



**Spatial evaluation of volcanic ash forecasts**

N. J. Harvey and  
H. F. Dacre

This discussion paper is/has been under review for the journal Atmospheric Chemistry and Physics (ACP). Please refer to the corresponding final paper in ACP if available.

# Spatial evaluation of volcanic ash forecasts using satellite observations

N. J. Harvey and H. F. Dacre

Department of Meteorology, University of Reading, Earley Gate, P.O. Box 243, Reading, RG6 6BB, UK

Received: 23 June 2015 – Accepted: 20 August 2015 – Published: 11 September 2015

Correspondence to: N. J. Harvey (n.j.harvey@reading.ac.uk)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Abstract

The decision to close airspace in the event of a volcanic eruption is based on hazard maps of predicted ash extent. These are produced using output from volcanic ash transport and dispersion (VATD) models. In this paper an objective metric to evaluate the spatial accuracy of VATD simulations relative to satellite retrievals of volcanic ash is presented. The metric is based on the fractions skill score (FSS). This measure of skill provides more information than traditional point-by-point metrics, such as success index and Pearson correlation coefficient, as it takes into the account spatial scale over which skill is being assessed. The FSS determines the scale over which a simulation has skill and can differentiate between a “near miss” and a forecast that is badly misplaced. The idealised scenarios presented show that even simulations with considerable displacement errors have useful skill when evaluated over neighbourhood scales of 200–700 km<sup>2</sup>. This method could be used to compare forecasts produced by different VATDs or using different model parameters, assess the impact of assimilating satellite retrieved ash data and evaluate VATD forecasts over a long time period.

## 1 Introduction

Volcanic ash provides a significant hazard to aircraft by reducing visibility and causing both temporary engine failure and permanent engine damage. The presence of ash disrupts air traffic and can result in large financial losses to the aviation industry. The 2010 eruption of Eyjafjallajökull disrupted European airspace for thirteen days, causing the cancelation of over 95 000 flights and an estimated global financial loss of USD 5 billion (Oxford-Economics, 2010).

In the event of an eruption, the decision to close airspace is based on information provided by one of the nine Volcanic Ash Advisory Centres (VAACs). The VAACs issue hazard maps of predicted ash cloud extent based on forecasts from Volcanic Ash Transport and Dispersion models (VATDs). After the large-scale disruption caused by

## Spatial evaluation of volcanic ash forecasts

N. J. Harvey and  
H. F. Dacre

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Spatial evaluation of volcanic ash forecasts

N. J. Harvey and  
H. F. Dacre

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



the 2010 Eyjafjallajökull eruption in Iceland, new guidelines were brought in by the UK Civil Aviation Authority requiring predictions of ash concentration values. A small number of studies have been performed to evaluate forecasts of ash concentration, however they almost exclusively use ground based measurements at point locations or data from short research flights (Dacre et al., 2011; Devenish et al., 2012; Folch et al., 2012; Grant et al., 2012; Kristiansen et al., 2012; Webster et al., 2012; Dacre et al., 2013) and thus can only evaluate the model at a limited number of locations.

Satellite observations of volcanic ash clouds are vital for tracking the transport of the erupted ash. The high temporal and spatial resolution of the data lends itself to data assimilation and model verification. Satellite imagery is an invaluable tool for forecasters and is used qualitatively by VAACs to give an indication of the accuracy of the location of the ash cloud predicted by VATDs. However, these comparisons are carried out manually and do not provide an objective measure of the skill of the VATD forecasts. Therefore it is not easily possible to compare the skill of forecasts made at different times or by different models, or to assess the impact of changing the value of a model input or parameterisation. The large spatial coverage of the satellite observations provides an opportunity to quantitatively evaluate forecasts over a large spatial scale.

The evaluation of a 2-D forecast field presents many challenges. Straightforward summary statistics, such as root-mean-square-error, and binary skill score measures based on hits, misses, false alarms and correct rejections which are used to evaluate forecast performance at a particular point are not always easy to interpret and can lead to an underestimation of forecast skill. For example, if a volcanic plume is forecast to have the perfect shape but is displaced due to small errors in wind speed, metrics that compare each point in space and time (known as point-by-point in this paper) would yield low values as the feature is not in the correct place at the correct time. This problem has given rise to a host of other techniques to evaluate model skill, each suitable for evaluating different aspects of the forecast (see Gilleland et al., 2010 for

a review of these techniques). In this paper the spatial accuracy of the VATD forecasts is being assessed and therefore a neighbourhood technique is used.

The perceived accuracy of any forecast depends on the scale over which it is being assessed (if a spatial tolerance is acceptable). For example, it is easier to predict the presence of ash in a large area than a small one. Previous studies using point locations and point-by-point metrics to evaluate forecasts of volcanic ash fail to recognise forecasts that contain useful information unless it is in exactly the right place and at the right time. Many forecasts do have valuable information about the ash cloud in spite of small positional errors. For example, Webster et al. (2012) found an increase in agreement between simulated and observed ash concentrations if a “buffer zone” accounting for positional errors in the simulated ash cloud was used. Similarly Dacre et al. (2011) showed that if a temporal error of 9 h (equating to approximately 100 km displacement in space) was taken into account then the simulated ash column loadings match well with lidar observations.

The aim of this paper is to develop an evaluation metric that can determine the spatial accuracy of volcanic ash forecasts. This metric utilises a neighbourhood-based measure of skill called the fractions skill score (FSS) (Roberts and Lean, 2008). This skill score was developed for the verification of precipitation forecasts produced by numerical weather prediction (NWP) models. This technique has been chosen as it relaxes the requirements for exact matching between forecasts and observations; the fractional coverage of simulated ash within an area needs to match the fractional coverage of the satellite retrieved ash to be counted as correct. It also provides users with information on the scale at which an acceptable level of skill is attained. To illustrate the use of this new technique VATD simulations made using the Numerical Atmospheric-dispersion Modelling Environment (NAME) (Jones et al., 2007) of the ash cloud from the 2010 Eyjafjallajökull eruption are evaluated against SEVIRI satellite observations made on 14 May 2010.

## Spatial evaluation of volcanic ash forecasts

N. J. Harvey and  
H. F. Dacre

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## 2 NAME simulations

NAME is the operational VATD used by the London VAAC. It is a Lagrangian particle dispersion model originally developed in response to the 1986 Chernobyl disaster. Particles, each representing a mass of volcanic ash, are released from a source. The particles are passively advected by 3-D wind fields provided by, in this case, the UK Met Office global NWP model analysis updated every 6 h and forecast fields updated every 3 h. The effect of turbulence is represented by stochastic perturbations to the particle trajectories based on semi-empirical turbulence profiles. NAME also includes parameterisations of sedimentation, dry deposition and wet deposition. The ash concentrations are calculated by summing the mass of particles in the model grid boxes and over 1 h. In this study the model grid boxes are  $0.375^\circ$  latitude by  $0.5625^\circ$  longitude (approximately  $40\text{ km} \times 40\text{ km}$ ).

To predict the transport and dispersion of ash, information about the volcanic eruption is required. These are known as eruption source parameters (ESPs) and include plume rise height, mass eruption rate, vertical profile of the plume, particle density and particle size distribution. In the simulations presented in this paper the plume height is based on observations by the Icelandic Meteorological Office's C-band radar (Arason et al., 2011) located at Keflavík International Airport. It is assumed that the ash was distributed uniformly throughout the height of the plume. The mass eruption rate is given by an empirical relationship based on the plume height given by Mastin et al. (2009). The ash density is assumed to be  $2500\text{ kg m}^{-3}$  and the particles are assumed to have a diameter of  $1\text{--}3\text{ }\mu\text{m}$ . The choice of model parameters used here are similar to those used in Grant et al. (2012) but the technique presented here could be applied to any VATD simulation.

ACPD

15, 24727–24749, 2015

### Spatial evaluation of volcanic ash forecasts

N. J. Harvey and  
H. F. Dacre

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



### 3 SEVIRI satellite observations

The Spinning Enhanced Visible and Infrared Imager (SEVIRI) is mounted on the geosynchronous Meteosat Second Generation (MSG) satellite. It has 12 spectral channels and provides high temporal (15 min) and spatial (3 km resolution at the equator) observations. The high temporal and spatial resolution makes these observations ideally suited to evaluating the transport of volcanic ash following an eruption.

The volcanic ash measurements used in this paper are retrieved using the algorithm of Francis et al. (2012) which utilises three long-wave window channels centred at 8.7, 10.8 and 12.0  $\mu\text{m}$  to discriminate between meteorological cloud and ash cloud. Where ash is detected this algorithm determines ash layer top pressure, ash column loading and ash effective radius. In this paper ash column loading is used to determine the horizontal accuracy of the simulated ash clouds. It is important to note that the detection of volcanic ash by satellite is dependent on the optical depth of the cloud and the physical properties of the ash. Optically thin ash clouds and ash particles smaller than 0.2  $\mu\text{m}$  may not be detected. Following this, the minimum detection limit of ash is considered to be in the range of 0.2–1.0  $\text{g m}^{-2}$  (Francis et al., 2012; Prata and Prata, 2012). Other factors, namely the thermal contrast between the ash and the underlying surface, satellite viewing angle, ash cloud height and the presence of other absorbers (e.g. water, ice and sulphur dioxide), also affect the detection and retrieval of ash properties (Millington et al., 2012). A case study comparison for 17 May 2010 between retrieved column loadings and airborne lidar data is presented in Francis et al. (2012). The mass column loading values are in reasonable agreement with maximum values of 0.7–0.8  $\text{g m}^{-2}$  in both data sets. The column loading values derived in Francis et al. (2012) are also qualitatively comparable to those presented in Thomas and Prata (2011). By applying their retrieval algorithm Dubuisson et al. (2014) found comparable values to Francis et al. (2012) for mean effective radius, plume height and mass loading for 6 May 2010.

For comparison with NAME the satellite retrieved column integrated loadings are averaged on to a regular  $0.375^\circ \times 0.5625^\circ$  grid and averaged over a period of 5 h centred

## Spatial evaluation of volcanic ash forecasts

N. J. Harvey and  
H. F. Dacre

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



on the verification time. This time averaging is used to smooth the SEVIRI ash observations which can be very patchy. The choice of a 5 h averaging time was based on the results of some simple data denial experiments. The results of these experiments can be found in Appendix A.

## 4 The evaluation method

There are many neighbourhood skill scores described in the literature (see Ebert, 2008; Gilleland et al., 2010 for an overview). The method used in this paper is based on the FSS developed by Roberts and Lean (2008) to test the skill of high resolution precipitation forecasts (e.g. Roberts, 2008; Mittermaier and Roberts, 2010) and is routinely computed for that purpose in the operational verification suite at the UK Met Office (Mittermaier et al., 2013). It compares fractional coverage in the forecast field with fractional coverage in the observational field for a specified precipitation threshold and over a range of neighbourhood sizes to determine the spatial scale over which a simulation can be considered skillful.

The evaluation is performed in two stages. First the simulation and satellite fractions (where fractions are the fractional coverage of a specified neighbourhood size in which pixels exceed a pre-defined threshold) are generated, then these fractions are compared using FSS. Here we focus on a case study day of 14 May 2010 during the Eyjafjallajökull eruption. Figure 1a shows the detected ash column loadings by SEVIRI at 00:00 UTC on 14 May. The ash cloud was detected in a coherent plume extending south-eastwards from Iceland to the northwest of the UK. There is also a small patch of ash detected north of Iceland. Figure 1b shows the corresponding NAME simulated ash column loading at the same time. A visual comparison of the satellite and NAME ash clouds suggests that at this time there is good agreement in the location of the maximum ash column loadings.

## Spatial evaluation of volcanic ash forecasts

N. J. Harvey and  
H. F. Dacre

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## 4.1 Stage 1: generating the fractional coverage

In general, NAME simulates a more extensive ash cloud structure than the satellite observations. This is largely due to the minimum detection limit of the satellite observations. Therefore, to perform a meaningful quantitative evaluation between the simulated and satellite retrieved ash cloud, a threshold must be applied to the NAME column loadings. In the case of precipitation forecasts a 95th percentile threshold is commonly used. This threshold selects the highest 5 % of radar and simulated precipitation accumulations in the domain independently. This is done to remove any bias in precipitation amounts when the focus is to look at the spatial accuracy of the forecast only. In the case of volcanic ash a fixed percentile threshold is not appropriate due to the artificial cut off in the distribution of retrieved ash column loadings due to the detection limit of the satellite. This cut off can be seen in Fig. 2a. Ash column loadings less than  $0.2 \text{ g m}^{-2}$  are not retrieved during the period 7–16 May 2010.

The satellite retrieved values of ash column loading often have large errors associated with them (P. N. Francis, personal communication, 2014). Therefore the values can be considered as a binary ash/no ash detection flag. The detection limit means that there are far more grid boxes populated with ash in the simulations than in the satellite observations. Therefore to ensure a fair comparison with the satellite the number of simulated ash grid boxes used in the comparison is restricted to match the number of grid boxes with observed ash (i.e. the area of ash cloud being compared in both the NAME simulation and satellite observations is the same at each evaluation time). For example, if there are 250 grid boxes with satellite retrieved ash then the 250 NAME grid boxes with the highest ash column loading are used in the comparison. This will be referred to as pixel matching in this paper and is equivalent to using a time varying percentile threshold (Fig. 2b). The fraction of the domain covered by satellite retrieved ash varies between 3.4 and 14.6 % giving a percentile threshold of 85.4–96.6 %. An example of how this pixel matching modifies the NAME ash distribution is shown in Fig. 1c. In this case the number of satellite pixels containing ash is 422, giving a per-

### Spatial evaluation of volcanic ash forecasts

N. J. Harvey and  
H. F. Dacre

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





centile threshold of 94.6% and a NAME concentration threshold of  $0.6 \text{ g m}^{-2}$  at this time (comparable to the stated minimum detection limit of Francis et al., 2012; Prata and Prata, 2012) when assuming a distal fine ash fraction (DFAF) of 3%. DFAF is the percentage of the ash vented from the volcano that undergoes long range transport (Dacre et al., 2011; Grant et al., 2012; Devenish et al., 2012). Note that the ash column loading threshold can vary from  $0.2\text{--}1.2 \text{ g m}^{-2}$  at this time when using other plausible DFAFs of 1 and 6% respectively (Fig. 2b).

The fraction of grid points containing ash for different sized square neighbourhoods centred on each gridbox are then calculated for both the pixel matched NAME data and satellite observations. In this paper neighbourhood sizes of  $40\text{--}1160 \text{ km}^2$  are considered.

## 4.2 Stage 2: computing the FSS

The FSS is calculated in the following way:

$$\text{FSS} = 1 - \frac{\text{FBS}}{\text{FBS}_{\text{ref}}} \quad (1)$$

(Roberts and Lean, 2008) where the Fractions Brier Score (FBS) is a variation on the Brier Score (Brier, 1950) in which both the simulated and observed probabilities (or fractions) can have any value between 0 and 1. FBS is given by:

$$\text{FBS} = \frac{1}{N} \sum_{j=1}^N (O_j - M_j)^2 \quad (2)$$

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

$$\text{FBS}_{\text{ref}} = \frac{1}{N} \left[ \sum_{j=1}^N O_j^2 + \sum_{j=1}^N M_j^2 \right]. \quad (3)$$

FBS<sub>ref</sub> is the largest FBS that could be obtained from the simulated and observed fraction which occurs when there is no collocation of non-zero fractions. A FSS of 1 indicates a perfect match between the modelled and observed fractions whilst a FSS of 0 indicates a complete mismatch. In general, a forecast with FSS > 0.5 is considered skillful (Roberts and Lean, 2008).

The FSS, calculated using a 40 km<sup>2</sup> neighbourhood (the grid scale), at 00:00 UTC on 14 May 2010 is 0.51 indicating that the NAME simulation has skill in capturing the satellite retrieved spatial distribution of volcanic ash at this scale. This objective measure agrees with the subjective visual comparison of Fig. 1a and c which shows fairly good spatial agreement in the location of the ash cloud at the 40 km<sup>2</sup> scale.

## 5 What if the simulated ash cloud is displaced from the satellite-retrieved ash cloud?

One vital input parameter for a VATD is the height of the plume. At the time of eruption this can be uncertain and can evolve throughout the eruption period. The use of an incorrect plume height could result in ash being transported in a different direction and at a different speed than it experiences in reality due to changes in windspeed and direction with height. In this section a set of idealised scenarios are presented where the NAME simulated ash plume is artificially stretched and squashed to represent the possible impact of an incorrect plume height. The transformations used are shown in Fig. 3 and are performed in the following way:

$$\text{new longitude} = s(\text{longitude} - E_{\text{lon}}) + E_{\text{lon}} \quad (4)$$

$$\text{new latitude} = (\text{latitude} - E_{\text{lat}})/s + E_{\text{lat}} \quad (5)$$

where  $s$  is a stretching factor and  $E_{\text{lat}}$  and  $E_{\text{lon}}$  are the latitude and longitude of Eyjafjalajökull. The NAME simulated ash cloud is interpolated on to this transformed grid. Note that the stretching transformation is applied to the NAME output before pixel matching

## Spatial evaluation of volcanic ash forecasts

N. J. Harvey and  
H. F. Dacre

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





## Spatial evaluation of volcanic ash forecasts

N. J. Harvey and  
H. F. Dacre

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



pare the output from two different VATDs with different eruption source parameters for the 1992 Mount Spur eruption. PCC is also known as the linear correlation coefficient. A simulation with a PCC value of 1 has complete correlation between the simulated and measured ash cloud. PCC is one of the measures calculated by Kristiansen et al. (2012) to evaluate and compare the skill of several different VATDs.

For all the skill metrics the highest values are for the simulation with no stretch. The simulation with stretch factor 1.2 has the next highest values of skill. In the case of no stretch and stretch factor 1.2 the FSS values are greater than the 0.5 threshold for skill, the PCC values fall within the bounds Kristiansen et al. (2012) consider skillful and the SI values are within the range Webley et al. (2009) found in their analysis of the impact of the vertical distribution of ash and ash particle size distribution. The SI and PCC for both stretch factor 0.5 and stretch factor 2 are very low and, by chance, equal, however by visual inspection the stretch factor 0.5 ash cloud appears to more closely match the satellite retrieved ash than the stretch factor 2 ash cloud. This is supported by the FSS score for the stretch factor 0.5 ash cloud having a higher FSS than the stretch factor 2 cloud at smaller spatial scales. This highlights the fact that point-by-point measures are unable to distinguish between a simulation that is a near-miss or a simulation that is completely wrong, although they do still pick out the “best” simulation in this instance.

## 6 Summary and conclusions

In this paper it has been shown that a neighbourhood-based metric fractions skill score (FSS) is suitable for evaluating simulations of volcanic ash clouds using satellite observations. This measure of skill provides more information than traditional point-by-point metrics, such as success index and Pearson correlation coefficient, as it takes into account spatial scale over which skill is being assessed and can be used to determine the spatial scale over which the VATD model should be believed. In the case study presented here (00:00 UTC, 14 May 2010) the NAME simulation had skill (FSS > 0.5) at neighbourhood scale of 40 km<sup>2</sup> (the grid resolution). Even simulations with consid-

erable displacement errors have skill when using larger neighbourhood sizes of 200–700 km<sup>2</sup>. The advantage of this kind of evaluation is that the objectively determined results for a set of idealised displacement scenarios are often much more similar to a subjective visual inspection of the simulations.

Although the evaluation in this paper has focussed on a single idealised set of scenarios the FSS method could, in principle, be used to evaluate forecasts over a longer period of time. It could also be used to compare forecasts with different ESPs or model parameters, or forecasts from an ensemble of simulations performed with different models, input meteorology and emissions, or assess the impact of assimilation of satellite data. This will be the focus of future studies. The assimilation could be for the ESPs (e.g. Stohl et al., 2011) or the distribution of ash downstream from the volcano (e.g. Wilkins et al., 2015). The methodology presented could also be extended to the distribution of sulphur dioxide following an eruption or to forecasts of other dispersion events, for example, after a nuclear incident or a forest fire.

## Appendix: SEVIRI retrieval smoothing time

This section describes the data denial experiments used to determine the SEVIRI smoothing time used in this study. In these experiments satellite retrieved ash column loadings at a verification time ( $t_0$ ) were considered the “truth” and compared using the root-mean-squared-error (RMSE) to satellite retrieved ash column loadings with 50% of the pixels randomly removed and replaced with a smoothed field using observations up to 8 h before and after  $t_0$ . This was done for each hour in the period 8–14 May 2010. This experiment was performed 50 times using different random sampling to assess the spread in the RMSE due to different areas in the plume being replaced.

Figure A1 shows the results of the data denial experiments. The symbols show the median RMSE value and the boxes indicate the interquartile range. There are several interesting points to note. Firstly, there is a large spread between different days. This is due to the time varying mass eruption rate of the volcano and changing meteorological

## Spatial evaluation of volcanic ash forecasts

N. J. Harvey and  
H. F. Dacre

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





**Spatial evaluation of volcanic ash forecasts**N. J. Harvey and  
H. F. Dacre

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Eyjaflajallajökull eruption using lidar observations and NAME simulations, *J. Geophys. Res.*, 116, D00U03, doi:10.1029/2011JD015608, 2011. 24729, 24730, 24735

Dacre, H. F., Grant, A. L. M., and Johnson, B. T.: Aircraft observations and model simulations of concentration and particle size distribution in the Eyjaflajallajökull volcanic ash cloud, *Atmos. Chem. Phys.*, 13, 1277–1291, doi:10.5194/acp-13-1277-2013, 2013. 24729

Devenish, B., Thomson, D., Marengo, F., Leadbetter, S., Ricketts, H., and Dacre, H.: A study of the arrival over the United Kingdom in April 2010 of the Eyjaflajallajökull ash cloud using ground-based lidar and numerical simulations, *Atmos. Environ.*, 48, 152–164, 2012. 24729, 24735

Dubuisson, P., Herbin, H., Minvielle, F., Compiègne, M., Thieuleux, F., Parol, F., and Pelon, J.: Remote sensing of volcanic ash plumes from thermal infrared: a case study analysis from SEVIRI, MODIS and IASI instruments, *Atmos. Meas. Tech.*, 7, 359–371, doi:10.5194/amt-7-359-2014, 2014. 24732

Ebert, E. E.: Fuzzy verification of high-resolution gridded forecasts: a review and proposed framework, *Meteorol. Appl.*, 15, 51–64, doi:10.1002/met.25, 2008. 24733

Folch, A., Costa, A., and Basart, S.: Validation of the FALL3D ash dispersion model using observations of the 2010 Eyjaflajallajökull volcanic ash clouds, *Atmos. Environ.*, 48, 165–183, 2012. 24729

Francis, P. N., Cooke, M. C., and Saunders, R. W.: Retrieval of physical properties of volcanic ash using Meteosat: a case study from the 2010 Eyjaflajallajökull eruption, *J. Geophys. Res.*, 117, D00U09, doi:10.1029/2011JD016788, 2012. 24732, 24735

Gilleland, E., Ahijevych, D., Brown, B. G., and Ebert, E. E.: Verifying forecasts spatially, *B. Am. Meteorol. Soc.*, 91, 1365–1373, 2010. 24729, 24733

Grant, A. L. M., Dacre, H. F., Thomson, D. J., and Marengo, F.: Horizontal and vertical structure of the Eyjaflajallajökull ash cloud over the UK: a comparison of airborne lidar observations and simulations, *Atmos. Chem. Phys.*, 12, 10145–10159, doi:10.5194/acp-12-10145-2012, 2012. 24729, 24731, 24735

Jones, A., Thomson, D., Hort, M., and Devenish, B.: The UK Met Office's next-generation atmospheric dispersion model, NAME III, in: *Air Pollution Modeling and its Application XVII*, Springer, 580–589, 2007. 24730

Kristiansen, N. I., Stohl, A., Prata, A. J., Bukowiecki, N., Dacre, H., Eckhardt, S., Henne, S., Hort, M. C., Johnson, B. T., Marengo, F., Neining, B., Reitebuch, O., Seibert, P., Thomson, D. J., Webster, H. N., and Weinzierl, B.: Performance assessment of a volcanic ash

**Spatial evaluation of volcanic ash forecasts**N. J. Harvey and  
H. F. Dacre

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



transport model mini-ensemble used for inverse modeling of the 2010 Eyjafjallajökull eruption, *J. Geophys. Res.*, 117, D00U11, doi:10.1029/2011JD016844, 2012. 24729, 24738

Mastin, L., Guffanti, M., Servranckx, R., Webley, P., Barsotti, S., Dean, K., Durant, A., Ewert, J., Neri, A., Rose, W., Schneider, D., Siebert, L., Stunder, B., Swanson, G., Tupper, A., Volentik, A., and Waythomas, C.: A multidisciplinary effort to assign realistic source parameters to models of volcanic ash-cloud transport and dispersion during eruptions, *J. Volcanol. Geoth. Res.*, 186, 10–21, 2009. 24731

Millington, S. C., Saunders, R. W., Francis, P. N., and Webster, H. N.: 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, 2012. 24732

Mittermaier, M. and Roberts, N. M.: Intercomparison of spatial forecast verification methods: identifying skillful spatial scales using the fractions skill score, *Weather Forecast.*, 25, 343–354, 2010. 24733

Mittermaier, M., Roberts, N., and Thompson, S. A.: A long-term assessment of precipitation forecast skill using the fractions skill score, *Meteorol. Appl.*, 20, 176–186, doi:10.1002/met.296, 2013. 24733

Oxford-Economics: The Economic Impact of Air Travel Restrictions Due to Volcanic Ash, Abbey House, Oxford, UK, 2010. 24728

Prata, A. and Prata, A.: Eyjafjallajökull volcanic ash concentrations determined using spin enhanced visible and infrared imager measurements, *J. Geophys. Res.-Atmos.*, 117, D00U23, doi:10.1029/2011JD016800, 2012. 24732, 24735

Roberts, N. M.: Assessing the spatial and temporal variation in the skill of precipitation forecasts from an NWP model, *Meteorol. Appl.*, 15, 163–169, 2008. 24733

Roberts, N. M. and Lean, H. W.: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events, *Mon. Weather Rev.*, 136, 78–97, 2008. 24730, 24733, 24735, 24736

Stohl, A., Prata, A. J., Eckhardt, S., Clarisse, L., Durant, A., Henne, S., Kristiansen, N. I., Minikin, A., Schumann, U., Seibert, P., Stebel, K., Thomas, H. E., Thorsteinsson, T., Tørseth, K., and Weinzierl, B.: 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, doi:10.5194/acp-11-4333-2011, 2011. 24739



## Spatial evaluation of volcanic ash forecasts

N. J. Harvey and  
H. F. Dacre

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Stunder, B. J., Heffter, J. L., and Draxler, R. R.: Airborne volcanic ash forecast area reliability, *Weather Forecast.*, 22, 1132–1139, 2007. 24737

Thomas, H. E. and Prata, A. J.: Sulphur dioxide as a volcanic ash proxy during the April–May 2010 eruption of Eyjafjallajökull Volcano, Iceland, *Atmos. Chem. Phys.*, 11, 6871–6880, doi:10.5194/acp-11-6871-2011, 2011. 24732

Webley, P., Stunder, B., and Dean, K.: 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. Geoth. Res.*, 186, 108–119, 2009. 24737, 24738

Webster, H. N., Thomson, D. J., Johnson, B. T., Heard, I. P. C., Turnbull, K., Marenco, F., Kristiansen, N. I., Dorsey, J., Minikin, A., Weinzierl, B., Schumann, U., Sparks, R. S. J., Loughlin, S. C., Hort, M. C., Leadbetter, S. J., Devenish, B. J., Manning, A. J., Witham, C. S., Haywood, J. M., and Golding, B. W.: Operational prediction of ash concentrations in the distal volcanic cloud from the 2010 Eyjafjallajökull eruption, *J. Geophys. Res.*, 117, doi:10.1029/2011JD016790, 2012. 24729, 24730

Wilkins, K., Mackie, S., Watson, M., Webster, H., Thomson, D., and Dacre, H.: Data insertion in volcanic ash cloud forecasting, *Ann. Geophys.-Italy*, 57, doi:10.4401/ag-6624, 2015. 24739

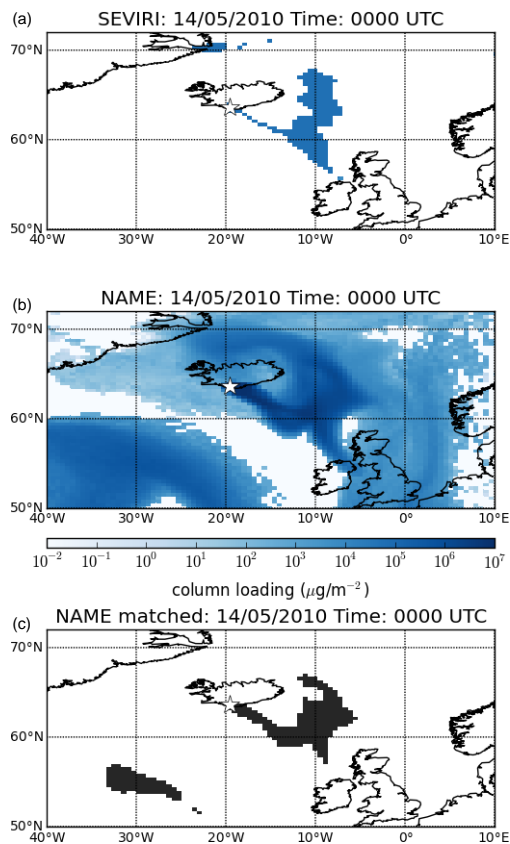
## Spatial evaluation of volcanic ash forecasts

N. J. Harvey and  
H. F. Dacre

**Table 1.** The value of success index (SI), Pearson correlation coefficient (PCC), FSS for a neighbourhood of 600 km<sup>2</sup> and the scale at which the FSS reaches a value of 0.5 for the scenarios presented in Fig. 3.

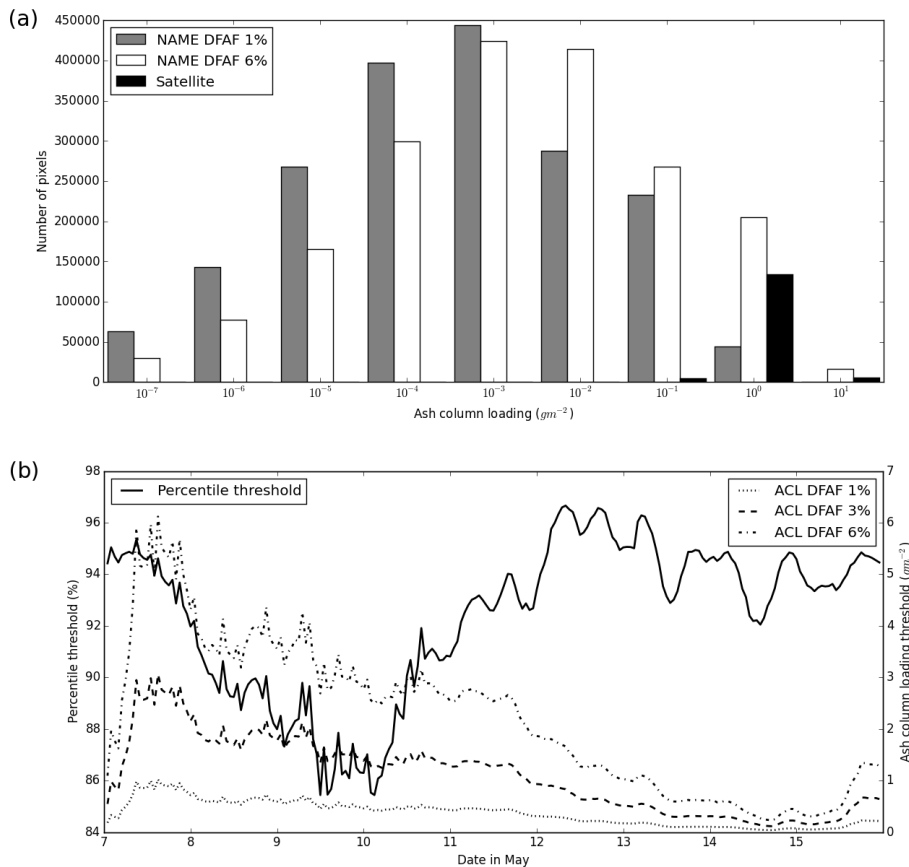
Simulated Ash Distribution	SI	PCC	Skill Score FSS (600 km <sup>2</sup> )	Scale
(a) No stretch	0.33	0.48	0.77	40 km <sup>2</sup>
(b) Stretch factor 0.5	0.06	0.07	0.40	700 km <sup>2</sup>
(c) Stretch factor 1.2	0.24	0.35	0.71	160 km <sup>2</sup>
(d) Stretch factor 2	0.06	0.07	0.29	960 km <sup>2</sup>

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)

**Figure 1.** Ash column loading at 00:00 UTC on 14 May 2010 **(a)** by the satellite (with 5x smoothing), **(b)** simulated by NAME, **(c)** NAME simulated ash cloud after pixel matching (i.e. black indicates pixels selected in satellite matching process). Panel **(a)** uses the colour scale shown in panel **(b)**.

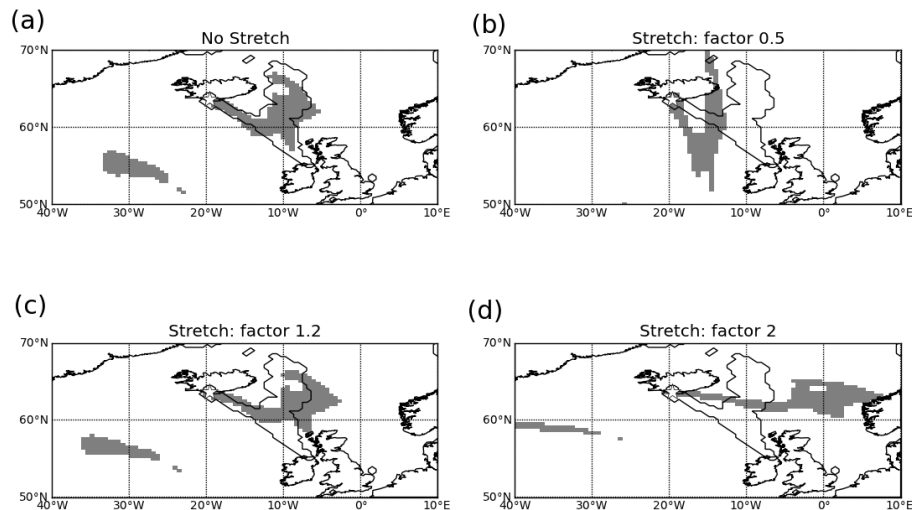
## Spatial evaluation of volcanic ash forecasts

N. J. Harvey and  
H. F. Dacre

**Figure 2.** (a) Number of pixels as a function of column loading for 7–16 May 2010 for both NAME (distal fine ash fraction (DFAF) of 6% (white) and DFAF of 1% (grey)) and satellite observations (black). (b) Time evolution of the percentile threshold (solid line) and minimum ash column loading calculated by applying the pixel matching technique (DFAF 1% (dotted line), DFAF 3% (dashed line), DFAF 6% (dot-dash line)).

## Spatial evaluation of volcanic ash forecasts

N. J. Harvey and  
H. F. Dacre

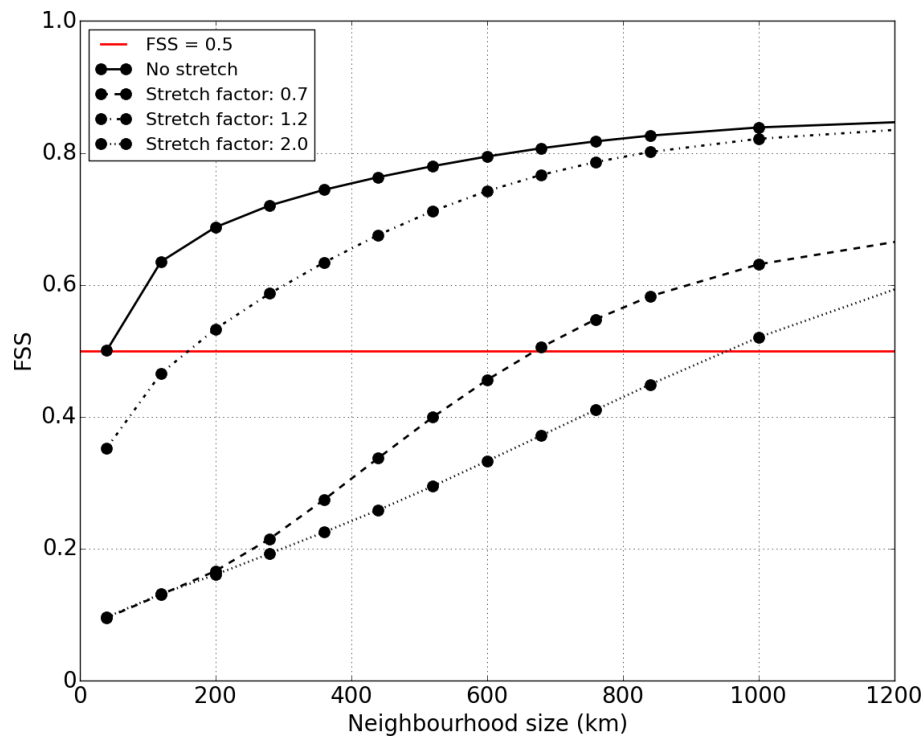


**Figure 3.** The pixel matched NAME ash cloud (grey shading) compared to the satellite retrieved ash cloud (black outline) with (a) no stretch, (b) stretch factor 0.5, (c) stretch factor 1.2, (d) stretch factor 2.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

## Spatial evaluation of volcanic ash forecasts

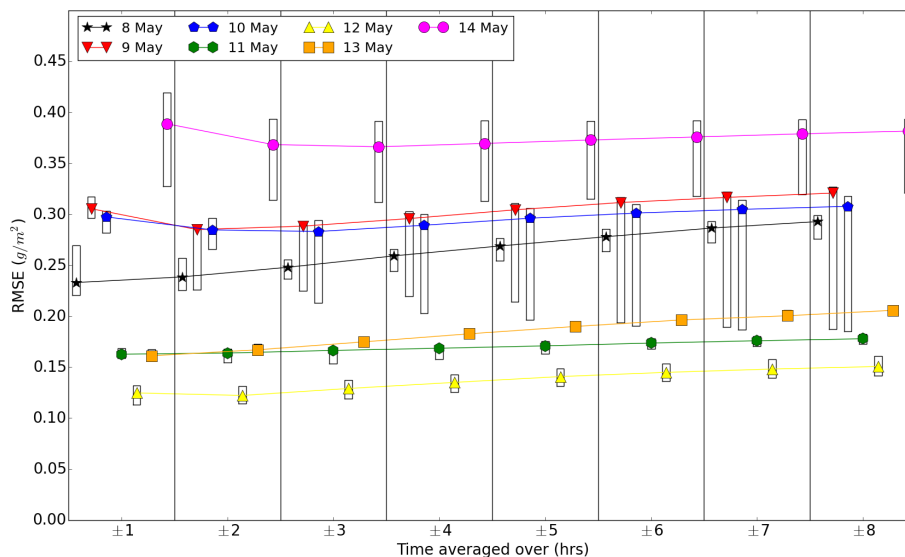
N. J. Harvey and  
H. F. Dacre

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

**Figure 4.** The FSS as a function of neighbourhood size for each of the three translations (dashed line: stretch factor 0.5, dot-dash line: stretch factor 1.2 and dotted line: stretch factor 2) compared to the original NAME simulation (solid black line) shown in Fig. 3a.

## Spatial evaluation of volcanic ash forecasts

N. J. Harvey and  
H. F. Dacre



**Figure A1.** The median RMSE between the SEVIRI observations at  $t_0$  (“truth”) and the truth with 50% of the pixels randomly replaced by the time averaged observations for each day 8 May 2010–14 May 2010 (8 May: grey stars, 9 May: grey downward-pointing triangles, 10 May: grey pentagons, 11 May: grey hexagons, 12 May: grey upward-pointing triangles, 13 May: grey circles, 14 May: grey squares). Each random replacement is repeated 50 times and the error bars show the interquartile range of the RMSE from these iterations.

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)
