University of Reading

How accurate are volcanic ash simulations of the 2010 Eyjafjallajökull eruption?

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Dacre, H. F. ORCID: https://orcid.org/0000-0003-4328-9126, Harvey, N. J. ORCID: https://orcid.org/0000-0003-0973-5794, Webley, P. W. and Morton, D. (2016) How accurate are volcanic ash simulations of the 2010 Eyjafjallajökull eruption? Journal of Geophysical Research: Atmospheres, 121 (7). pp. 3534-3547. ISSN 2169-8996 doi: 10.1002/2015JD024265 Available at https://centaur.reading.ac.uk/65532/

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

To link to this article DOI: http://dx.doi.org/10.1002/2015JD024265

Publisher: American Geophysical Union

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

www.reading.ac.uk/centaur

University of Reading

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE
10.1002/2015JD024265

Key Points:
- Objective spatial verification metric agrees with subjective visual assessment of model skill
- Model skill decreases as volcanic ash residence time increases
- Errors in the wind fields decrease model skill, particularly near bifurcation points

Supporting Information:
- Movie S1
- Figure S1
- Figure S2
- Figure S3
- Supporting Information S1

Correspondence to:
H. F. Dacre,
h.f.dacre@reading.ac.uk

Citation:
Dacre, H. F., N. J. Harvey, P. W. Webley, and D. Morton (2016), How accurate are volcanic ash simulations of the 2010 Eyjafjallajökull eruption?, *J. Geophys. Res. Atmos.*, *121*, 3534–3547, doi:10.1002/2015JD024265.

Received 23 SEP 2015
Accepted 7 MAR 2016
Accepted article online 11 MAR 2016
Published online 8 APR 2016

The copyright line for this article was changed on 26 JUL 2016 after original online publication.

©2016. The Authors.
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

How accurate are volcanic ash simulations of the 2010 Eyjafjallajökull eruption?

H. F. Dacre[1], N. J. Harvey[1], P. W. Webley[2], and D. Morton[3]

[1]Department of Meteorology, University of Reading, Reading, UK, [2]Geophysical Institute, University of Alaska Fairbanks, Fairbanks, Alaska, USA, [3]Boreal Scientific Computing LLC, Fairbanks, Alaska, USA

Abstract In the event of a volcanic eruption the decision to close airspace is based on forecast ash maps, produced using volcanic ash transport and dispersion models. In this paper we quantitatively evaluate the spatial skill of volcanic ash simulations using satellite retrievals of ash from the Eyjafjallajökull eruption during the period from 7 to 16 May 2010. We find that at the start of this period, 7–10 May, the model (FLEXible PARTicle) has excellent skill and can predict the spatial distribution of the satellite-retrieved ash to within 0.5° × 0.5° latitude/longitude. However, on 10 May there is a decrease in the spatial accuracy of the model to 2.5°× 2.5° latitude/longitude, and between 11 and 12 May the simulated ash location errors grow rapidly. On 11 May ash is located close to a bifurcation point in the atmosphere, resulting in a rapid divergence in the modeled and satellite ash locations. In general, the model skill reduces as the residence time of ash increases. However, the error growth is not always steady. Rapid increases in error growth are linked to key points in the ash trajectories. Ensemble modeling using perturbed meteorological data would help to represent this uncertainty, and assimilation of satellite ash data would help to reduce uncertainty in volcanic ash forecasts.

1. Introduction

The April–May 2010 eruption of Eyjafjallajökull caused unprecedented closure of European and North Atlantic airspace with global economic losses of $5 billion [*Oxford Economics*, 2010]. In the event of an eruption the decision to close airspace is based on forecast ash maps, produced using volcanic ash transport and dispersion (VATD) models and available observations of volcanic ash concentration. These observations can be derived from both remote sensing data (e.g., radar, lidar, and satellite radiometers) and in situ measurements (e.g., dropsondes and research aircraft).

Observations from automated systems such as satellites and their sensors are a vital tool in the tracking of hazardous volcanic plumes. Their high spatial and temporal resolutions make them particularly suitable for near real-time model validation [*Langmann et al.*, 2013]. Matching satellite imagery and VATD model forecasts is, however, not straightforward because the satellite observations and VATD model forecasts are often not readily comparable and are frequently presented on different map projections. To overcome these issues, *Millington et al.* [2012] used ash concentrations from VATD model forecasts, together with temperature and humidity profiles from weather forecasts, as input to a fast radiative transfer model. They computed radiances from which simulated satellite images were created allowing direct comparison with the observed satellite imagery. These images provide a useful tool for forecasters in the London Volcanic Ash Advisory Centre. For example, satellite imagery can give an indication of the accuracy in the location of the forecast ash from earlier VATD model forecasts. A close collocation of the VATD model and satellite ash signals gives confidence in the VATD model forecasts, while a mismatch in locations indicates that the forecaster needs to adjust the current and future volcanic advisories [*Millington et al.*, 2012].

While the above comparison of VATD model output and satellite imagery is useful, it must be applied manually and thus it does not offer an objective measure of the spatial accuracy of the VATD model forecast. This makes it difficult to (i) compare the skill of forecasts made at different times, (ii) compare the skill of forecasts made by different models, (iii) assess the behavior of data assimilation methods, and (iv) quantify the sensitivity of forecast skill to changes in model formulation. The aim of this paper is to use an objective measure to assess quantitatively the spatial and temporal variation in the skill of volcanic ash simulations when compared to satellite observations.

The perceived accuracy of any simulation depends on the scale over which it is being assessed. For example, it is easier to predict the presence of volcanic ash somewhere within a large area than a small one. Almost all of the quantitative verification of volcanic ash simulations in the literature [e.g., *Grant et al.*, 2012; *Webster et al.*, 2012; *Folch et al.*, 2012; *Dacre et al.*, 2015] have been concerned with assessing the performance at point locations (i.e., where there are measurements). The problem with this is that in many situations such a verification approach fails to recognize when a simulation contains useful information unless it happens to be correct at a particular location. Point-based verification techniques are unable to discriminate between simulations in which the volcanic ash is nearly in the correct place and simulations that are wrong by a larger margin.

In response to the problem outlined above, *Harvey and Dacre* [2015] adapted the neighborhood-based verification method of *Roberts and Lean* [2008] to assess the spatial skill of VATD simulations. In this paper we apply this neighborhood-based verification method to volcanic ash simulations made using a VATD model and quantify the sensitivity of the simulations to changes in model formulation. VATD model simulations of the Eyjafjallajökull eruption are compared to satellite observations made during the period 7–16 May 2010.

2. FLEXPART Model Simulations

Simulations of the transport of volcanic ash in the atmosphere are carried out using FLEXPART (FLEXible PARTicle dispersion model), a Lagrangian particle dispersion model. FLEXPART includes transport, diffusion, wet and dry depositions, and radioactive decay of tracers released from point, line, area, or volume sources with tracer concentrations calculated on a three-dimensional longitude-latitude grid [*Stohl et al.*, 2005]. It was originally developed to simulate the dispersion of dangerous substances from point sources and has been validated against several large-scale tracer experiment data including the Cross Appalachian Tracer Experiment (CAPTEX), the Across North American Tracer Experiment and the European Tracer Experiment [*Stohl et al.*, 1998]. Since then it has been applied in a large number of studies on atmospheric transport including the transport of volcanic gases and particles. For example, *Prata et al.* [2007] used satellite measurements and FLEXPART to determine the mass loading, dispersion, and transport of sulfur dioxide, SO_2, from the volcanic cloud emitted during the May 2006 eruption of Soufriere Hills volcano, Montserrat; *Eckhardt et al.* [2008] used FLEXPART to develop a volcanic source inversion method to constrain the vertical profile of SO_2 injection into the atmosphere; and more recently, FLEXPART was used to estimate the volcanic ash source strength as a function of altitude and time from the 2010 Eyjafjallajökull eruption [*Stohl et al.*, 2011; *Kristiansen et al.*, 2012].

The emission of volcanic ash is modeled by releasing model particles, with each model particle representing a mass of volcanic ash. The model ash particles are carried along by the wind. In this paper, FLEXPART version 8.2 is driven using the three-dimensional (3-D) winds and thermodynamic fields from the GFS (Global Forecast System) produced by National Centers for Environmental Prediction using $1° \times 1°$ latitude/longitude analysis fields updated every 3 h. Ash concentrations are computed by summing the mass of ash particles in areas of $0.5° \times 0.5°$ latitude/longitude, integrated over the entire depth of the atmosphere and over a time period of 1 h. The simulations account for gravitational particle settling [*Naeslund and Thaning*, 1991] and wet and dry depositions [*Stohl et al.*, 2005]. The model simulation starts at 06 UTC on 1 May 2010 and continues until 06 UTC on 17 May 2010.

Several eruption source parameters (ESPs) must be specified to characterize the volcanic emission. The ESPs are as follows: plume rise height, vertical ash emission profile, particle size distribution, ash density, and mass eruption rate. The plume rise height input is taken from measurements provided by the Icelandic Meteorological Office's C-band radar [*Arason et al.*, 2011]. The plume height from the radar varies on a range of timescales from minutes to hours. In this paper no attempt has been made to follow every fluctuation in the radar data. Plume height variations on timescales of 6 h or more are represented. Values of plume height have been chosen to pass through the upper end of the scatter in the 5 min radar values while not reflecting the most extreme values. The volcano plume rise height variation in time used in the simulation is shown in Figure 1. The ash emission profile represents the vertical distribution of ash mass released in the column above the volcano vent. In these simulations the emission profile is uniform from the volcano vent to the plume rise height unless otherwise specified. The particle size distribution specifies the mass fraction of ash in different particle size bins. In these simulations the particles are all assumed to have a diameter of 2 μm unless otherwise specified and the ash density is assumed to be 2500 kg m^{-3}. The mass eruption rate is specified using the plume rise height relationship of *Mastin et al.* [2009]. In order to make quantitative simulations we assume that 95% of the total emitted ash falls out close to the volcano source due to sedimentation of large ash

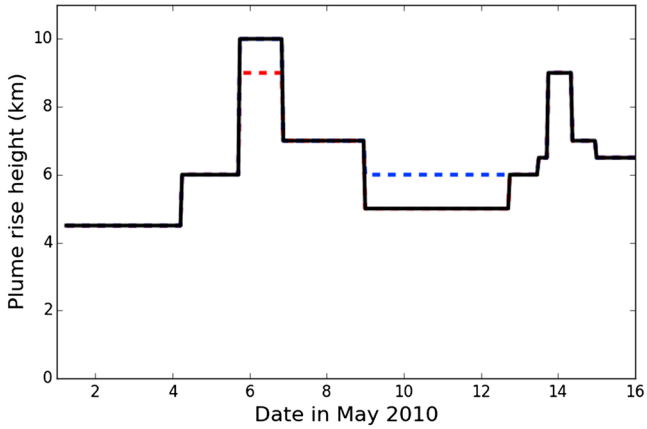

Figure 1. Time series of plume rise height used in the FLEXPART simulations. Modifications to the plume rise height used in section 7.1 are shown as red and blue dashes.

particles (>100 μm) and aggregation of fine ash. This is based on the results of *Dacre et al.* [2011] and *Devenish et al.* [2012] for the Eyjafjallajökull eruption.

3. SEVIRI Satellite Observations

For comparison with the FLEXPART simulations we use volcanic ash retrievals calculated using data from the infrared channels of the Spinning Enhanced Visible and Infrared imager (SEVIRI) instrument mounted onboard Meteosat Second Generation (MSG) satellite. Retrieved quantities are calculated using the radiative transfer model described in *Francis et al.* [2012]. SEVIRI provides high spatial (3 km at the equator), temporal (15 min repeat cycle), and multispectral (12 channels) data making it very suitable for the retrieval of volcanic ash data. The scheme of *Francis et al.* [2012] is based on the difference in brightness temperature between 2 channels centered at 10.8 μm and 12.0 μm, which can be used to discriminate ash from water or ice clouds [*Prata*, 1989]. The strength of the volcanic ash signal is, however, dependent on the optical depth of the ash cloud as well as the physical properties of the ash particles (shape, size distribution, and refractive indices). Other factors affecting the signal include the thermal contrast between the ash cloud top and underlying surface, the presence of other absorbers (water and ice), and satellite-viewing angle (with increased path length through the ash cloud leading to stronger signals) [*Millington et al.*, 2012]. For comparison with the FLEXPART simulations the satellite-observed column-integrated ash loadings are averaged onto a $0.5^\circ \times 0.5^\circ$ latitude/longitude grid and composited over a period of 5 h centered on the verification time. This temporal compositing is performed to create a continuous time series. *Harvey and Dacre* [2015] show that the choice of a 5 h time window is sufficient to fill gaps while maintaining a high correlation with the noncomposited fields. Figures S1 and S2 in the supporting information show a comparison of 1 h and 5 h composited satellite data and the sensitivity of the Fractions Skill Score (described in section 4) to this compositing time window.

4. Fractions Skill Score Methodology

This section contains a brief summary of the neighborhood-based verification method developed by *Harvey and Dacre* [2015]. The method is based on the Fractions Skill Score (FSS) developed by *Roberts and Lean* [2008] to test the skill of high-resolution precipitation forecasts made using the UK Met Office Numerical Weather Prediction model [*Roberts*, 2008; *Mittermaier and Roberts*, 2010]. It compares fractional area coverage in the forecast field with fractional area coverage in the observational field for a specified threshold and a range of neighborhood sizes in order to find a neighborhood size over which the model can be considered skillful. This metric is related to the more general Fisher linear discriminant analysis statistic [*Fisher*, 1936]. The FSS is given by

$$\mathrm{FSS} = 1 - \frac{\mathrm{FBS}}{\mathrm{FBS}_{\mathrm{ref}}} \tag{1}$$

where the Fractions Brier Score (FBS) is a variation on the Brier Score in which both the forecast and observed probabilities (fractions) can have any value between 0 and 1. The FBS is given by

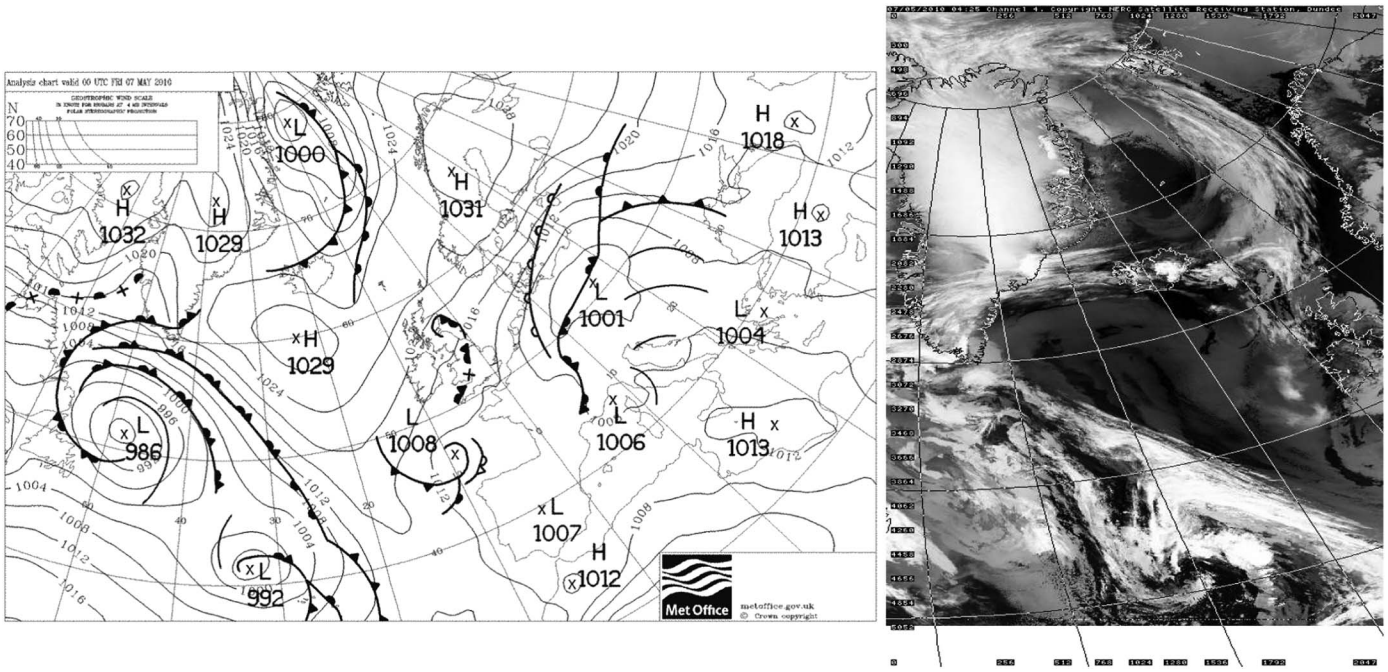

Figure 2. (a) UK Met Office surface analysis chart at 00 Z on 7 May 2010. Mean sea level pressure isobars overlaid with surface fronts. (b) MODIS infrared satellite image at 04:25 Z on 7 May 2010. Provided by the Dundee satellite receiving station.

$$\mathrm{FBS} = \frac{1}{N}\sum_{j=1}^{N}\left(O_j - M_j\right)^2 \tag{2}$$

M_j and O_j are the modeled and observed fractions, respectively, at each point, with values between 0 and 1. N is the number of pixels in the verification area. $\mathrm{FBS}_{\mathrm{ref}}$ is given by

$$\mathrm{FBS}_{\mathrm{ref}} = \frac{1}{N}\left[\sum_{j=1}^{N} O_j^2 + \sum_{j=1}^{N} M_j^2\right] \tag{3}$$

A FSS of 1 indicates a perfect match between the observed and modeled fractions, while a FSS of 0 indicates no overlap. In general, a FSS > 0.5 indicates that the model has skill in capturing the observed distribution. In this paper the neighborhood size is taken to be 5.5° × 5.5° latitude/longitude unless otherwise specified. A simulation that has skill at this neighborhood size could predict the presence of ash regionally in the UK (i.e., distinguish between London, Manchester, and Edinburgh airports). A simulation with skill only at larger scales is considered to be unusable for making decisions to close airspace. Further details on the application of the method to evaluate volcanic ash simulations and comparison with point-based verification metrics can be found in *Harvey and Dacre* [2015].

5. Results

First, we focus on an individual day during the 2010 Eyjafjallajökull eruption, 7 May 2010. Figure 2a shows the UK Met Office surface analysis chart at 00 Z on 7 May 2010. A large region of high pressure is positioned in the North Atlantic with low-pressure systems located to the west and south of the high. Figure 2b shows the MODIS satellite infrared image at 04:25 Z on 7 May 2010. Low-level stratus clouds can be seen colocated with the high-pressure region, and high-level cirrus outflow can be seen above the occluded front to the west of the high.

Figure 3a shows the ash column loading detected by SEVIRI centered on a verification time of 00 Z on 7 May as a binary ash/no-ash field. The ash was detected in a coherent plume extending southward from Iceland to the west of the UK. There is also a narrow band of volcanic ash oriented northwest-southeast between 30 and 20°N plus incoherent speckles of ash over Europe. It is likely that some of these speckles are noise generated by the satellite detection algorithm. Figure 3b shows the FLEXPART-simulated ash column loading at the same

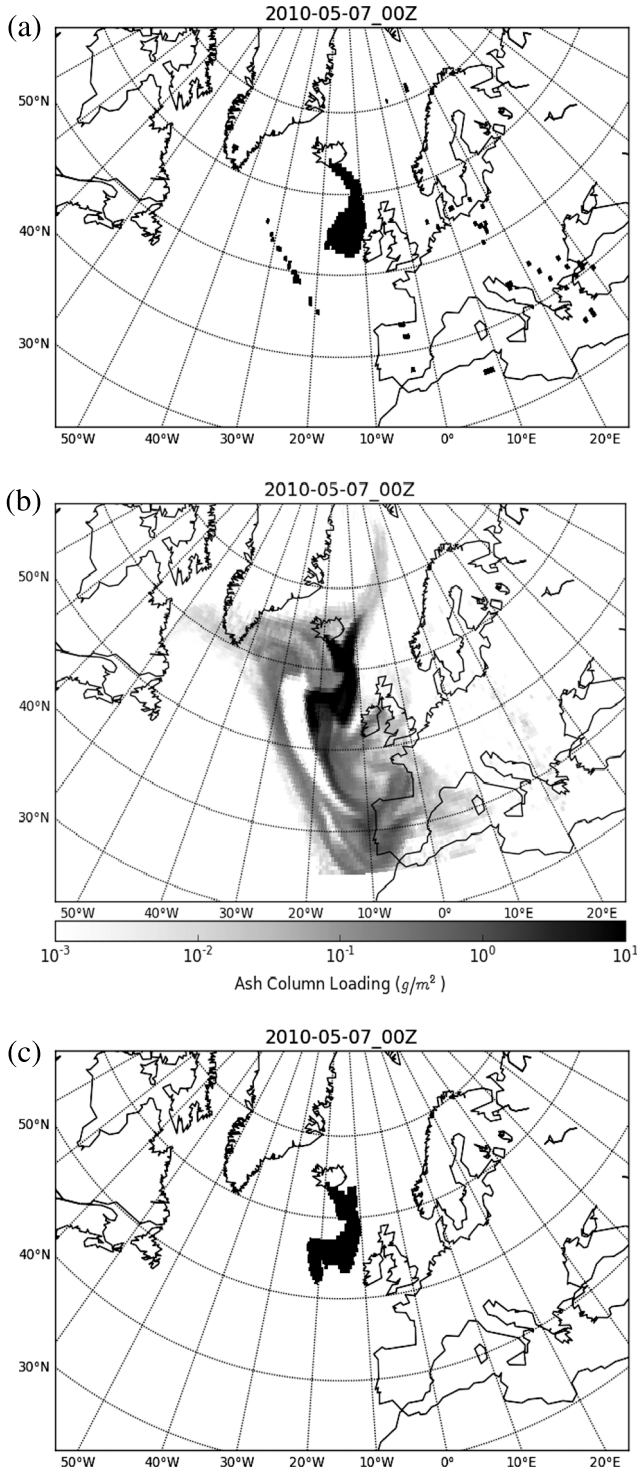

Figure 3. Ash column loading at 00 Z on 7 May 2010 (a) 5 h composite of satellite-detected ash clouds, (b) simulated by FLEXPART model, and (c) FLEXPART-simulated ash cloud after pixel matching (pm-FLEXPART).

time. The sharp boundaries of the ash cloud seen at 35°S, 30°E, and 60°W delineate the southern, eastern, and western boundaries of the model domain. A visual comparison of the satellite and FLEXPART ash clouds suggests that at this time, the satellite and model-simulated ash agree well in the location of maximum ash column loading although the modeled ash cloud is significantly more coherent and extensive than in the satellite ash cloud. This is a result of the minimum detection limit of the satellite [*Webley et al.*, 2012]. In order to perform a meaningful quantitative evaluation of the skill of the FLEXPART simulation, it is necessary to choose an ash column load threshold which represents the time-varying detection limit of the satellite. As shown by *Harvey and Dacre* [2015], a suitable method for evaluating volcanic ash model simulations is to select only the model pixels with the highest ash column loading such that the model ash cloud area is equal to that of the observed ash cloud area at each evaluation time. This method is known as pixel matching. Pixel matching removes bias from the simulation and is equivalent to using a time-varying percentile threshold. The fraction of the domain covered by satellite-retrieved ash varies between 1.2% and 6.7% during 7–16 May, giving a percentile threshold range of 93.3–098.9%.

Figure 3c shows the equivalent FLEXPART ash cloud as shown in Figure 3b after pixel matching (pm) has been applied as a binary ash/no-ash field. These images will be referred to as pm-FLEXPART images for the remainder of the paper. The FSS, calculated at 5.5° × 5.5° latitude/longitude resolution, at this time is 0.81 indicating that the FLEXPART simulation has considerable skill in capturing the satellite-detected spatial distribution of volcanic ash. This objective evaluation of the model skill agrees with our visual comparison of Figures 3a and 3c.

5.1. 8–10 May 2010

Figures 4a–4c show the ash cloud from the Eyjafjallajökull eruption, as detected by the SEVIRI instrument at 00 Z on 8, 9, and 10 May, respectively. An animation of the hourly satellite images shows a smooth evolution of the detected signal from which the following analysis is summarized (see supporting information for animation). At 00 Z on 8 May (Figure 4a) the ash was detected in a coherent plume extending southward from Iceland to the west of the UK, some of the ash particles were transported anticyclonically around the North Atlantic high-pressure center (Figure 2a), while a smaller fraction of the ash was advected cyclonically around a weak low over southern France toward the Mediterranean. There is also a coherent signal detected over the central Mediterranean at this time. Twenty-four hours later (Figure 4b), some of the ash has recirculated around the high-pressure center and is detected to the west of Iceland, while the ash traveling toward the Mediterranean is no longer detected. The ash in the central Mediterranean 24 h earlier is now detected over the eastern Mediterranean. Finally, by 00 Z on 10 May, (Figure 4c) the ash previously to the west of Iceland begins a secondary circulation around the high-pressure center and is advected southward. The coherent patch of ash in the Mediterranean is no longer detected in the domain.

Figures 4g–4i show the FLEXPART-simulated ash cloud after pixel matching has been applied (pm-FLEXPART). A visual comparison of the satellite and pm-FLEXPART images suggests that as for 7 May, the satellite and model-simulated ash spatial distributions agree well during the period 8–10 May 2010. The pixel matching identifies the main part of the FLEXPART ash cloud in the North Atlantic but does not select the ash over the Mediterranean as the ash column loading values are lower than in the North Atlantic (Figures 4d–4f). Despite this the FSS at 00 Z on 8, 9, and 10 is 0.81, 0.81, and 0.79, respectively, demonstrating that the model has considerable skill during this period.

5.2. 11–13 May 2010

Figures 5a–5c show the ash as detected by the SEVIRI instrument at 00 Z on 11, 12, and 13 May 2010. At 00 Z on 11 May (Figure 5a) ash is detected over a large part of the North Atlantic as it is advected anticyclonically around the high-pressure system in the North Atlantic. By 00 Z on 12 May (Figure 5b) the fragmented patches of ash detected south of 45°N the previous day (labeled 1 in Figure 5a) have been advected south out of the domain and the coherent ash between 20–40°W and 45–60°N (labeled 2 in Figures 5a and 5b) is now south of 50°N. The narrow band of ash extending southward from Iceland (labeled 3 in Figures 5a and 5b) is now separated from Iceland and is only detected south of 60°N. This narrow plume continues to travel south and by 00 Z on 13 May (Figure 5c) is detected below 45°N. At this time a coherent ash cloud (labeled 4 in Figure 5c) is also detected extending southeast from Iceland and then turning cyclonically around the developing low-pressure system which moves from west to east across Iceland on 12 and 13 May.

A visual comparison of the satellite and corresponding FLEXPART images (Figures 5d–5f) suggests that the model skill decreases during this period. Figures 5g–5i show the FLEXPART ash column loading after pixel

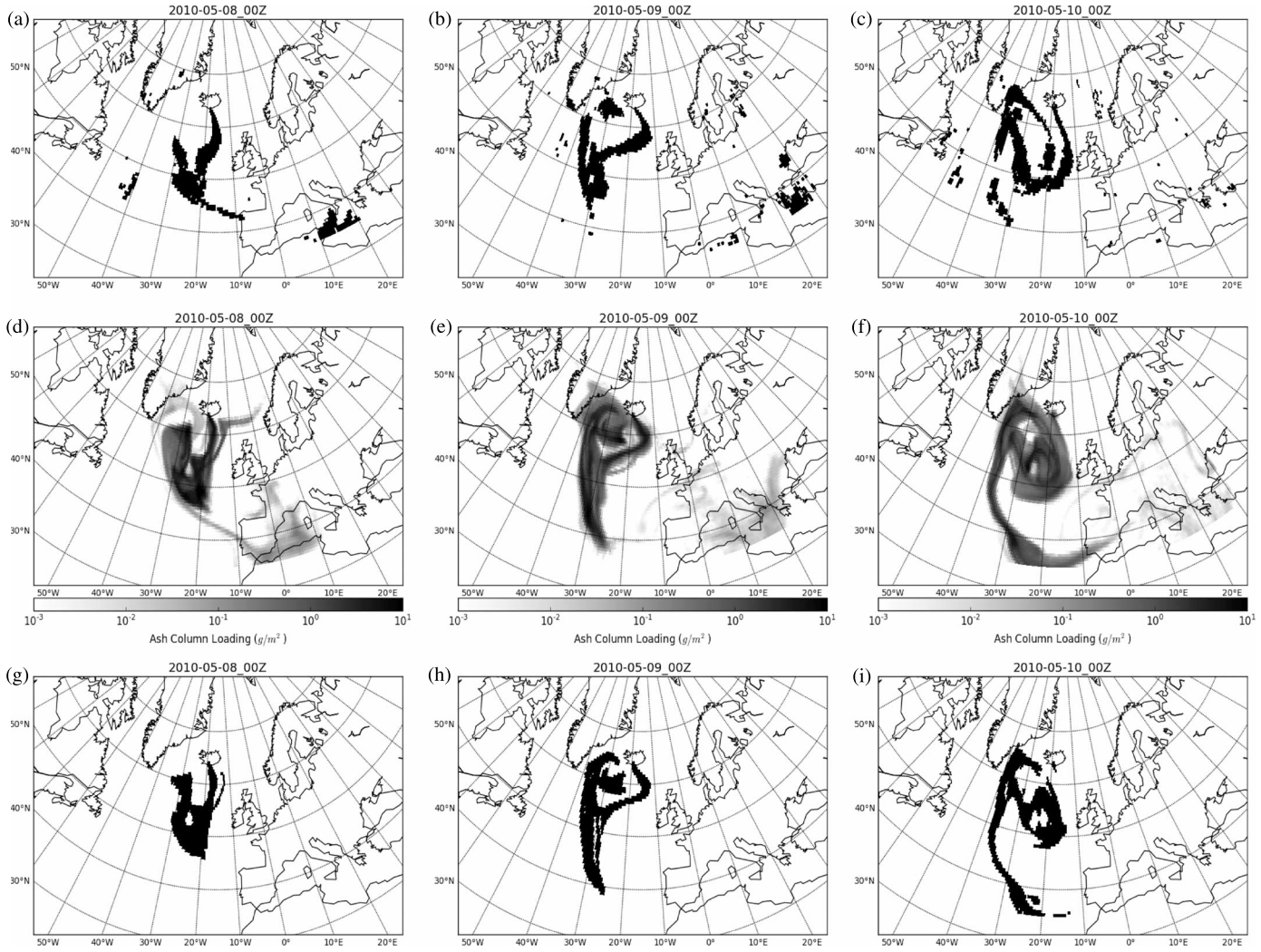

Figure 4. (a–c) The 5 h composite of satellite-detected ash clouds, (d–f) FLEXPART-simulated ash clouds, (g–i) FLEXPART-simulated ash clouds after pixel matching (pm-FLEXPART). Figures 4a, 4d, and 4g 00 Z on 8 May 2010. Figures 4b, 4e, and 4h 00 Z on 9 May 2010. Figures 4c, 4f, and 4i 00 Z on 10 May 2010.

matching has been applied (pm-FLEXPART). At 00 Z on 11 May, the pm-FLEXPART image contains a similar ash distribution to the satellite image resulting in a FSS of 0.7. However, small differences can be identified between the pm-FLEXPART and SEVIRI images. For example, we find that (i) the ash identified in region 1 is less patchy in the pm-FLEXPART image than that seen in the satellite-detected image and it is located farther west; (ii) the ash identified in region 2 is farther north in the pm-FLEXPART image than in the satellite image; and (iii) the ash identified in region 3 does not extend as far south in the pm-FLEXPART image as in the satellite image.

By 00 Z on 12 May, the differences between the pm-FLEXPART and satellite images have increased resulting in a FSS of 0.3. This low FSS is because (i) the ash identified in region 1 has not been transported south out of the domain, as seen in the satellite animation, but remains as a coherent structure between 35 and 43°N in the pm-FLEXPART image; (ii) the ash identified in region 2 is now 10° farther north than that identified in the satellite image; and (iii) the ash identified in region 3 is both less extensive and too far north when compared to the satellite image.

Finally, by 00 Z on 13 May we find that in the pm-FLEXPART image, rather than the ash in regions 1 and 2 being advected south out of the domain, they have combined and are recirculating anticyclonically around the high-pressure system in the North Atlantic. The ash in region 3 is missing entirely from the pm-FLEXPART image, and the ash in region 4 is well located but does not extend in a cyclonic direction around the

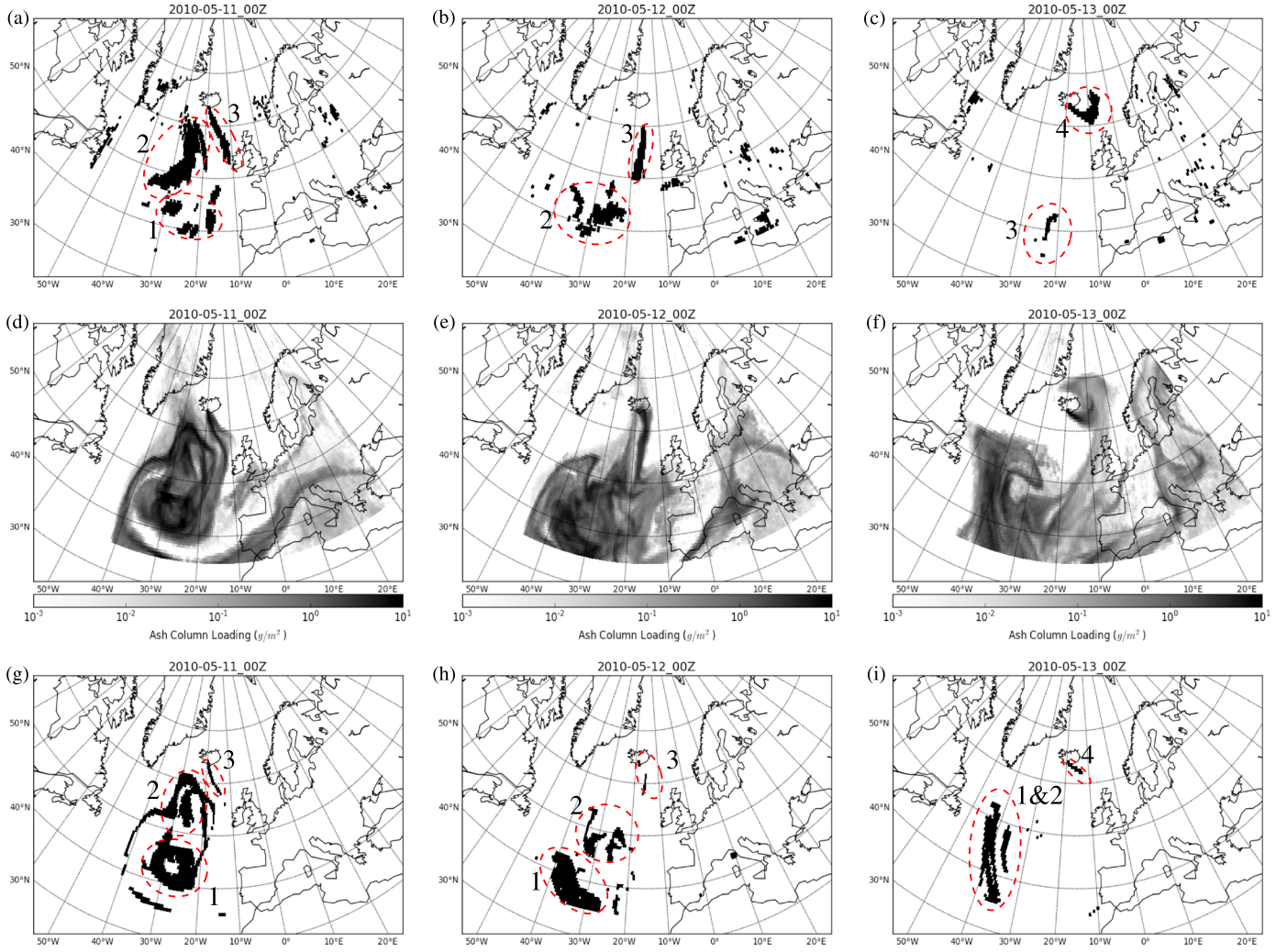

Figure 5. (a–c) The 5 h composite of satellite-detected ash clouds, (d–f) FLEXPART-simulated ash clouds, and (g–i) FLEXPART-simulated ash clouds after pixel matching (pm-FLEXPART). Figures 5a, 5d, and 5g 00 Z on 11 May 2010. Figures 5b, 5e, and 5h 00 Z on 12 May 2010. Figures 5c, 5f, and 5i 00 Z on 13 May 2010.

low-pressure system centered over Iceland. All of these discrepancies result in a very low FSS of 0.1 reflecting the very large differences between the satellite and pm-FLEXPART images.

The peak concentration of simulated ash in region 1 at 00Z on 11 May 2010 is located between 2 and 4 km above sea level. Figure 6 shows the GFS 700 hPa geopotential height at this time. As the ash is transported anticyclonically around the ridge a mismatch develops in the location of the ash in the satellite and pm-FLEXPART images. In particular, the ash located between 38 and 48°N is located farther west in the pm-FLEXPART image than in the satellite image. The colocation of the simulated ash with a bifurcation point in the atmosphere results in a modeled ash evolution which is very different from the satellite ash evolution. We can detect this 2-D horizontal flow separation by calculating the velocity gradient perpendicular to the flow

$$\frac{\partial \mathbf{v}}{\partial n} = \frac{1}{q^2}\left[v^2\frac{\partial u}{\partial x} - uv\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right) + u^2\frac{\partial v}{\partial y}\right] \tag{4}$$

where **v** is the velocity vector, q is the wind speed, n is distance in the direction perpendicular to the flow, and x and y are distances in longitude and latitude directions, respectively. Where this diagnostic is positive, the atmospheric flow separates, and where it is negative, the flow contracts. Figure 6 shows the positive part of this field at 00 Z on 11 May 2010. The black rectangle highlights the region in which ash in the FLEXPART simulation starts to deviate from the SEVIRI-detected ash position and is associated with large vertically integrated (800–500 hPa) flow separation. As a result of a small location error, the FLEXPART-simulated ash continues to travel anticyclonically around the high-pressure center, while the satellite-observed ash travels southward.

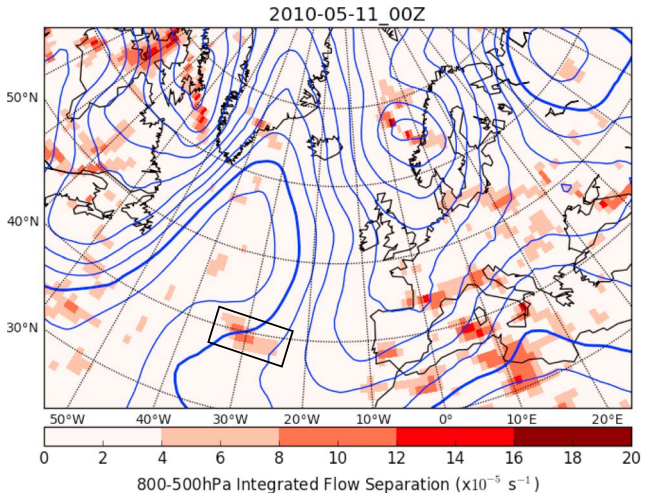

Figure 6. GFS analysis at 00 Z on 11 May 2010. The 700 hPa geopotential height contours (every 40 m, 3210 m contour thick) overlaid on 800–500 hPa vertically integrated 2-D flow separation. Black rectangle outlines the region of high flow separation corresponding to region 1 ash location in Figures 5a and 5d.

6. How Accurate are the Volcanic Ash Cloud Simulations?

Figure 7a shows the entire time evolution of the FSS at hourly intervals between 00 Z on 7 May and 00 Z on 16 May 2010 for three different neighborhood sizes. We define a skillful simulation as having a FSS > 0.5. The pm-FLEXPART and satellite ash cloud images are closely matched during the period 7–10 May 2010. FLEXPART can simulate the spatial distribution of the observed ash to within 0.5° × 0.5° latitude/longitude during this period. However, on 10 May there is a decrease in the spatial accuracy of the model. By 00 Z on 11 May FLEXPART can only skillfully simulate the presence of observed ash to within 2.5° × 2.5° latitude/longitude. The skill of the model continues to decrease, and between 12 Z on 11 May and 18 Z on 14 May the spatial accuracy of the model is <5.5° × 5.5° latitude/longitude and is therefore considered unusable for making decisions to close airspace.

The decrease in skill is followed by a period of gradual improvement in the match between FLEXPART and satellite ash locations over a period of 72 h. The increase in skill, starting from 18 Z on 12 May, is caused by increasing mass eruption rate from 18 Z on 12 May to 12 Z on 14 May (Figure 1). This results in ash column loading values that are higher close to the volcano than the dispersed ash in the North Atlantic. Therefore, the misplaced ash identified by the pixel matching method between 11 and 13 May is gradually replaced by correctly located ash close to the volcano, hence the increase in FSS (see supporting information Figure S3).

7. Why Does the Accuracy of the Volcanic Ash Cloud Simulations Decrease?

The mismatch in the locations of the pm-FLEXPART and satellite ash cloud images which occur between 10 and 13 May could result from (i) errors in the eruption source parameters, (ii) satellite detection limits, (iii) errors in the input meteorology, or (iv) a combination of the above factors. In this section we use the FSS to diagnose the most important causes of the model and satellite discrepancies for the period studied.

7.1. Eruption Source Parameters

Errors in the eruption source parameters (plume rise height, ash emission profile, mass eruption rate, and particle size distribution) could lead to a mismatch in the location of the modeled and satellite-observed ash. The particle size distribution specifies the mass fraction of ash with different diameters. In the control simulation the ash particles are assumed to have a diameter of 2 μm. As such, their sedimentation rates are negligible and they effectively act as passive tracers transported by the 3-D winds. If the ash particles emitted by the volcano have larger diameters, they will fall faster and thus potentially change the evolution of the ash cloud as they experience different wind speeds/directions or are deposited to the surface. Figure 7b shows the FSS for FLEXPART simulations in which a range of ash particle sizes from 0.3 to 30 μm is used, corresponding to the best match to in situ observations for the Eyjafjallajökull eruption [*Dacre et al.*, 2013]. In this simulation therefore, lighter and heavier particles have the potential to be transported in different directions due

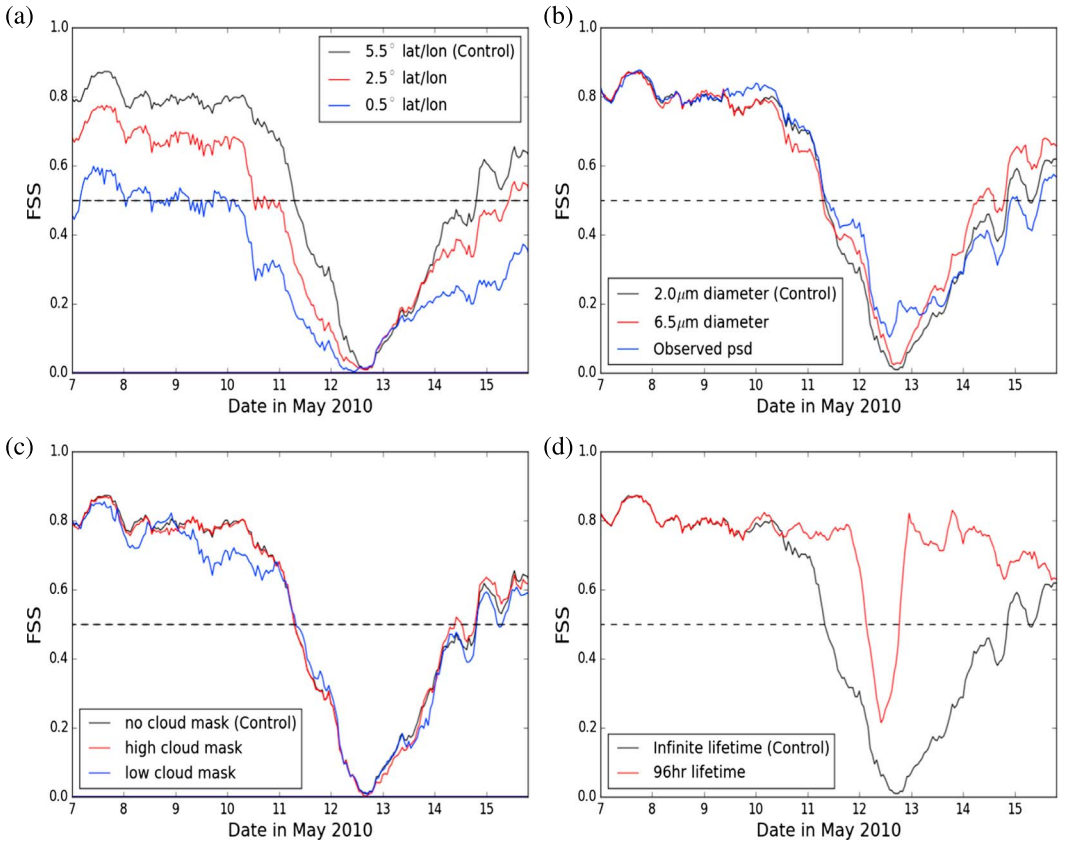

Figure 7. Evolution of the Fractions Skill Score (FSS) during the period 00 Z 7 May to 00 Z 16 May 2010. (a) Varying neighborhood size, (b) varying particle size, (c) varying cloud mask, and (d) varying ash residence time (lifetime). The dashed line represents a FSS of 0.5, the level at which the model is considered skillful.

to vertical wind shear. However, the simulation shows a very similar FSS evolution to the control simulation (as most of the ash has a diameter of <10 μm and thus low sedimentation velocities) suggesting that the particle size distribution is not responsible for the decrease in skill.

Errors in the plume rise height can also result in an incorrect plume transport direction due to the effects of vertical wind shear and will cause an underestimation/overestimation of the mass eruption rate (due to the fact that the mass eruption rate is calculated from the plume rise height equation specified in *Mastin et al.* [2009]). As the pixels are ranked according to their column loading prior to pixel matching at each time, pixels containing ash released during any period of underestimation/overestimation of the mass eruption rate will be ranked lower/higher as a result of the time-varying nature of the emission. Therefore, the pixel matching will be affected by any underestimation or overestimation of the plume rise height. Simulations using modified plume rise height evolution input prior to the period of reduced skill (red dash in Figure 1) and during the period of reduced skill (blue dash in Figure 1) were performed. These modifications to the plume rise height evolution were chosen as those most likely to affect the pixels leading to low FSS in the control simulation, i.e., the mislocated pixels in the North Atlantic. However, the modified plume rise height simulations did not show any improvement in the FSS. This is likely to be because the period studied is dominated by a high-pressure synoptic situation during which vertical wind shear is relatively small. Similarly, the FSS does not show any sensitivity to the assumed ash emission profile which represents the vertical distribution of ash released in the column above the volcano vent (not shown).

7.2. Satellite Detection Limits

The strength of the volcanic ash signal detected by the satellite is dependent on a variety of ash particle and atmospheric variables. For example, the presence of high-level water and ice clouds can obscure the ash cloud [*Rose et al.*, 2000]. To determine the effect of water and ice clouds, a high cloud mask was created using hourly cloud top height data from the Meteosat Second Generation (MSG) satellite. Locations in which the

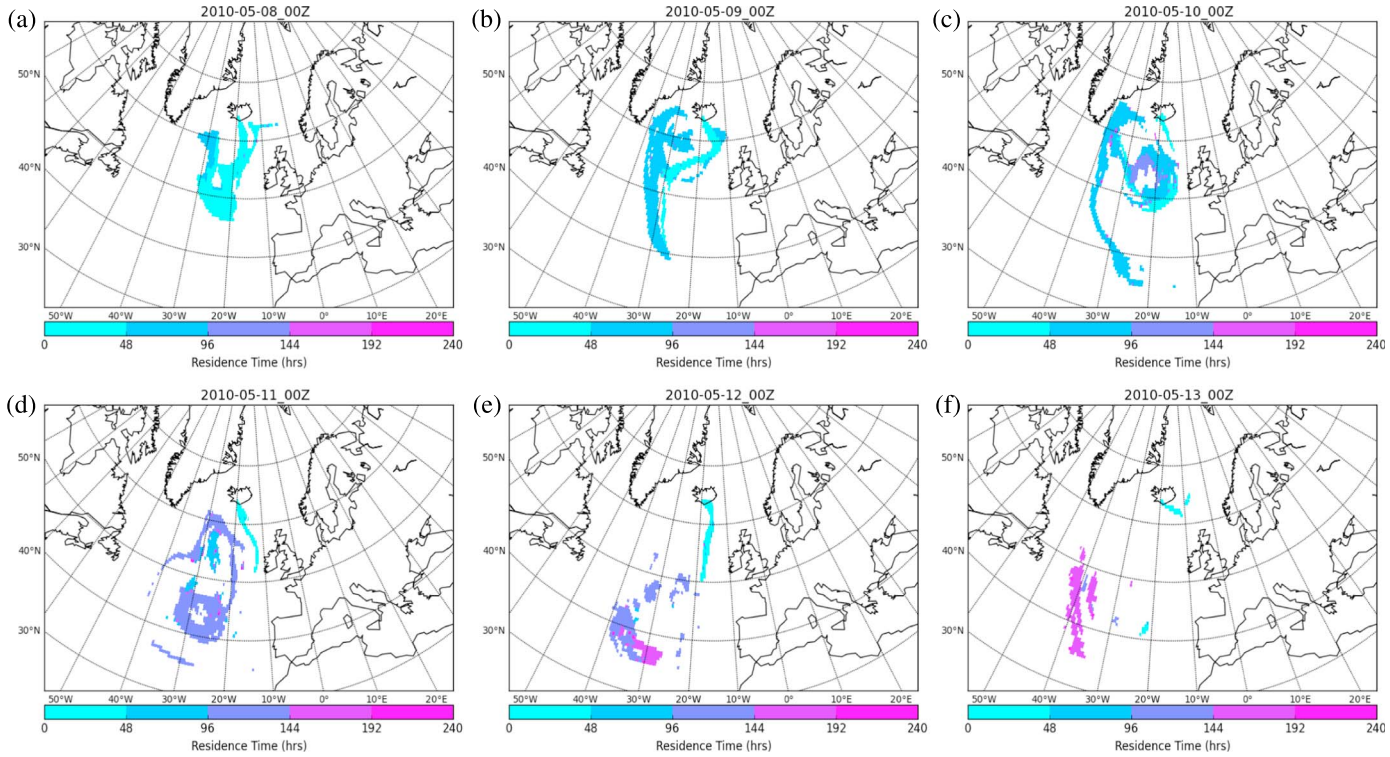

Figure 8. FLEXPART-simulated average ash residence time at 00 Z on (a–f) 8–13 May, respectively.

meteorological cloud top height was greater than the height of maximum ash concentration in the FLEXPART simulation at the equivalent time were included in the mask and removed from the FLEXPART ash cloud. This resulted in the removal of between 12.3% and 22.8% of the FLEXPART ash containing pixels, depending on the amount of high cloud at each hour. Following the application of the high cloud mask, pixel matching was applied and the FSS calculated for the new pm-FLEXPART image. Removing the pixels which were obscured by clouds made virtually no difference to the FSS (Figure 7c).

The strength of the volcanic ash signal detected by the satellite can also be dependent on the thermal contrast between the ash cloud and the underlying surface. As shown in Figure 2b there are extensive regions of low-level meteorological clouds colocated with the high-pressure center. Locations in which the meteorological cloud top height was 1 km lower or less than the height of maximum ash concentration in the FLEXPART simulation at the equivalent time were included in the low cloud mask and removed from the FLEXPART ash cloud. This resulted in the removal of between 2.0% and 9.1% of the FLEXPART ash containing pixels, depending on the amount of low cloud at each hour. Following the application of the low cloud mask, pixel matching was applied and the FSS calculated for the new pm-FLEXPART image. Removing the pixels in which the thermal contrast between the ash cloud and the underlying cloud top surface was likely to be low also did not improve the FSS (Figure 7c). The effect of the high and low cloud masks is not significant as the percentage of pixels removed by the cloud masks during this period is small compared to the percentage removed due to pixel matching. Thus, the satellite detection limits due to the presence of water and ice in the atmosphere are unlikely to be responsible for the decrease in model skill.

7.3. Input Meteorology

As the model skill decreases gradually over a 3 day period it is possible that small errors in the wind fields may be responsible for the increasing location errors in the pm-FLEXPART images. Figures 5d–5f suggest that ash in regions 1 and 2 is responsible for the mismatch in ash locations. Figure 8 shows the residence time of the ash particles in the pm-FLEXPART image, where the ash residence time is the length of time the ash has been in the atmosphere since it was released from the volcano. At 00 Z on 8 May the ash residence time is less than 48 h, but as it travels around the high-pressure system it ages and by 00 Z on 11 May 2010 ash west of 30°W has an average residence time of more than 96 h. To investigate the effect of accumulating errors in the wind fields on the FSS, we remove ash particles from the simulation when they have reached specified residence

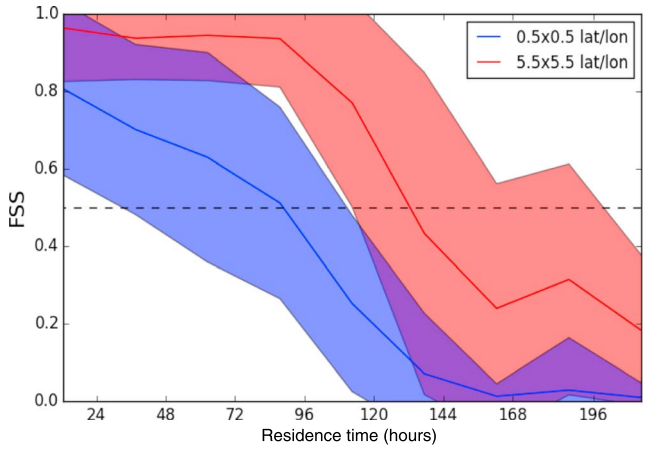

Figure 9. FSS as a function of ash residence time calculated for a neighborhood size of 0.5° × 0.5° latitude/longitude (blue) and 5.5° × 5.5° latitude/longitude (red). The shading represents ±1 standard deviation away from the mean (solid line).

times. Figure 7d shows the FSS when ash particles with residence times greater than 96 h are removed. In this simulation the FSS remains high (> 0.5) until 3 Z on 12 May, 28 h later than the control simulation. Furthermore, the period over which FLEXPART is unskillful (FSS < 0.5) is much shorter (15 h compared to 84 h for the control simulation). This suggests that it is indeed the ash in regions 1 and 2 that is responsible for the decrease in FSS.

To further investigate the effect of ash residence time on the FSS, we calculate the FSS separately for ash of different residence times. As the number of pixels in the verification area varies, depending on the residence time, the modeled and observed fractions at each point (M_j and O_j, respectively, in equation (2)) are treated as binary values instead of fractions. M_j is set to 1 if the fraction of ash in the neighborhood is > 0, and O_j is set to 1 if the fraction of observed ash in the neighborhood is > 0.

Figure 9 shows the average FSS as a function of ash residence time for two neighborhood sizes. For the smallest neighborhood size (0.5° × 0.5° latitude/longitude) there is a steady decrease in the FSS with increasing residence time, for the first 96 h, consistent with the hypothesis that small errors in the wind fields accumulate over time leading to increased differences in the location of the ash and hence to a decrease in FSS. The average FSS for residence times of less than 24 h is 0.8, indicating that FLEXPART is able to predict the presence of ash accurately for newly emitted ash. As the residence time of the ash increases, the mean FSS decreases and reaches a value of 0.5 for ash with residence times between 72 and 96 h. This shows that FLEXPART is able to predict the presence of ash in a 0.5° × 0.5° latitude/longitude area with skill for ash with residence times up to 4 days. However, it should be noted that the standard deviation on the FSS is large. At some times FLEXPART is skillful at this scale only for residence times up to 2 days, while at others it is skillful for residence times up to 5 days. Increasing the neighborhood size increases the FSS as it is easier to predict the presence of ash in a 5.5° × 5.5° latitude/longitude area than it is within a smaller area. For both neighborhood sizes there is a rapid decrease in FSS for residence times > 92 h. This decrease in skill is linked to the key bifurcation point in the ash trajectories described in section 7.3. These results are consistent with other studies that have shown that the accuracy of meteorological data can play a large role in the accuracy of dispersion model output [*Stohl et al.*, 1998; *Dacre*, 2010; *Arnold et al.*, 2015].

8. Discussion and Conclusions

In this paper we evaluate the spatial and temporal accuracy of ash cloud simulations performed using the FLEXPART volcanic ash transport and dispersion model. FLEXPART-simulated ash clouds are compared to satellite retrievals of ash using a spatial verification method known as the Fractions Skill Score (FSS) which varies from 0 (no match) to 1 (perfect match). The period of simulation (7–16 May 2010) can be split into two periods.

1. *7–10 May 2010.* FLEXPART compares well to the satellite ash cloud capturing the anticyclonic transport of ash around a high-pressure system located in the North Atlantic. This good agreement is captured by the spatial verification metric resulting in an average FSS of 0.75 (evaluated at 5.5° × 5.5° latitude/longitude).

2. *11-16 May.* FLEXPART and satellite ash locations begin to diverge resulting in a sharp drop in FSS from 0.75 to 0 over a period of 48 h. This decrease in skill is followed by a period of gradual improvement in the match between FLEXPART and satellite ash locations over a period of 72 h.

We investigated the potential reasons leading to this loss of skill in the FLEXPART simulations. We find that errors in the eruption source parameters (plume rise height, ash emission profile, and particle size distribution) cannot explain the reduction in skill. We also find that potential satellite detection limits (due to the presence of high-level meteorological cloud obscuring the ash cloud or the lack of thermal contrast between the ash cloud and the underlying surface) were also unable to explain the reduction in skill. Finally, we investigated potential errors in the input wind fields. We find that there is a decrease in the FSS as the ash residence time increases. This suggests that small errors in the wind fields result in location errors in the simulated ash which accumulate with time and are thus responsible for the decrease in model skill.

For this period (7–16 May 2010), FLEXPART has skill in predicting the location of ash within a 0.5° × 0.5° latitude/longitude area up to an ash residence time of 3 days (±1 day). For ash residence times of 3–5 days the spatial accuracy of the FLEXPART simulations decreases to within a 5.5° × 5.5° latitude/longitude area and beyond residence times of 5 days there is little correspondence between the FLEXPART ash and the satellite retrieved ash. These simulations are performed using analyzed wind fields, and thus, they provide an upper limit on the spatial and temporal predictability of the model during this period.

The rapid decrease in model skill is related to a bifurcation point which affects the predictability of the evolution of the atmosphere [*Lorenz*, 1963]. On 11 May there is ash very close to a bifurcation point, and as a result the FLEXPART simulation is able to produce an ash cloud evolution which is very different from the satellite ash evolution. This could be explored with ensemble modeling where many simulations are run with slightly different meteorological initial conditions. Large spread in the ensemble leads to low confidence in the forecasts and vice versa. It also implies that assimilation of satellite ash data, such as that described in *Wilkins et al.* [2016], could lead to an improvement in the ash location predictability.

Acknowledgments
We are grateful to Pete Francis and Mike Cooke at the UK Met Office for making their satellite-derived volcanic ash data available. We thank Nigel Roberts from the UK Met Office for helping us calculate the Fractions Skill Score. We also thank Alan Grant, Peter Clark, and Jonathan Dehn for useful discussions on this work. This work was supported by a US-UK Lloyds of London Fulbright Scholar Award. Natalie Harvey gratefully acknowledges funding from NERC grant NE/J017221/1 Probability, Uncertainty, and Risk in the Environment. The GFS analysis fields are available from ftp://nomads.ncdc.noaa.gov. The FLEXPART model can be downloaded from https://flexpart.eu/. The SEVIRI satellite data and FLEXPART simulation output are available from the authors (h.f.dacre@reading.ac.uk).

References

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

Arnold, D., C. Maurer, G. Wotawa, R. Draxler, K. Saito, and P. Seibert (2015), *J. Environ. Radioact.*, *139*, 212–225, doi:10.1016/j.jenvrad.2014.02.01.

Dacre, H. F. (2010), Evaluating the ability of a numerical weather prediction model to forecast tracer concentrations during ETEX 2, *44*, *1-10*, 294–303.

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. Atmos.*, *116*, D00U03, doi:10.1029/2011JD015608.

Dacre, H. F., A. L. M. Grant, and B. T. Johnson (2013), Aircraft observations and model simulations of particle size distribution in the Eyjafjallajökull volcanic ash cloud, *Atmos. Chem. Phys.*, *13*, 1277–1291.

Dacre, H. F., A. L. M. Grant, N. J. Harvey, D. J. Thomson, H. N. Webster, and F. Marenco (2015), Volcanic ash layer depth: Processes and mechanisms, *Geophys. Res. Lett.*, *42*, 637–645, doi:10.1002/2014GL062454.

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.*, *48*, 152–164.

Eckhardt, S., A. J. Prata, P. Seibert, K. Stebel, and A. Stohl (2008), Estimation of the vertical profile of sulfur dioxide injection into the atmosphere by a volcanic eruption using satellite column measurements and inverse transport modeling, *Atmos. Chem. Phys.*, *8*, 3881–3897.

Fisher, R. A. (1936), The use of multiple measurements in taxonomic problems, *Ann. Eugenics*, *1*, 179–188, doi:10.1111/j.1469-1809.1936.tb02137.x.

Folch, A., A. Costa, and S. Basart (2012), Validation of the FALL3D ash dispersion model using observations of the 2010 Eyjafjallajökull volcanic ash clouds, *Atmos. Environ.*, *48*, 165–183.

Francis, P. N., M. C. Cooke, and R. W. Saunders (2012), Retrieval of physical properties of volcanic ash using Meteosat: A case study from the 2010 Eyjafjallajökull eruption, *J. Geophys. Res.*, *117*, D00U09, doi:10.1029/2011JD016788.

Grant, A. L. M., H. F. Dacre, D. J. Thomson, and F. Marenco (2012), Horizontal and vertical structure of the Eyjafjallajökull ash cloud over the UK: A comparison of airborne lidar observations and simulations, *Atmos. Chem. Phys.*, *12*, 10145–10159.

Harvey, N. J., and H. F. Dacre (2015), Spatial evaluation of volcanic ash forecasts using satellite observations, *Atmos. Chem. Phys. Discuss.*, *15*, 24,727–24,749, doi:10.5194/acpd-15-24727-2015.

Kristiansen, N. I., et al. (2012), Performance assessment of a volcanic ash transport model mini-ensemble used for inverse modeling of the 2010 Eyjafjallajökull eruption, *J. Geophys. Res.*, *117*, D00U11, doi:10.1029/2011JD016844.

Langmann, B., A. Folch, M. Hensch, and V. Matthias (2013), Volcanic ash over Europe during the eruption of Eyjafjallajökull on Iceland, April–May 2010, *Atmos. Environ.*, *48*, 1–8, doi:10.1016/j.atmosenv.2011.03.054.

Lorenz, E. N. (1963), Deterministic nonperiodic flow, *J. Atmos. Sci.*, *20*, 130–141.

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.*, *117*, D00U17, doi:10.1029/2011JD016770.

Mittermaier, M., and N. Roberts (2010), Intercomparison of spatial forecast verification methods: Identifying skillful spatial scales using the fractions skill score, *Weather Forecasting*, *25*, 343–354.

Naeslund, E., and L. Thaning (1991), On the settling velocity in a nonstationary atmosphere, *Aerosol Sci. Tech.*, *14*, 247–256.

Oxford Economics (2010), The economic impacts of air travel restrictions due to volcanic ash, *Rep. Airbus*., Oxford Economics, Oxford. [Available at http://www.oxfordeconomics.com.]

Prata, A. J., S. A. Carn, A. Stohl, and J. Kerkmann (2007), Long range transport and fate of a stratospheric volcanic cloud from Soufriere Hills volcano, Montserrat, *Atmos. Chem. Phys.*, *7*, 5093–5103.

Prata, A. J. (1989), Observations of volcanic ash clouds in the 10–12 μm window using AVHRR/2 data, *Int. J. Remote Sens.*, *10*, 751–761.

Roberts, N. (2008), Assessing the spatial and temporal variation in the skill of precipitation forecasts from an NWP model, *Meteorol. Appl.*, *15*, 163–169.

Roberts, N. M., and H. W. Lean (2008), Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events, *Mon. Weather Rev.*, *136*, 78–97.

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. A*, *358*, 1585–1606.

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.

Stohl, A., et al. (2005), Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2, *Atmos. Chem. Phys.*, *5*, 2461–2474.

Stohl, A., M. Hittenberger, and G. Wotawa (1998), Validation of the Lagrangian particle dispersion model FLEXPART against large-scale tracer experiment data, *Atmos. Environ.*, *32*, 4245–4264.

Webley, P. W., T. Steensen, M. Stuefer, G. Grell, S. Fareitas, and M. Pavolonis (2012), Analyzing the Eyjafjallajökull 2010 eruption using satellite remote sensing, lidar and WRF-Chem dispersion and tracking model, *J. Geophys. Res.*, *117*, D00U26, doi:10.1029/2011JD016817.

Webster, H. N., D. J. Thomson, B. T. Johnson, I. P. C. Heard, K. Turnbull, F. Marenco, and B. W. Golding (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.

Wilkins, K. L., I. M. Watson, N. I. Kristiansen, H. N. Webster, D. J. Thomson, H. F. Dacre, and A. J. Prata (2016), Using data insertion with the NAME model to simulate the 8 May 2010 Eyjafjallajökull volcanic ash cloud, *J. Geophys. Res.*, *121*, 306–323, doi:10.1002/2015JD023895.