in it are not, however, well characterized.

[17] The relationship between actual local peak concentrations and modeled mean concentrations over large volumes and time periods is also rather variable and hard to predict. Differences between means and peaks clearly depend on the averaging period and volume over which the mean concentrations are taken. Analysis of observations can give insight into the variability in ash concentrations and the expected relationship between observed mean and peak ash concentrations. Many observations show evidence of patchiness and/or thin layers of ash [*Ansmann et al.*, 2010; *Harrison et al.*, 2010; *Marenco and Hogan*, 2011; *Marenco et al.*, 2011; *Schumann et al.*, 2011; *Turnbull et al.*, 2012; B. Johnson et al., In-situ observations of volcanic ash clouds from the FAAM aircraft during the eruption of Eyjafjallajökull in 2010, submitted to *Journal of Geophysical Research*, 2012; R. J. Hogan et al., Combined lidar and sunphotometer retrievals of ash particle size and mass concentration from the Eyjafjallajökull volcano, manuscript in preparation, 2012]. For example, the airborne measurements presented by Johnson et al. (submitted manuscript, 2012) and *Schumann et al.* [2011] showed that peak ash concentrations were typically a factor of 2 or 3, and sometimes a factor of 5, greater than the mean concentrations obtained when transiting through an ash layer. Also, *Marenco and Hogan* [2011] found ash layers over the UK which were only a few hundred meters deep [see also *Dacre et al.*, 2011; *Devenish et al.*, 2012]. Furthermore, the resolution of the driving meteorology, the parameterization of sub-scale processes, the uniform representation of the source in the vertical and the time resolution of variations in the source, all affect the model's ability to resolve fully the patchy nature and fine structure of the observed ash cloud, even with little explicit averaging in the model. For example, *Devenish et al.* [2012] found that the ash cloud from Eyjafjallajökull at Exeter (50.7°N, 3.7°W) and Cardington (52.1°N, 0.4°W), UK, on 15–16 April 2010 was predicted by NAME to be about 1500 m thick, even when the model's explicit vertical averaging interval was as small as 100 m, while the ground based lidars at these locations showed a thin layer with a typical thickness of 200 to 300 m. All else being equal this will lead to NAME underpredicting the peak concentrations by a factor of 5 to 8. With output averaged over larger space-time volumes, and over deeper layers in particular, the ratio of the peak concentration to the resolved mean concentration (hereafter referred to as the peak-to-mean ratio) will be larger. For simplicity we retained the same space-time averaging volumes as used previously (about 40 km $\times$ 40 km $\times$ 20,000 ft $\times$ 6 hours) and, in order to predict likely peak ash concentrations, we multiplied the model's mean concentrations by a peak-to-mean factor (detailed below) to account for the unresolved structures.

[18] On about 17–19 April 2010, a preliminary comparison of model results with the available observations from the Eyjafjallajökull ash cloud suggested that observed local peak ash concentrations were comparable to the modeled deep layer mean concentrations before the near-source fall-out factor was applied (i.e., assuming that all the mass emitted was present in the form of particles less than 100 μm in diameter which survive near-source fall-out). In this particular case, assuming that the emission rate is approximately correct, the distal fine ash fraction must therefore be roughly equal to the reciprocal of the peak-to-mean ratio (i.e., the two factors roughly cancel one another out). This is consistent with the rough estimates of these factors made above. The assumption that these two factors cancel was used for predicting peak concentrations during the 2010 Eyjafjallajökull eruption. For definiteness we can regard this approach as assuming a 5% distal fine ash fraction and a factor of 20 peak-to-mean ratio, although of course it is only the product of these factors that affects the predictions of peak concentrations. Initially there was much uncertainty in the predictions but as more data became available confidence grew that this 'tuning' was able to predict peak concentrations within about one and a half orders of magnitude (although with the understanding that the tuning should be continually assessed against the observational evidence and adjusted if necessary, in particular to account for different eruption characteristics and improved understanding of the source). We note that tuning of the distal fine ash fraction may correct to some extent for errors

in the eruption rate or in the particle size distribution. Further comparisons between the predictions made using this approach and observations are presented in section 3.4 below. Note that although the observations used in the tests below are, of course, not independent of the observations used to increase confidence during the eruption, the majority of them are independent of the observations available when the tuning was first done.

[19] Following discussions with the aviation regulators, forecasts of peak ash concentration were produced from about 19 April. The presentation of these forecasts evolved into a format which showed three regions where peak concentrations were predicted to be in the ranges 200–2000 $\mu g/m^3$, 2000–4000 $\mu g/m^3$ and above 4000 $\mu g/m^3$. These ranges were chosen by the aviation regulators following advice from aircraft engine manufacturers and discussions of the likely uncertainties in the predicted concentrations. The way these predictions were used is beyond the scope of this paper. To enable comparisons with the previous approach in which there was no attempt to predict concentrations, we note that, with the peak ash concentration modeling approach outlined above, the 200 $\mu g/m^3$ contour is approximately equivalent to the previous ash extent prediction, at least at the top of each plume rise height range used in the VAFTAD table of threshold concentrations.

2.2. Changes to the Modeling Approach Made After the 2010 Eyjafjallajökull Eruption

[20] After the eruption of Eyjafjallajökull some further changes were made to the operational procedure for estimating peak ash concentrations. The mass emission rate estimation method was changed from the power law fit to the calibrated VAFTAD thresholds to the *Mastin et al.* [2009] formula. Also the vertical resolution of the NAME concentrations was increased to 25 flight levels (~2,500 ft) in order to resolve some aspects which were previously just treated through the peak-to-mean factor. At the same time, the peak-to-mean factor was reduced from 20 to 10 in order to account for the increased resolution. This was based on model-observation comparisons similar to those presented below and on some case studies of how the NAME ash cloud thickness varied with vertical resolution similar to the case studies presented by *Devenish et al.* [2012]. The 5% distal fine ash fraction was retained, with the expectation as before that this factor would be re-tuned if the evidence indicated that this was appropriate (as was done during the 2011 Grimsvötn eruption). The advice provided to aviation (both the official volcanic ash advisories and the supplementary peak ash concentration charts) was still (at least prior to any forecaster intervention) given over the original deep layers but was based on the maximum of the various 25FL sublayer model predictions within each deep layer. These increases in resolution require a factor of 8 increase in the number of NAME model particles to prevent an increase in statistical noise levels and hence the 25FL sublayer model has a higher computational cost. The product of the distal fine ash fraction and the peak-to-mean ratio results in an overall factor which is reduced by only a factor of two from that assumed with the scheme described in section 2.1. This suggests that the 25FL sublayer model output resolves some more structure in the ash cloud than the section 2.1 scheme does, but not a large amount more. The revised scheme also provides more detailed information on the predicted height of the ash cloud. As with the approach described in section 2.1, this approach is tested below, although here the test data are now not independent of the observations used to design the approach.

3. Comparison of Predictions With Observations

[21] In this section we compare model predictions with a range of observations. Section 3.1 summarizes the observations used, section 3.2 summarizes the model configurations, section 3.3 discusses uncertainties and section 3.4 presents the results.

3.1. Observations

[22] A large number of observations were collected during the 2010 Eyjafjallajökull eruption and a subset of these have been used in this study. The observations data set used here is summarized in Table 1 and consists of ground-based measurements obtained using lidars and sunphotometers, airborne measurements from instrumentation onboard research aircraft (the University of Manchester Cloud and Aerosol Spectrometer [CAS] and the Leosphere ALS 450 lidar, both onboard the FAAM [Facility for Airborne Atmospheric Measurements] BAe-146 aircraft, and the Forward Scattering Spectrometer Probe [FSSP] on the DLR [Deutsches Zentrum für Luft- und Raumfahrt] Falcon aircraft) and a balloon ascent (Hertfordshire University) at Stranraer (54.8°, 4.9°W). All observations in the data set are selected on the basis that there is evidence that they represent local peak ash concentrations. The data set is, however, not unbiased. The aircraft observations which dominate the data set are limited to 13 days on which aircraft measurement flights took place and to locations over the UK, Germany and the surrounding seas. Also, for the FAAM BAe-146 aircraft for example, the flight paths were planned with the intention of flying in areas predicted to contain ash while, because of safety considerations, avoiding regions where peak ash concentrations were predicted to be greater than 2000 $\mu g/m^3$. Also, the flight paths were altered if high concentrations of ash (greater than 1000 $\mu g/m^3$) or SO_2 (greater than 75 ppbv) were actually encountered. The ground-based observations are few in number and at fixed locations. The CAS measurements from the FAAM aircraft are local, i.e., not remotely sensed, observations obtained during vertical profiles (Johnson et al., submitted manuscript, 2012). The FSSP measurements from the DLR aircraft are similar [*Schumann et al.*, 2011]. The FAAM lidar measurements are, however, remotely sensed and hence more extensive [*Marenco et al.*, 2011]. The three ground lidar measurements at Exeter on 16 and 18 April 2010 and at Aberystwyth (52.4°N, 4.2°W) on 4 May 2010 and the ground lidar and sunphotometer measurement at Chilbolton on 16 April 2010 are described by *Marenco and Hogan* [2011], *Marenco et al.* [2011] and Hogan et al. (manuscript in preparation, 2012), respectively. The balloon ascent at Stranraer is described by *Harrison et al.* [2010].

[23] Determining estimates of volcanic ash concentrations from observations is far from trivial. Ash concentrations are not directly sampled but are inferred from optical measurements of the type commonly used to measure clouds and aerosols, using various assumptions about the ash particle

Table 1. The Observations Data Set Used in the Model-Observations Comparison[a]

Source	Location and Date(s)	Number of Observations	Horizontal Averaging	Vertical Averaging	Time Averaging	Reference
CAS (FAAM BAe-146 aircraft)	UK and surrounding seas 21, 22 Apr, 4, 5, 14, 16, 17, 18 May	35	1 km	50 m		Johnson et al. (submitted manuscript, 2012)
Lidar (FAAM BAe-146 aircraft)	UK and surrounding seas 4, 5, 14, 16, 17, 18 May	107	9 km	45 m		*Marenco et al.* [2011]
Balloon ascent	Stranraer, UK (54.8°N, 4.9°W) 19 Apr	1		50 m		*Harrison et al.* [2010]
Ground Lidar	Exeter, UK (50.7°N, 3.5°W) 16, 18 Apr	2		45 m	10 min	*Marenco and Hogan* [2011]
Ground Lidar	Aberystwyth, UK (52.4°N, 4.2°W) 4 May	1		100 m	10 min	*Marenco et al.* [2011]
Ground Lidar and Sunphotometer	Chilbolton, UK (51.1°N, 1.4°W) 16 Apr	1		30 m	30 s	Hogan et al. (manuscript in preparation, 2012)
FSSP (DLR Falcon aircraft)	Germany, UK and surrounding seas 19, 22, 23 Apr, 2, 9, 13, 16, 17, 18 May	21	1 km	50 m		*Schumann et al.* [2011]

[a]CAS is Cloud and Aerosol Spectrometer. FSSP is Forward Scattering Spectrometer Probe. The averaging time for the aircraft lidar data has been translated into horizontal averaging, the averaging times for the CAS and FSSP data, which were obtained during aircraft profiles, have been translated into horizontal and vertical averaging, and the averaging time for the balloon ascent has been translated into vertical averaging.

shapes and densities and about the ash refractive index. These assumed properties are based on previous mineral dust research and have proved successful in optical closure studies comparing in-situ optical properties with remote sensing and satellite retrievals [*Newman et al.*, 2012]. Lidar estimates of volcanic ash mass also rely on the particle size distribution being known from in-situ measurements and/or constrained by further remote sensing information such as the use of additional lidar wavelengths or sunphotometers.

[24] Estimated ash concentrations are obtained over various spatial and temporal scales (see Table 1) with some time averaging being required to remove noise and, for some measurement methods, increase counting statistics. The observations have been screened for cloud contamination and, in most cases, an attempt to distinguish volcanic ash from background aerosols has been made. The ash concentration measurements are thought to be, in the main, within a factor of 2 of the true value [*Marenco et al.*, 2011; *Schumann et al.*, 2011; *Turnbull et al.*, 2012; Johnson et al., submitted manuscript, 2012] and are believed to be measurements of ash which is less than six days old. The effects of aggregation and ice contamination are not assessed within the uncertainty estimate. Furthermore, any measurement reporting a low ash concentration below a certain value (e.g., 20 μg/m^3 for the CAS instrument onboard the FAAM aircraft and 10 μg/m^3 for the FSSP instrument onboard the DLR aircraft) should be regarded as 'little or no observable ash' since the reported values may reflect the background concentration of other aerosols or fluctuations associated with poor sampling statistics or signal-to-noise ratio. The CAS estimate of $\sim$5000 μg/m^3 on 14 May is suspected to be an overestimate due to possible contamination of the data by ice but is likely to be within a factor of 4 of the true value (Johnson et al., submitted manuscript, 2012). Also the mass concentration estimate from the balloon ascent relied on a compact optical particle counter with three wide particle size bins [*Harrison et al.*, 2010] leading, potentially, to larger uncertainties than a factor of two.

3.2. Model Configurations

[25] All of the NAME model runs used here are post-event reruns using analyzed global 3-hourly MetUM meteorological data (strictly analyses every 6 hours alternated with 3 hour forecasts), as opposed to the operational runs undertaken at the time of the eruption which used meteorological forecasts out to 6 days (combined with analyses).

[26] The total ash emission rate (M in kg/s) is determined using the relationship between plume rise height and emission rate given by *Mastin et al.* [2009],

$$M = 140.8H^{1/0.241}, \quad (1)$$

where H is the plume rise height above the vent in km. Emission rates determined using this relationship compare well with the smooth power law fit to the VAFTAD thresholds, as used during the 2010 Eyjafjallajökull eruption, and with the relationship given by *Sparks et al.* [1997] (Figure 2). For the range of plume rise heights observed during the 2010 Eyjafjallajökull eruption, the emission rates obtained with equation (1) and with the smooth power law fit to the VAFTAD thresholds agree to within 15% except when the plume rise is low ($H < 3$ km) and emission rates are small.

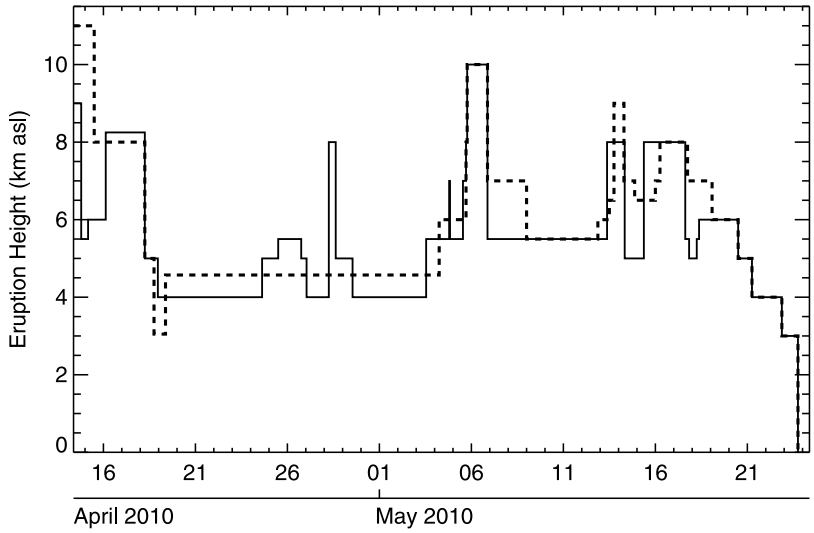

Figure 3. The plume rise height time profile (above sea level, asl) determined during the event (dashed line) and determined post-event (solid line).

[27] The plume rise height H is specified as a function of time using advice provided by the Icelandic Meteorological Office during the eruption. This advice was based on a range of data sources, of which the radar at Keflavík airport [*Arason et al.*, 2011] was the most important. Views on the most appropriate plume rise values evolved as more information became available, and here we use the last estimates that were made during the course of the April-May 2010 eruption (shown by the dashed line in Figure 3). Figure 3 shows the plume rise height above sea level; H (the height above the vent) is calculated from this using a summit height of 1666 m. There are significant uncertainties in these plume rise values as a result of the discrete scanning elevation angles of the radar (thought to result in a measurement uncertainty of 1–2 km), occasions when the plume was obscured from the radar by cloud or, for low plumes, by mountains and the curvature of the earth, occasions with missing radar scans, and, in strong winds, when the maximum plume height is reached some distance downwind. Also, the maximum observed plume rise height may not be the height at which ash is injected laterally into the atmosphere.

[28] Three different model set-ups (or schemes), which we call the deep layer, 25FL layer and hybrid schemes, have been tested. The schemes are described below. The deep layer scheme is essentially the operational scheme used at the end of the 2010 Eyjafjallajökull eruption (i.e., at the end point of the evolution presented in section 2.1) and the hybrid scheme is essentially the revised operational scheme described in section 2.2. The 25FL layer scheme has not been used operationally. All three schemes use model concentrations averaged over 6 hours and over $\sim$40 km $\times$ 40 km horizontal regions and are now described.

[29] 1. In the deep layer scheme, average model concentrations are obtained over deep layers (FL000 - FL200, FL200 - FL350 and FL350 - FL550). The distal fine ash fraction and peak-to-mean ratio given in section 2.1 are used. With the exception of the use of analyzed meteorological data, the tracking of more NAME model particles to reduce statistical noise, and the determination of the emission rate from the relationship given by *Mastin et al.* [2009] instead of from the smooth power law fit to the VAFTAD thresholds, this scheme is identical to the operational set-up in use at the end of the 2010 Eyjafjallajökull eruption as described in section 2.1.

[30] 2. In the 25FL layer scheme, average model concentrations are obtained over layers of 25FL depth and the distal fine ash fraction and peak-to-mean ratio given in section 2.2 are used. Unlike the scheme described in section 2.2, however, the 25FL layer model predictions are used directly rather than taking the maximum of these predictions over deep layers.

[31] 3. The hybrid scheme follows the 25FL layer scheme except that the peak ash concentrations are calculated over deep layers (FL000 - FL200, FL200 - FL350 and FL350 - FL550) using the maximum of the 25FL sublayer model predictions within each deep layer. With the exception of the use of analyzed meteorological data, this scheme is identical to the operational set-up described in section 2.2.

3.3. Uncertainties

[32] There are many uncertainties in the modeling of the ash concentrations and in the observations which mean that significant differences between model results and measurements should be expected. Uncertainties which have already been discussed include uncertainties (1) in estimates of the plume rise height, (2) in estimates of the mass emission rate from the plume rise height, (3) in the vertical source profile, (4) in the near-source fall-out and the distal fine ash fraction, (5) in the peak-to-mean ratio, (6) in analyzing the observations to infer the concentrations, and (7) due to the representativeness of the sample of observations.

[33] We discuss here a further significant source of differences between modeled and observed ash concentrations, namely the effect of ash cloud position errors. Even relatively small errors in the predicted ash cloud position or extent can result in significant differences between model and observation at a fixed time and place, especially for relatively narrow volcanic ash clouds. Ash cloud position

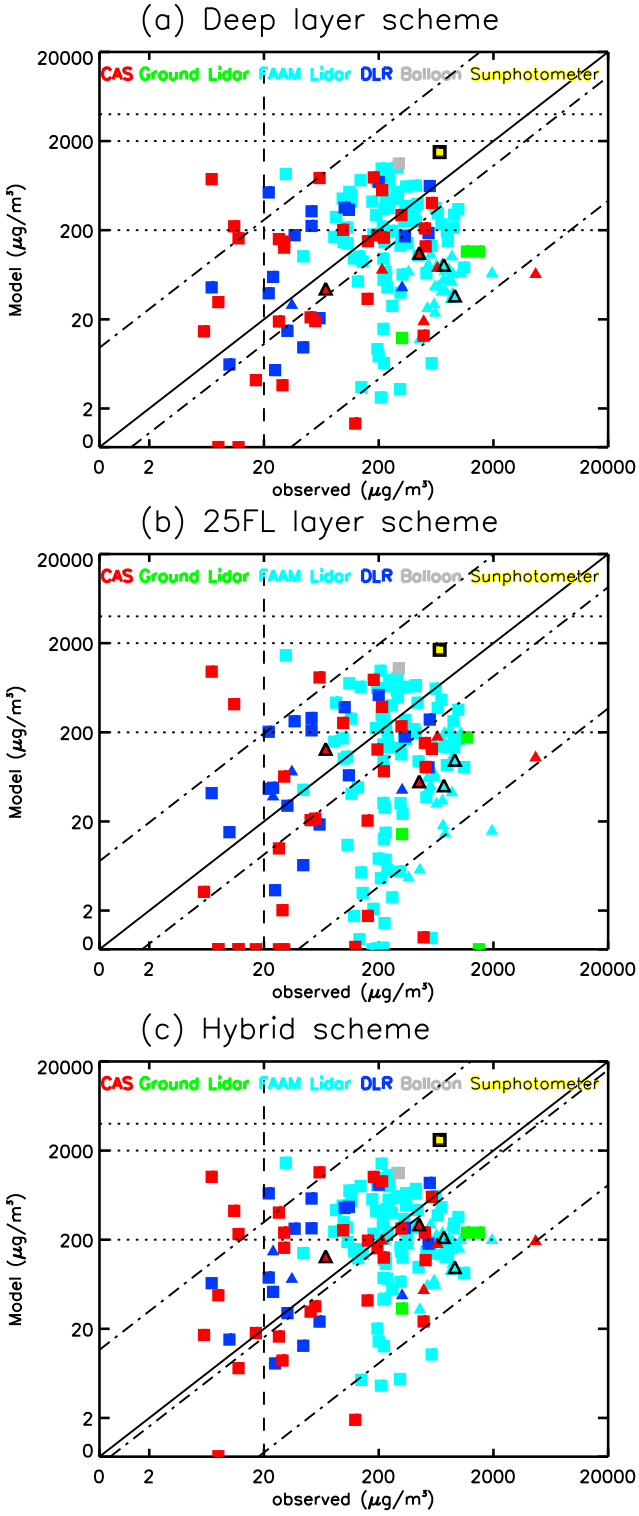

Figure 4. Modeled peak concentrations (obtained using the three model schemes) versus the peak observed concentrations described in Table 1. Squares denote observations below FL200 and triangles denote observations between FL200 and FL350. The color of the square or triangle is used here to denote the observation type: red, CAS; dark blue, DLR; light blue, FAAM lidar; green, ground lidar; grey, balloon; and yellow, sunphotometer and lidar. The squares and triangles outlined in black are discussed in section 4.1.

errors can be caused by errors in the driving meteorology, errors in the modeled dispersion or by uncertainties in the effective ash source (e.g., plume height, vertical distribution of ash). Although winds from modern Numerical Weather Prediction systems are generally well modeled (especially analyzed winds), slight errors in the driving winds can result, after a large transport distance, in significant errors in the predicted ash cloud position. Also, small errors in the ash height can amplify into large errors in predicted ash concentrations when the wind speed and/or direction varies with height.

[34] Two examples of ash cloud position errors during the 2010 Eyjafjallajökull eruption have been studied. On 15–16 April, a tendril of volcanic ash, located over the UK, was being advected mainly along its length but also with a relatively small lateral velocity. A relatively small error in the modeled ash cloud position therefore resulted in a large timing error in the predicted arrival of the ash cloud at the ground-based lidar locations [*Dacre et al.*, 2011; *Devenish et al.*, 2012]. Similarly, Devenish et al. (submitted manuscript, 2012) show NAME simulations of the ash cloud located over and to the north of Scotland on 14 May which have the cloud positioned too far to the west and not extending far enough south when compared to satellite imagery. As a result, aircraft observations intended to be at the edge of the ash cloud were probably closer to the cloud center and hence are likely to be significantly higher than the corresponding model predicted ash concentrations.

3.4. Results

[35] A comparison between predicted and observed peak ash concentrations is presented for the observations summarized in Table 1. The NAME model predictions are obtained for the three schemes described in section 3.2 and the appropriate 25FL layer or deep layer modeled value is selected based on the reported height of the observation. In addition, the appropriate ~40 km × 40 km horizontal gridbox and 6 hour time averaging period is selected so that model predictions and observations are compared for the same location, altitude and time.

[36] We note that because the height at which the model and observation are compared is the height at which the peak concentration is *observed*, this will tend, in the presence of ash cloud height errors, to cause underpredictions (since ash cloud height errors will mean that the predicted peak concentration does not occur at the altitude of the observed peak concentration). If the comparison was done between model and observation with both at the height of the *model* peak, this would tend to lead to overpredictions. These effects will be larger for the 25FL layer scheme than for the other schemes. Similar issues arise due to errors in the ash cloud's horizontal position or to timing errors.

[37] Figure 4 shows scatterplots of modeled versus observed peak ash concentrations for the three model schemes. The axes are linear below 2 $\mu g/m^3$ and logarithmic above this value (with continuous gradient at the transition) since the lower values cannot be easily distinguished on a fully linear plot whereas zero values cannot be shown on a fully logarithmic plot. The dashed vertical line at 20 $\mu g/m^3$ denotes the value below which confidence in the CAS observations is low as discussed in section 3.1. The solid diagonal line is the 1–1 line and the horizontal dotted lines denote the boundaries

Table 2. Statistical Comparison Between Model Predictions, for the Three Model Schemes, and Observations[a]

Model Scheme	Percentage in Agreement	Percentage of Overpredictions	Percentage of Underpredictions	Geometric Mean Bias	Geometric s.d.
Deep layer	24[b], 43[c]	25[b], 15[c]	51[b], 42[c]	0.53	4.97
25FL layer	23[b], 67[c]	17[b], 2[c]	60[b], 30[c]	0.43	4.79
Hybrid	29[b], 50[c]	29[b], 22[c]	42[b], 28[c]	0.81	4.43

[a]Observations are from the data set described in Table 1.
[b]Agreement assessed using a factor of 2 uncertainty in the observations but no model uncertainty.
[c]Agreement assessed using uncertainty in both the observations and in the model predictions. The considered uncertainty in the model predictions is due to positional errors in the ash cloud of up to two grid-boxes in each horizontal direction and, for the 25FL layer scheme only, one grid-box in the vertical.

between the concentration zones (modeled concentrations less than 200 $\mu g/m^3$, 200–2000 $\mu g/m^3$, 2000–4000 $\mu g/m^3$ and greater than 4000 $\mu g/m^3$). The diagonal dot-dashed lines are described below.

[38] The scatter is large as would be expected for a dispersion problem of this type where there are many possible sources of error. The general magnitude of the predicted peak ash concentrations agrees reasonably well with the observations. This indicates that the overall conversion factor, i.e., the product of the assumed distal fine ash fraction and the peak-to-mean ratio, which also accounts for errors in the source, is within the correct range. The results are comparable for the different model schemes, with modeled and observed values lying within about one and a half orders of magnitude of each other, except that in the 25FL layer scheme there are a number of predictions of zero or near-zero peak ash concentrations which do not agree well with the observations. As discussed above, this is to be expected – the 25FL layer scheme is more sensitive to errors in the altitude of the predicted peak concentration and hence, when the peak concentrations are not predicted to occur at roughly the correct altitude, this scheme has a tendency to underpredict at the height of the observed peak.

[39] Table 2 shows the percentage of modeled values overpredicting, underpredicting and in agreement with the observations for the three model schemes. Agreement is assessed in two ways. In the first, an uncertainty in the observations of a factor of 2 is assumed and the modeled and observed values are said to be in agreement if the predicted peak ash concentration lies within a factor of two of the observation. In the second, some attempt has been made to account for uncertainties in the modeled peak ash concentrations due to slight positional errors in the predicted ash cloud by considering the variability in the modeled concentrations over nearby model output grid-boxes. An ash cloud positional error of up to two grid-boxes (40 km resolution) in each horizontal direction and, for the 25FL layer scheme only, one grid-box up or down (25FL resolution) in the vertical direction has been considered. Agreement is said to occur if the uncertainty ranges for the model prediction and the observation overlap.

[40] In addition, the geometric mean bias and geometric standard deviation (s.d.) of the error, representing the bias and the spread in Figure 4, are presented in Table 2. These are defined as $\exp(\mu)$ and $\exp(\sigma)$ where μ and σ are the mean and standard deviation of the error in $\ln(\max(\text{concentration}, 20\ \mu g/m^3))$. The modeled and observed values are limited to 20 $\mu g/m^3$ for the purpose of calculating these statistics in order to prevent the noise in the low concentrations having an undue influence on the statistics. This may lead to a small underestimate in the geometric standard deviation. Alternative approaches are possible here such as excluding small values (although this is problematic when only one of the observed and predicted values is small) or replacing the log function with the distance on the scatterplot (which is linear for small values) but these choices are no less arbitrary and the 20 $\mu g/m^3$ limit has the advantage of simplicity. A geometric mean bias of 1 denotes no bias and a geometric standard deviation of 1 denotes no scatter. The three dot-dashed diagonal lines in Figure 4 show the bias (μ) plus or minus two standard deviations (σ).

[41] When no account is taken of uncertainty in the modeled values, between 23% and 29% of modeled and observed values are found to be in agreement. The 25FL layer scheme performs similarly to the other model schemes by this measure; however, of the values not in agreement, this scheme shows a greater fraction of model underpredictions with 60% of the observations underpredicted, presumably mainly due to errors in the predicted ash cloud's vertical position. For all three schemes, there is a significant improvement in the percentage of modeled and observed values in agreement when uncertainty in the modeled values (due to slight positional errors in the predicted ash cloud) is considered. This improvement is especially marked for the 25FL layer scheme and suggests that there would be significant benefit in using a 'buffer zone' with this scheme, i.e. there would be benefit in using the maximum of the values over nearby grid-boxes in any hazard assessment. When both uncertainty in the observations and uncertainty in the predicted ash cloud are considered, the results for the three schemes range from 43% of values in agreement for the deep layer scheme to 67% of values in agreement for the 25FL scheme. This suggests there is value in using thin layers and that it may be better to use these directly (with a buffer zone) rather than taking the maximum over a much deeper layer.

[42] In general, the hybrid scheme gives higher predicted peak concentrations than the deep layer scheme, typically by about 50%. This suggests the tuning of the two schemes is slightly different and that perhaps the peak-to-mean ratios in the two schemes should differ by a factor of about three instead of two. This difference in the peak-to-mean ratio implies that, with 25FL output, NAME can resolve concentrations about three times as large as with ~200FL output.

[43] The geometric mean bias is less than 1 which suggests a general underprediction by the model. We note that most of the large underpredictions (i.e., underpredictions by a factor of at least 10) for the deep layer and hybrid schemes occur on May 4, 5 and 14. These are days when we know there are significant ash cloud position errors as discussed by Devenish et al. (submitted manuscript, 2012) and A. L. M.

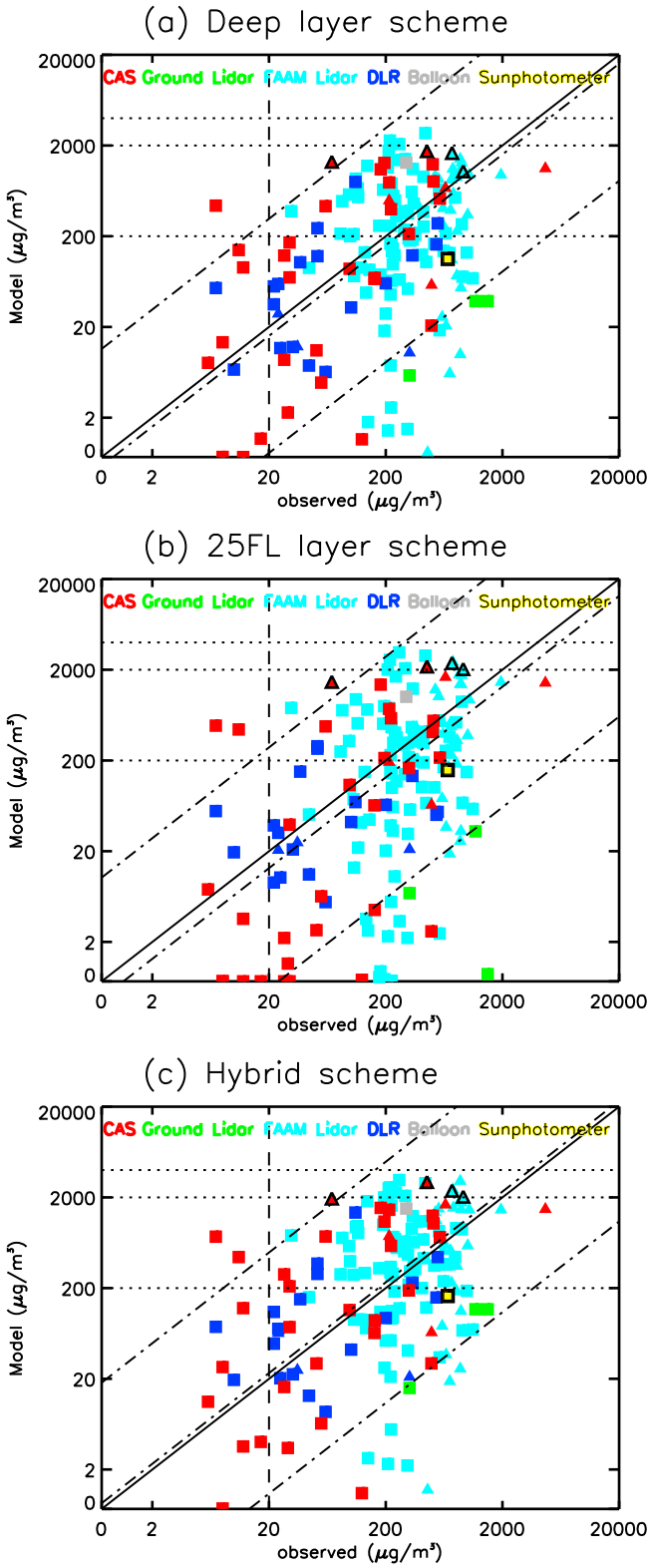

Figure 5. Same as Figure 4 but using the revised time profile of plume height.

Grant et al. (Horizontal and vertical structure of the Eyjafjallajökull ash cloud over the UK, a comparison of airborne lidar observations and simulations, manuscript in preparation, 2012). In fact 24% of the observations used here are from May 14, 15% are from May 5 and 11% are from May 4, illustrating the uneven sampling of the ash cloud inherent in the available data.

[44] The deep layer and hybrid schemes have geometric standard deviations between 4 and 5. In other words, for these schemes, the error is equal to the geometric standard deviation when the model prediction and observation differ by a factor of about 4 or 5 (once correction has been made for the bias). The precise definition of geometric standard deviation chosen here is particularly sensitive to the 20 μg/m^3 concentration limit in situations where there are many values below this limit (e.g. the 25FL layer scheme) and so this measure does not fully reflect the increased scatter seen in the 25FL scheme scatterplot.

4. Sensitivity to the Source Properties

[45] It seems likely that a large proportion of the uncertainty in the predicted ash concentrations is due to uncertainties in the effective source properties. This includes uncertainties in the ash emission rate, the plume rise height, the vertical distribution of the released ash, the particle size distribution and the distal fine ash fraction. We assess here the sensitivity of the predicted concentrations to the source properties, in particular to uncertainties in the time profile of the plume height and to uncertainties in the vertical distribution of ash at the source. Model predictions are compared with the same observational data set as considered in section 3 above.

4.1. Sensitivity to the Time Profile of Plume Height

[46] Following the end of the 2010 eruption of Eyjafjallajökull, detailed information on the observed plume height from radar [*Arason et al.*, 2011], from pilot reports (Pireps) and from Icelandic coastguard observations was gathered and studied. A revised best-guess time profile of the plume height was produced (see the solid line in Figure 3) which differs slightly from that used in section 3.4 (referred to below as the 'original' profile).

[47] Figure 5 shows scatterplots of model predictions obtained using the revised plume-height time profile versus the observations. Results for each of the three model schemes are shown. Comparing with the equivalent scatterplots obtained using the original plume-height time profile (Figure 4), we see that the differences in model predictions are sometimes large, illustrating the potential for high sensitivity of the model predictions to the uncertainties in the plume height.

[48] Large differences are seen between the predicted ash concentrations, obtained using the original and the revised plume-height time profiles, on 16 April (ground lidar and sunphotometer at Chilbolton), 19 April (DLR aircraft) and 14 and 16 May (FAAM aircraft). For example, on 14 May 2010 the lidar and CAS instruments onboard the FAAM aircraft gave estimated peak concentrations of 69 μg/m^3 (CAS), 451 μg/m^3 (CAS), 740 μg/m^3 (lidar) and 923 μg/m^3 (lidar) at FL209, FL234, FL226 and FL202, respectively. Using the original plume-height time profile (the dashed line in Figure 3), the corresponding predicted peak ash concentrations were 44/130/130 μg/m^3, 111/56/298 μg/m^3, 82/51/214 μg/m^3 and 37/98/98 μg/m^3, respectively for the deep layer scheme/25FL layer scheme/hybrid scheme (see

Table 3. Statistical Comparison Between Model Predictions, Obtained Using the Revised Time Profile of Plume Height, and Observations[a]

Model Scheme	Percentage in Agreement	Percentage of Overpredictions	Percentage of Underpredictions	Geometric Mean Bias	Geometric s.d.
Deep layer	33[b], 61[c]	28[b], 16[c]	39[b], 23[c]	0.80	4.42
25FL layer	30[b], 75[c]	23[b], 3[c]	48[b], 22[c]	0.65	4.62
Hybrid	29[b], 66[c]	42[b], 20[c]	29[b], 14[c]	1.16	4.62

[a]Observations are from the data set described in Table 1.
[b]Agreement assessed using a factor of 2 uncertainty in the observations but no model uncertainty.
[c]Agreement assessed using uncertainty in both the observations and in the model predictions. The considered uncertainty in the model predictions is due to positional errors in the ash cloud of up to two grid-boxes in each horizontal direction and, for the 25FL layer scheme only, one grid-box in the vertical.

the four triangles marked with a black outline in Figure 4). Using the revised plume-height time profile (the solid line in Figure 3), the predicted peak ash concentrations are significantly higher, namely 1325/1474/1946 μg/m^3, 1738/2176/2962 μg/m^3, 1658/2376/2377 μg/m^3 and 1039/2038/2038 μg/m^3, respectively (see the four triangles marked with a black outline in Figure 5), showing up to a 50-fold increase. These significant differences in the predicted peak ash concentrations appear to be due to relatively minor changes in the plume-height time profile on 13 May. In the revised profile, the plume height increases to its maximum for the day just a few hours earlier than in the original profile (see Figure 3). This subtle difference in plume height over these few hours results in significantly different predicted concentrations of ash at observation points located over northern England and north of Scotland the next day. The changes in the predictions are significantly larger than expected from the changes in the ash emission rate per unit height alone. The heights of the observed peak concentrations are close to the plume rise height and are near the bottom of our FL200-FL350 deep layer. This is likely to enhance the sensitivity because the fraction of material above FL200 in the model will be strongly sensitive to the plume rise height. In addition Devenish et al. (submitted manuscript, 2012) have shown that increases in plume height lead to the ash cloud penetrating further south over Scotland and northern England, on this day, as a result of variations in the wind field with height. These sorts of sensitivity-enhancing effects are likely to occur in many situations.

[49] Another example concerns the peak concentration of 680 μg/m^3 estimated from the lidar and sunphotometer at Chilbolton on 16 April 2010 (Hogan et al., manuscript in preparation, 2012). Using the original plume-height time profile (the dashed line in Figure 3), the predicted peak ash concentrations were 1505/1686/2628 μg/m^3 for the deep layer scheme/25FL layer scheme/hybrid scheme (see the yellow square marked with a black outline in Figure 4). Using the revised plume-height time profile (the solid line in Figure 3), the predicted peak ash concentrations are considerably less, namely 112/157/164 μg/m^3 (see the yellow square marked with a black outline in Figure 5). In this case, the differences in predicted peak ash concentrations at Chilbolton are due to a reduction of the initial eruption height in the revised plume-height time profile (see Figure 3).

[50] Table 3 shows the degree of agreement between observed and modeled values obtained using the revised plume-height time profile. Encouragingly, there is a significant increase in the percentage of modeled and observed values in agreement, suggesting an increase in skill from a more accurate plume-height time profile. Using the revised plume-height time profile, the three schemes have between 61% and 75% of modeled and observed values in agreement when account is taken of the uncertainties in the observations and of the uncertainties in the modeled values (due to slight positional errors in the predicted ash cloud). There is also a reduction in the percentage of underpredictions and an increase in the percentage of overpredictions, so that the model predictions are a little more conservative than before. However, it should be remembered that the observations data set is not unbiased with around a quarter of the measurements made on 14 May (a day on which large differences exist between predicted concentrations obtained using the original and revised plume-height time profiles). The geometric mean biases are closer to 1 with less suggestion of an underprediction while the geometric standard deviations are similar to before.

4.2. Sensitivity to the Vertical Distribution of Ash at the Source

[51] The assumption of a uniform vertical distribution of released mass at the source does not capture the top-heavy weighting expected for the effective source [see, e.g., *Sparks et al.*, 1997]. The sensitivity of model predicted ash concentrations to the vertical distribution of ash is assessed using a simple top-heavy representation of the ash eruption column. The relationship between plume height and emission rate given by *Mastin et al.* [2009] is used as before. However, three-quarters of the mass is released uniformly over the top quarter of the eruption column and the remaining quarter of the mass is released uniformly over the bottom three-quarters of the eruption column, to give a top-heavy weighting of ash. The 'revised' plume-height time profile (described in section 4.1) is used for this sensitivity test. We note that in reality the exact vertical distribution of the effective ash source will vary with eruption characteristics and with the atmospheric profiles of wind speed, temperature and humidity. In particular, it is likely to be different for vertically rising and bent-over eruption columns [see, e.g., *Devenish et al.*, 2012, also submitted manuscript, 2012].

[52] Using the model schemes described in section 3.2 and, in particular, assuming that the distal fine ash fraction and the peak-to-mean ratio used by each scheme are the same as before, we find that the model now seems biased toward overprediction, as assessed from plots like those in Figure 4 (not shown). This suggests that the rather large peak-to-mean ratios required in earlier sections are partly a result of the (probably overly) smoothed-out nature of the uniform source profile assumed. Using a top-heavy weighting

Table 4. Statistical Comparison Between Model Predictions, Obtained Using the Top-Heavy Source Distribution With the Revised Time Profile of Plume Height, and Observations[a]

Model Scheme	Percentage in Agreement	Percentage of Overpredictions	Percentage of Underpredictions	Geometric Mean Bias	Geometric s.d.
Deep layer	29[b], 58[c]	24[b], 12[c]	47[b], 30[c]	0.63	4.59
25FL layer	25[b], 69[c]	22[b], 2[c]	53[b], 29[c]	0.58	4.82
Hybrid	26[b], 60[c]	34[b], 16[c]	40[b], 24[c]	0.87	5.00

[a]Observations are from the data set described in Table 1.
[b]Agreement assessed using a factor of 2 uncertainty in the observations but no model uncertainty.
[c]Agreement assessed using uncertainty in both the observations and in the model predictions. The considered uncertainty in the model predictions is due to positional errors in the ash cloud of up to two grid-boxes in each horizontal direction and, for the 25FL layer scheme only, one grid-box in the vertical.

of the ash column allows the peak-to-mean ratio to be reduced, with a value of 10 for the deep layer scheme and of 5 for the 25FL and hybrid schemes giving roughly unbiased predictions (in the sense that results are scattered more or less symmetrically about the 1:1 line).

[53] Table 4 shows the degree of agreement between modeled and observed values when these revised peak-to-mean ratios are used. The percentage of values in agreement is similar to, although slightly worse than, that obtained using the uniform vertical distribution of ash with the revised plume height (Table 3). The scatter between modeled and observed values as indicated by the geometric standard deviation and scatterplots (not shown) is also slightly worse than that seen in Table 3 and Figure 5. One might hope for an improvement in the results here because a certain amount of top-heavy weighting is expected to be more realistic. Together with the adjusted peak-to-mean ratio, one might expect this to slightly increase concentrations where material originating from near the top of the eruption column is present but to reduce concentration overpredictions in regions dominated by material from low down in the eruption column. However, although results may be more realistic in terms of the general character of the dispersing ash cloud, the less smoothed out nature of the predictions is likely to increase the scatter caused by errors in ash cloud position. This is true whether these errors arise from errors in plume rise height and in the vertical distribution of ash at the source or from other causes such as errors in the meteorology. Also the simple top-heavy weighting tested here is unlikely to be accurate. In the absence of accurate knowledge of the vertical distribution of ash at the source (a distribution which is likely to vary over time and between different volcanic eruptions), there seems little evidence to justify adopting a more detailed vertical profile in preference to a uniform profile.

4.3. Using a More Accurate Representation of the Source Properties Based On Inversion Modeling

[54] To obtain more accurate predictions of ash concentrations, it is probably necessary to have a more accurate representation of the effective source properties, namely the evolution with time of the plume height, the emission rate, the vertical distribution of ash, the particle size distribution and the distal fine ash fraction. These quantities are expected to vary between eruptions and over time for the same volcanic eruption. *Stohl et al.* [2011] developed an inversion method to determine volcanic ash source strengths as a function of time and altitude using satellite derived total column ash amounts and a Lagrangian dispersion model, together with an a priori estimate of the emission profile derived using the Plumeria model [*Mastin*, 2007], the local meteorology and the estimated plume rise height. This method is of course subject to errors in the satellite retrieval processes and in the dispersion modeling (including, in particular, in the meteorology used in the dispersion modeling), but is expected to yield a better representation of the effective source term. Using the Lagrangian model FLEXPART, *Stohl et al.* [2011] applied this method to estimate the source term for the 2010 Eyjafjallajökull eruption. Predicted ash concentrations were then obtained, both for this source term and for the *a priori* source term, again by using the Lagrangian model FLEXPART. Results were compared against observations from the DLR aircraft. Their scatterplot of modeled and observed values shows a large degree of scatter, not unlike that seen in the comparisons here. However, results with the source determined by the inversion method show an improvement in the agreement between modeled and observed ash concentrations compared to results from the a priori source.

[55] A similar source term for the 2010 Eyjafjallajökull eruption has been obtained by *Kristiansen et al.* [2012] using the inversion method of *Stohl et al.* [2011], total ash column retrievals from SEVIRI satellite measurements [*Prata*, 1989; *Stohl et al.*, 2011], and the NAME model. The particle size distribution used (see the solid line in Figure 1) is a specific particle size distribution for the 2010 Eyjafjallajökull eruption [*Stohl et al.*, 2011] and is based on measurements of volcanic ash deposits near to Eyjafjallajökull. In deriving this particle size distribution, *Stohl et al.* [2011] took account of the fact that the particle size distribution of the released ash is likely to be different to that of the ground deposits near to the source, with the latter being biased toward larger particles. In this derivation the effect of aggregation was neglected. The derived particle size distribution differs from that normally used in NAME for volcanic ash (see the dotted line in Figure 1) and is used for all the calculations in this section. A particle density of 2750 kg/m^3 is assumed which is also based on ground based measurements from the 2010 Eyjafjallajökull eruption and differs from that normally used in NAME for volcanic ash. The particle size range detected by the satellites is taken to be 2.8 to 28 μm and hence the inversion method only uses this part of the particle size distribution. The source term determined by the inversion method is then extrapolated over the full particle size range, 0.2 to 300 μm, of the assumed particle size distribution. The resulting source term is complex with emission rates varying frequently with height and time, namely with a time resolution of three hours and a vertical resolution of 650 m.

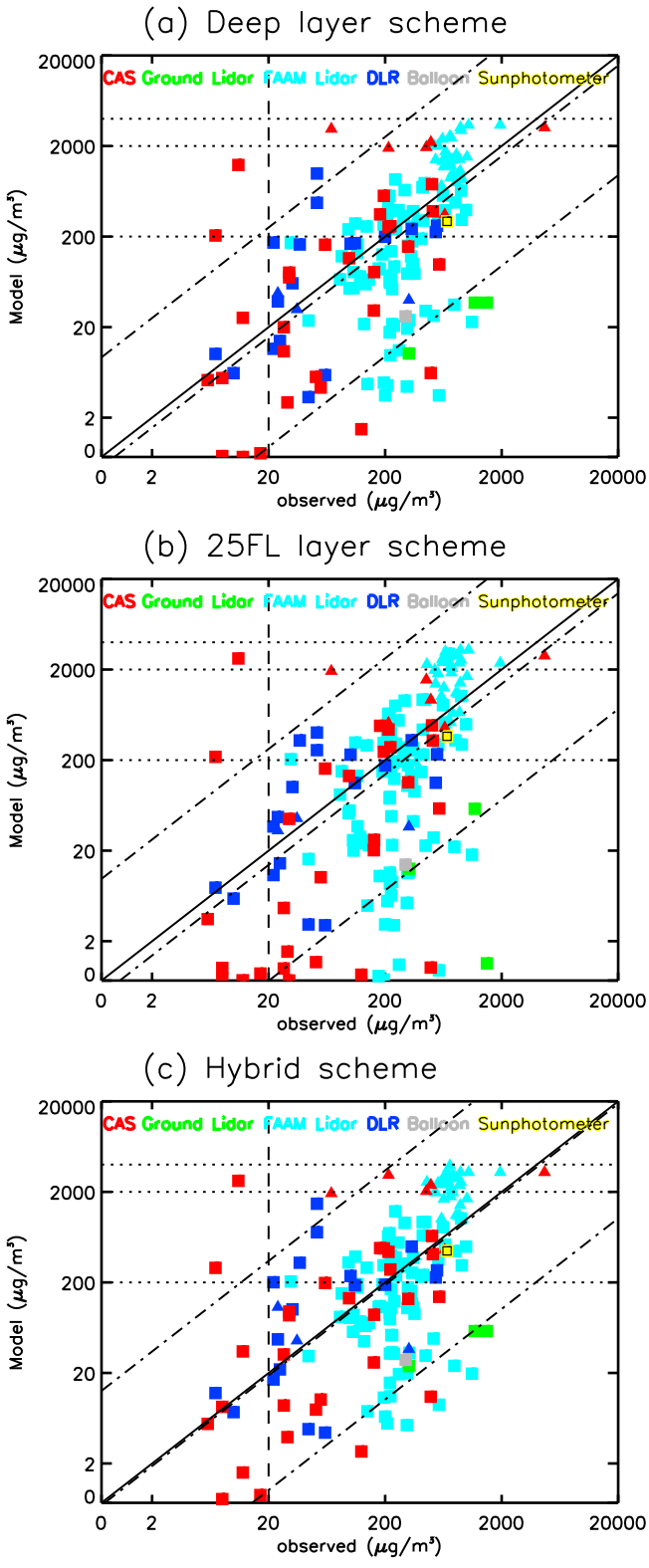

Figure 6. Same as Figure 4 but using the source term determined using the inversion method with a peak-to-mean ratio of 10 for the deep layer scheme and of 5 for the two schemes based on 25FL layer output.

[56] As the satellite retrievals of total column ash are obtained some distance from the source where the near-source fall-out of large particles is assumed to have occurred, the source term derived by the inversion method is the effective source and represents only the proportion of the emitted mass that survives the near-source fall-out. As such, a distal fine ash factor does not need to be applied to this source estimate. However, a peak-to-mean ratio will need to be applied to the NAME predicted mean concentrations to account for unresolved peak ash concentrations. As in section 4.2, it is expected that the source term from the inversion, with its detailed vertical structure, will result in a smaller peak-to-mean ratio being required than was needed with the uniform source. We adopt a value of 10 for the deep layer scheme and of 5 for both the 25FL layer and hybrid schemes. These values give roughly unbiased predictions (in the sense that results are scattered more or less symmetrically about the 1:1 line) and are consistent with the values adopted in section 4.2.

[57] Figure 6 shows scatterplots of predicted versus observed values of peak ash concentration using the inversion-derived source. The scatter is reduced a little from that seen previously when a uniform release profile was assumed and the emission rate was determined from the plume height. Table 5 shows the degree of agreement between modeled and observed values. When model uncertainty is not taken into account, more modeled and observed values are in agreement when the inversion-derived source is used than was the case for the simple uniform source, for both the 'original' (Table 2) and 'revised' (Table 3) plume-height time profiles. In addition, there is an improvement seen in the geometric standard deviation. This indicates that the source term determined by the inversion method gives an increase in skill. However, when model uncertainty due to slight positional errors in the predicted ash cloud is also taken into account, the percentage of modeled and observed values in agreement is similar to that obtained using the simple uniform source with the revised plume-height time profile (although greater than that obtained using the uniform source with the original plume-height time profile). This suggests that, because of the many errors and uncertainties involved, simple source terms with the use of a buffer zone may, in an operational setting, perform similarly to more complex source terms.

[58] Nonetheless, there are potential benefits in using the inversion-derived source term. As noted above, results with this source term, even with a buffer zone, do show an improvement over the results using the original plume-height time profile (see Figure 4 and Table 2) which was all that was available for making rapid predictions in real time. Also we note that both the mass emission rate versus plume height relationship and the distal fine ash fraction are likely to vary between volcanic eruptions and potentially between each phase of an eruption. The inversion method provides a way of automatically tuning for such changes. Hence, although the results presented here do not show a substantial benefit, the inversion approach may well be more beneficial for other eruptions and would avoid the need for manual retuning. Furthermore, the inversion method would be directly applicable for eruptions in poorly monitored volcanic regions where detailed observations of the plume rise height are not likely to be available. On the other hand, we

Table 5. Statistical Comparison Between Model Predictions, Obtained Using the Source Term Determined Using the Inversion Method, and Observations[a]

Model Scheme	Percentage in Agreement	Percentage of Overpredictions	Percentage of Underpredictions	Geometric Mean Bias	Geometric s.d.
Deep layer	44[b], 66[c]	21[b], 9[c]	35[b], 25[c]	0.77	4.03
25FL layer	35[b], 70[c]	25[b], 1[c]	40[b], 29[c]	0.70	4.37
Hybrid	37[b], 65[c]	32[b], 15[c]	32[b], 20[c]	0.94	4.29

[a]Observations are from the data set described in Table 1.
[b]Agreement assessed using a factor of 2 uncertainty in the observations but no model uncertainty.
[c]Agreement assessed using uncertainty in both the observations and in the model predictions. The considered uncertainty in the model predictions is due to positional errors in the ash cloud of up to two grid-boxes in each horizontal direction and, for the 25FL layer scheme only, one grid-box in the vertical.

note that there are limitations to this method when satellite observations of the volcanic ash cloud are not available due, say, to the ash cloud being obscured by cloud or being located in a region without high-quality satellite coverage. In addition, this post-event analysis of the event has enabled all satellite observations to be used in constraining the source term. In real-time forecasting only those observations known at the time of prediction would be available and this may impact on the accuracy of the source term.

5. Discussion

[59] The 2010 Eyjafjallajökull eruption is the first time ash concentrations within volcanic clouds have been predicted in an operational aviation-hazard context. Predicting volcanic ash concentrations accurately is an extremely difficult and challenging task for many reasons. Volcanic eruptions are complex processes with eruption intensities, ash properties (including particle size distributions) and eruption column dynamics likely to vary substantially over time and between different eruptions, including in response to varying atmospheric conditions. Processes such as aggregation are not understood well enough for reliable prediction and accurate eruptive information is often unknown. Despite these difficulties, the comparisons shown here between predicted and observed concentrations have demonstrated an ability to provide useful guidance on likely peak ash concentrations. This work has shown, however, that the uncertainties involved can be numerous and large. Errors in predicted ash concentrations of about one and a half orders of magnitude have been found. In general, errors of this size are to be expected and larger errors are possible, particularly in the initial stages of an eruption before observational evidence is obtained to assess the accuracy of estimates of the eruption strength and the distal fine ash fraction.

[60] An important advantage of giving quantitative predictions is that results can be compared more precisely with measurements leading to a more objective assessment of prediction skill and of potential improvements in modeling approaches. Previous approaches based on forecasting areas of 'visual ash' depend on imprecisely defined thresholds which makes predictions hard to test, verify and improve. Quantitative predictions also enable better interfacing with engine manufacturers estimates of hazard levels and make it possible to use any future improvements in such estimates. Of course there is a danger that quantitative predictions are interpreted as implying a high degree of accuracy. To avoid this danger, the uncertainties in quantitative forecasts must be communicated appropriately.

[61] In the operational schemes set out in sections 2.1 and 2.2 a certain degree of tuning is advisable for each eruption (or as conditions change during an eruption) guided by observations and geological information. As more information becomes available through the course of an eruption and tuning is performed, confidence in the predicted concentrations is likely to increase. This was put into practice during the 2011 eruption of Grimsvötn where satellite observations and volcanological information indicated that the ash emission rate and/or distal fine ash fraction assumed initially were too large for part of the eruption. We note also that, although the default assumptions in the schemes are not unduly conservative, there is less scope for underprediction than overprediction, in that the default fine ash fraction (5%) is toward the top of the range of 0.05% to 10% (in the sense that it may lead to an underprediction of a factor of 2 but could lead to an overprediction of a factor of 100, all else being equal). Of course even with tuning, the advice provided to aviation should be an assessment based on all the available evidence and not just raw model output.

[62] The need for a peak-to-mean ratio is an interesting aspect of the approach. Because the source of ash is localized, the scales associated with the concentration variations are much smaller than would be expected for many other atmospheric scalars such as water vapor or ozone, as exemplified by the thin layers of ash which often occur. As a result the resolved model concentrations depend more strongly on the effective resolution than is the case for many other scalars. The peak-to-mean ratio encompasses a range of uncertainties and errors in addition to representing the difference between local peak concentrations and mean concentrations in space and time. For example, the assumed uniform source profile requires a larger peak-to-mean ratio than the top heavy profile in section 4.2 or the profile derived by *Kristiansen et al.* [2012] because using a uniform profile reduces the mean concentrations within the model by spreading material more widely. The most appropriate value of the peak-to-mean ratio may also depend on the amount of sub-grid diffusion in the model. Within certain limits, the amount of sub-grid diffusion can be regarded as a matter of choice, following the ideas of *Mason and Callen* [1986]. A high value will smooth out the small scale structures and give lower peaks but may give better agreement with observations if these small scales are not predictable. In contrast a small value may retain more of the small scales and give higher peak predictions but may give worse agreement with observations if the peaks are in the wrong place. Conceptually these differences can be thought of as aiming to predict the ensemble mean

concentration for different size ensembles. We have not investigated this here but note that some investigations of the sensitivity to the sub-grid diffusion are presented by *Devenish et al.* [2012, also submitted manuscript, 2012]. The peak-to-mean ratio may also account for a bias caused by ash cloud positional errors since the observations are local peak concentrations while the model predictions at the same location may not be peak values in the presence of errors in the ash cloud position. Techniques such as using meteorological ensembles or inclusion of buffer zones may be a better way to account for the effects of ash cloud positional errors.

[63] The ash concentration forecasting method has been deliberately kept relatively simple in view of the many uncertainties. However, a number of possible improvements deserve consideration. Methods to address ash cloud position errors are one such possible improvement and are discussed in the previous paragraph. Another example is the use of plume rise models to account for the effect of variations in meteorology on the relation between the mass emission rate and the plume rise height (e.g., the Plumeria model [*Mastin*, 2007] which accounts for atmospheric profiles of temperature and humidity, but not of wind). In principle this could even attempt to account for cases where minor eruptions trigger deep convection (or perhaps could just give a warning that this might occur, as the eruption rate is unlikely to be strongly constrained by the total rise height in such cases). The use of a top-heavy emission profile (section 4.2) did not show any benefit in this study suggesting that there is little value in improving the vertical emission profile without case-specific information. However, it seems premature to rule this out. More promising is the inversion approach to source estimation which has been pioneered in the volcanic ash context by *Stohl et al.* [2011] and which provides a method to adjust the effective source strength (the product of the emission rate and the distal fine ash fraction) and the vertical profile of the emission automatically. Another issue worth noting is that the approach adopted here assumes that, except in the immediate vicinity of the volcano, the dispersion can be considered as passive – essentially we are assuming a scale separation between the near source volcanic plume dynamics and the larger scale meteorological processes. Such an assumption may well be untenable for really large eruptions such as Pinatubo 1991, where the umbrella cloud dynamics played a key role in the spread of the umbrella cloud which reached a diameter of over 1,100 km and intruded more than 200 km upwind [*Self et al.*, 1996]. Hence there is a case for considering using models of the volcanic plume dynamics for more than just relating rise height to emission rate. Even for much smaller eruptions, this might be useful for predictions close to the volcano. We also note that, although individually many of these improvements seem straightforward to develop given appropriate resources, the problem of integrating them operationally, ideally within a probabilistic risk framework, is very significant.

[64] The aim of the concentration forecasting methods considered in this paper has been to estimate local peak ash concentrations. However, there may be demand in the future for additional forecasts such as route-based integrated air concentrations. This does not completely remove the need for some accounting of differences between peak and mean concentrations because, although sometimes such fluctuations may be averaged out, this will not always be so, for example if an aircraft is flying for some time within a thin layer of ash.

[65] Finally we note that, while results are encouraging, the data all come from only one eruption. Further verification of concentration forecasting methods against future volcanic eruptions would improve understanding of this scientific area.

[66] **Acknowledgments.** The authors would like to thank the Icelandic Meteorological Office for the provision of invaluable information on the eruption (both during the eruption and subsequently), and Barbara Stunder (NOAA) for discussions about the table of thresholds developed for the VAFTAD model. The Facility for Airborne Atmospheric Measurements is jointly funded by the Met Office and the National Environment Research Council. R. S. J. Sparks acknowledges support from the European Research Council.

References

Ansmann, A., et al. (2010), The 16 April 2010 major volcanic ash plume over central Europe: EARLINET lidar and AERONET photometer observations at Leipzig and Munich, Germany, *Geophys. Res. Lett.*, *37*, L13810, doi:10.1029/2010GL043809.

Arason, P., G. N. Petersen, and H. Björnsson (2011), Observations of the altitude of the volcanic plume during the eruption of Eyjafjallajökull, April–May 2010, *Earth Syst. Sci. Data*, *3*, 9–17.

Boughton, B. A., J. M. Delaurentis, and W. E. Dunn (1987), A stochastic-model of particle dispersion in the atmosphere, *Boundary Layer Meteorol.*, *40*, 147–163.

Casadevall, T. J. (1994), The 1989–1990 eruption of Redoubt Volcano, Alaska: impacts on aircraft operations, *J. Volcanol. Geotherm. Res.*, *62*, 301–316.

Costa, A., A. Folch, and G. Macedonio (2010), A model for wet aggregation of ash particles in volcanic plumes and clouds: 1. Theoretical formulation, *J. Geophys. Res.*, *115*, B09201, doi:10.1029/2009JB007175.

Dacre, H. F., et al. (2011), Evaluating the structure and magnitude of the ash plume during the initial phase of the 2010 Eyjafjallajökull eruption using lidar observations and NAME simulations, *J. Geophys. Res.*, *116*, D00U03, doi:10.1029/2011JD015608.

Devenish, B. J., D. J. Thomson, F. Marenco, S. J. Leadbetter, H. Ricketts, and H. F. Dacre (2012), A study of the arrival over the United Kingdom in April 2010 of the Eyjafjallajökull ash cloud using ground-based lidar and numerical simulations, *Atmos. Environ.*, doi:10.1016/j.atmosenv.2011.06.033, in press.

Durant, A. J., W. I. Rose, A. M. Sarna-Wojcicki, S. Carey, and A. C. M. Volentik (2009), Hydrometeor-enhanced tephra sedimentation: Constraints from the 18 May 1980 eruption of Mount St Helens, *J. Geophys. Res.*, *114*, B03204, doi:10.1029/2008JB005756.

Folch, A., A. Costa, A. Durant, and G. Macedonio (2010), A model for wet aggregation of ash particles in volcanic plumes and clouds: 2. Model application, *J. Geophys. Res.*, *115*, B09202, doi:10.1029/2009JB007176.

Galmarini, S., F. Bonnardot, A. Jones, S. Potempski, L. Robertson, and M. Martet (2010), Multi-model vs. EPS-based ensemble atmospheric dispersion simulations: A quantitative assessment on the ETEX-1 tracer experiment case, *Atmos. Environ.*, *44*, 3558–3567.

Guffanti, M., T. J. Casadevall, and K. Budding (2010), Encounters of aircraft with volcanic ash clouds; A compliation of known incidents, 1953–2009, *U. S. Geol. Surv. Data Ser.*, *545*, 16 pp.

Hanstrum, B. N., and A. S. Watson (1983), A case study of two eruptions of Mount Galunggung and an investigation of volcanic eruption cloud characteristics using remote sensing techniques, *Aust. Meteorol. Mag.*, *31*, 171–177.

Harrison, R. G., K. A. Nicoll, Z. Ulanowski, and T. A. Mather (2010), Self-charging of the Eyjafjallajökull volcanic ash cloud, *Environ. Res. Lett.*, *5*, 024004, doi:10.1088/1748-9326/5/2/024004.

Heffter, J. L., and B. J. B. Stunder (1993), Volcanic Ash Forecast Transport and Dispersion (VAFTAD) model, *Weather Forecast.*, *8*, 534–541.

Hobbs, P. V., L. F. Radke, J. H. Lyons, R. J. Ferek, D. J. Coffman, and T. J. Casadevall (1991), Airborne measurements of particle and gas emissions from the 1990 volcanic eruptions of Mount Redoubt, *J. Geophys. Res.*, *96*, 18,735–18,752.

International Civil Aviation Organization (2002), *Manual on volcanic ash, radioactive material and toxic chemical clouds*, 1st ed., Doc. 9691, 159 pp., Montreal, Que., Canada.

Jones, A., D. Thomson, M. Hort, and B. Devenish (2007), The U.K. Met Office's next generation atmospheric dispersion model, NAME III, in *Air Pollution Modeling and its Application XV11: Proceedings of the*

27th NATO/CCMS International Technical Meeting on Air Pollution Modeling and its Application, edited by Borrego, C., and A.-L. Norman, pp. 580–589, Springer, New York.

Kristiansen, N. I., et al. (2012), Performance assessment of a volcanic ash transport model mini-ensemble used for inverse modelling of the 2010 Eyjafjallajökull eruption, *J. Geophys. Res.*, doi:10.1029/2011JD016844, in press.

Leadbetter, S., and M. Hort (2011), Volcanic ash hazard climatology for an eruption of Hekla Volcano, Iceland, *J. Volcanol. Geotherm. Res.*, *199*, 230–241.

Marenco, F., and R. J. Hogan (2011), Determining the contribution of volcanic ash and Boundary Layer aerosol in backscatter lidar returns: a three-component atmosphere approach, *J. Geophys. Res.*, *116*, D00U06, doi:10.1029/2010JD015415.

Marenco, F., B. Johnson, K. Turnbull, S. Newman, J. Haywood, H. Webster, and H. Ricketts (2011), Airborne lidar observations of the 2010 Eyjafjallajökull volcanic ash plume, *J. Geophys. Res.*, *116*, D00U05, doi:10.1029/2011JD016396.

Maryon, R. H., D. B. Ryall, and A. L. Malcolm (1999), The NAME 4 Dispersion Model: Science documentation, *Met Off. Turbulence Diffus. Note 262*, 45 pp., Met Off., Exeter, U. K.

Mason, P. J., and N. S. Callen (1986), On the magnitude of the subgrid-scale eddy coefficient in large-eddy simulations of turbulent channel flow, *J. Fluid Mech.*, *162*, 439–462.

Mastin, L. G. (2007), A user-friendly one-dimensional model for wet volcanic plumes, *Geochem. Geophys. Geosyst.*, *8*, Q03014, doi:10.1029/2006GC001455.

Mastin, L. G., 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, S. C., R. W. Saunders, P. N. Francis, and H. N. Webster (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.*, doi:10.1029/2011JD016770, in press.

Newman, S., L. Clarisse, D. Hurtmans, F. Marenco, B. T. Johnson, K. F. Turnbull, S. Havemann, A. J. Baran, D. O'Sullivan, and J. M. Haywood (2012), A case study of observations of volcanic ash from the Eyjafjallajökull eruption: 2. Airborne and satellite radiative measurements, *J. Geophys. Res.*, doi:10.1029/2011JD016780, in press.

Potempski, S., et al. (2008), Multi-model ensemble analysis of the ETEX-2 experiment, *Atmos. Environ.*, *42*, 7250–7265.

Prata, A. J. (1989), Radiative transfer calculations for volcanic ash clouds, *Geophys. Res. Lett.*, *16*, 1293–1296.

Pruppacher, H. R., and J. D. Klett (1999), *Microphysics of Clouds and Precipitation*, 2nd ed., Kluwer Acad., Dordrecht, Netherlands.

Rose, W. I., G. J. S. Bluth, and G. G. J. Ernst (2000), Integrating retrievals of volcanic cloud characteristics from satellite remote sensors: A summary, *Philos. Trans. R. Soc. Ser. A*, *358*, 1585–1606.

Ryall, D. B., and R. H. Maryon (1998), Validation of the UK Met Office's NAME model against the ETEX dataset, *Atmos. Environ.*, *32*, 4265–4276.

Schumann, U., et al. (2011), Airborne observations of the Eyjafjalla volcano ash cloud over Europe during air space closure in April and May 2010, *Atmos. Chem. Phys.*, *11*, 2245–2279, doi:10.5194/acp-11-2245-2011.

Self, S., J.-X. Zhao, R. E. Holasek, R. C. Torres, and A. J. King (1996), The atmospheric impact of the 1991 Mount Pinatubo eruption, in *Fire and Mud—Eruptions and Lahars of Mount Pinatubo, Philippines*, edited by C. G. Newhall and R. S. Punongbayan, pp. 1089–1115, Univ. of Wash. Press, Seattle.

Sparks, R. S. J., M. I. Bursik, S. N. Carey, J. S. Gilbert, L. S. Glaze, H. Sigurdsson, and A. W. Woods (1997), *Volcanic Plumes*, John Wiley, Chichester, U. K.

Stohl, A., et al. (2011), Determination of time- and height-resolved volcanic ash emissions and their use for quantitative ash dispersion modeling: the 2010 Eyjafjallajökull eruption, *Atmos. Chem. Phys.*, *11*, 4333–4351.

Thomson, D. J., and A. R. Jones (2011), A new puff modelling technique for short range dispersion applications, *Int. J. Environ. Pollut.*, *44*, 156–163.

Tupper, A., J. S. Oswalt, and D. Rosenfeld (2005), Satellite and radar analysis of the volcanic-cumulonimbi at Mount Pinatubo, Philippines, 1991, *J. Geophys. Res.*, *110*, D09204, doi:10.1029/2004JD005499.

Tupper, A., J. Davey, P. Stewart, B. Stunder, R. Servranckx, and A. Prata (2006), Aircraft encounters with volcanic clouds over Micronesia, Oceania, 2002–03, *Aust. Meteorol. Mag.*, *55*, 289–299.

Turnbull, K. F., B. T. Johnson, F. Marenco, J. M. Haywood, A. Minikin, B. Weinzierl, H. Schlager, U. Schumann, S. J. Leadbetter, and A. M. Woolley (2012), A case study of observations of volcanic ash from the Eyjafjallajökull eruption: 1. In situ airborne observations., *J. Geophys. Res.*, doi:10.1029/2011JD016688, in press.

Underwood, B. Y. (1999), A review of dry and wet deposition, *AEAT-5382*, AEA Technology, London.

Webley, P. W., B. J. B. Stunder, and K. G. Dean (2009), Preliminary sensitivity study of eruption source parameters for operational volcanic ash cloud transport and dispersion models - A case study of the August 1992 eruption of the Crater Peak vent, Mount Spurr, Alaska, *J. Volcanol. Geotherm. Res.*, *186*, 108–119.

Webster, H. N., and D. J. Thomson (2002), Validation of a Lagrangian model plume rise scheme using the Kincaid data set, *Atmos. Environ.*, *36*, 5031–5042.

Webster, H. N., and D. J. Thomson (2012), Dry deposition modelling in a Lagrangian dispersion model, *Int. J. Environ. Pollut.*, in press.

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

B. J. Devenish, B. W. Golding, J. M. Haywood, I. P. C. Heard, M. C. Hort, B. T. Johnson, S. J. Leadbetter, A. J. Manning, F. Marenco, D. J. Thomson, K. Turnbull, H. N. Webster, and C. S. Witham, Met Office, FitzRoy Road, Exeter EX1 3PB, UK. (helen.webster@metoffice.gov.uk)

J. Dorsey, NCAS, University of Manchester, Oxford Road, Manchester M13 9PL, UK.

N. I. Kristiansen, NILU, Norwegian Institute for Air Research, PO Box 100, N-2027 Kjeller, Norway.

S. C. Loughlin, British Geological Survey, Edinburgh EH9 3LA, UK.

A. Minikin, U. Schumann, and B. Weinzierl, Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, D-82234 Oberpfaffenhofen, Germany.

R. S. J. Sparks, Department of Earth Sciences, University of Bristol, Bristol BS8 1RJ, UK.