Improved satellite detection of volcanic ash

Weather Science Technical Report 656

October 2023

Cameron Saint

Document history

Version	Purpose	Date
1.1	Published as SAN Tech Memo 35	06/05/22
1.2	Reformatted as Weather Science Technical Report 656	12/10/23

Prepared by

Cameron Saint, Scientist, 12/10/23

Reviewed by

Bob Tubbs, Manager of the Satellite Imagery team, 13/10/2023

Authorised for issue by

Simon Keogh, Strategic Head of Space Applications and Nowcasting, 31/10/23

Contents

Abstract

This Tech Memo summarises changes to the volcanic ash detection algorithm that improve the spaceborne observations of volcanic ash available to the London Volcanic Ash Advisory Centre (VAAC). Since the shutdown of significant parts of European airspace during the 2010 eruption of Eyjafjallajökull, there has been significant demand for information on volcanic plumes. Forecasts are obtained with dispersion models, such as the Numerical Atmospheric-dispersion Modelling Environment (NAME) used in the London VAAC, and observations are used in real-time to identify regions with airborne ash.

Geostationary satellites can view the entire plume with high temporal frequency, and several algorithms have been developed to automatically detect ash and retrieve its properties. The Met Office algorithm has been designed for the Spinning Enhanced Visible and Infrared Imager (SEVIRI) instrument mounted on-board Meteosat Second Generation (MSG). The method has since been extended to the instruments aboard three other geostationary satellites, GOES-16 (Geostationary Operational Environmental Satellite), GOES-17 and Himawari-8, as well as the MODIS (Moderate-resolution Imaging Spectro-radiometer) instruments aboard Aqua and Terra. The algorithm exploits the observed radiance in three infrared channels (8.7, 10.8 and 12 μm), by considering the Brightness Temperature Difference (BTD) between these channels. The BTD tests are combined with various thresholds on the ratios of effective absorption optical thickness (beta ratios). Based on the above quantities, we assign a confidence level to the volcanic ash detection. I will briefly explain how this data is used in the London VAAC and by developers of the NAME dispersion model.

The key improvements that I am researching are achieved through the refinement of detection thresholds and the addition of new tests designed to counter false detections primarily in desert regions, but also other areas such as at high satellite zenith angle. Here, I will show how the results compare to the current (at time of writing) operational set-up. The new scheme results in, on average, an 11% increase in correct detection of volcanic ash and an 85% reduction of false positives. The improved satellite imagery will be implemented operationally at the Met Office, and the VAAC will benefit from it during future eruptions.

1. Introduction

Volcanic ash is a hazard, especially for aircraft, and as such the aviation industry is very interested in accurate forecasts of dangerous levels of volcanic ash following an eruption. The primary products they are interested in are output from dispersion models. However, satellite data is crucial, both in assessing the accuracy of the dispersion model output and also for improving it. At the Met Office this is achieved using our inversion software InTEM, which effectively infers how much ash has been emitted at various vertical intervals above a volcano, and at what time, based on how much ash is visible in the satellite imagery. This information is then used to update the source term in the dispersion model. It is therefore important to have automated systems that can accurately identify ash in the imagery and evaluate how much ash is likely to be there. The method of detection at the Met Office has not changed significantly from the method described in [*Francis et al*, 2012]. Since then, many other more complex algorithms used around the world have been developed to accurately identify ash [*Pavolonis et al.*, 2015b] [*Piontek et al*. 2021] [*Mackie and Watson,* 2014], including methods that use information beyond the spectral data used in the current Met Office algorithm [*Francis et al*, 2012] [*Saint, 2020*].

Each detection method intends to maximise correct detection whilst minimising false detection. However, it is challenging to increase the number of correct detections without also increasing the number of false detections. Therefore, each automated algorithm must strike a balance between having the maximum ability to correctly detect pixels of ash with minimising the number of pixels incorrectly marked as ash. The most appropriate balance is influenced by the main uses of the products produced from the automated algorithm. At the Met Office, the data and imagery produced by the automated algorithm has two primary uses. These are to provide the data for InTEM and for forecasters to use when verifying the accuracy of NAME forecasts.

For InTEM it is critical to minimise false detection, especially close to the real ash plume. This is because InTEM will treat the false detection as correct detection and attempt to alter the source term to get ash to this location. However, not all false detection is problematic for InTEM. Where it would not be possible for ash from the source to have reached the location of false detection, InTEM does not use these observations and they don't impact the source term. Additionally, as the NAME and InTEM grids are lower resolution than the image grid there is a re-gridding process before data from the satellite is available to InTEM. False detection covering some of the pixels within an InTEM grid box may effectively be removed by the re-gridding process. In order for InTEM to use the observations, 50% of the pixels within a gridbox must be identified as ash or 90% of the pixels as ash or clear sky pixels. Furthermore, for these observations the average column loading of the ash and clear sky pixels within the gridbox is taken, so in cases where only some of the pixels contain false detection the impact of these false detections is lessened. These points are important to be aware of when developing improvements to the automated detection algorithm as total false detection statistics do not always reveal the full story of the amount of improvement one might expect from a new algorithm. The requirements for OpMets are similar although they are usually better able to deal with false detection than any software. This is because a forecaster is able to make a better assessment of whether it is reasonable for ash to be in a given location based on previous imagery and the meteorological conditions (i.e forecasters have a better assessment of the prior probability of ash). As a result, false detection that is not nearby to the ash plume is largely irrelevant to an OpMet. Therefore, for this purpose a balance where there is more detection, both correct and false, would likely be more suitable compared to the optimal balance for InTEM. However, OpMets primarily use the Dust RGB

imagery to identify ash, which means that the optimal balance for our algorithm is strongly based around producing the best possible data for InTEM.

2. Detection Method

2.1. How does the detection work?

The brightness temperature observed at the satellite is largely determined by the temperature at the location of emission. For example, if the emission is from the surface the brightness temperature will be that of the surface (i.e relatively warm). Different atmospheric absorbers, such as meteorological clouds or a volcanic ash plume, will be more absorptive at different wavelengths. As a result, it is often possible to determine the composition of an atmospheric absorber based on BTD tests. In particular, we expect ash of typical composition to produce a negative BTD between 10.8 µm and 12.0 µm, as ash is more absorbing at 10.8 µm than 12.0 µm, whereas the opposite is true for water and ice (as illustrated in Figure 1). These brightness temperature differences can be used to create composite imagery.

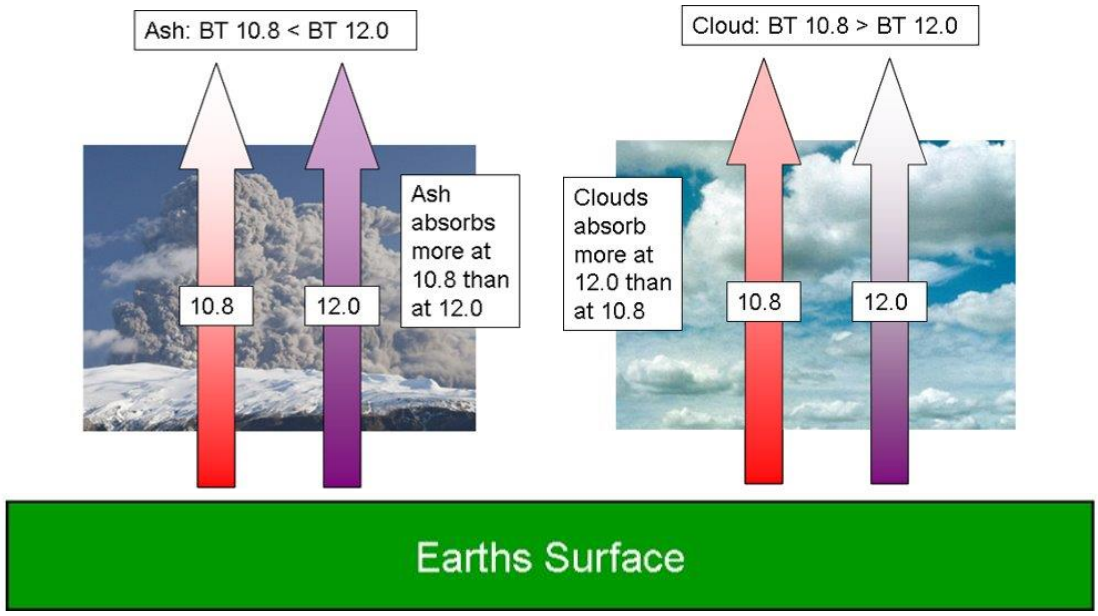

Figure 1 *Graphic illustrating why ash plumes can be differentiated from meteorological cloud with the BTD(10.8-12.0).*

The composite imagery includes a product known as the Dust RGB, where the ash present can be identified by red, orange, or yellow colours. The precise colour of the ash in the Dust RGB imagery is dependent on a variety of factors, including the height and amount of ash present. However, the amount of SO_2 present with the ash can have a significant impact on the colour seen in the Dust RGB. In particular, bright yellow colours usually indicate the presence of a significant amount of SO_2, although it is also possible for yellow colours to be indicative of more smaller ash particles being present. Figure 2 shows the BT information used to produce the Dust RGB alongside an example image with two regions of ash labelled. Further information on the Dust RGB product which was originally devised by EUMETSAT can be found in SA Tech Memo n.110 [Saint, 2020]. The BTD's are also used in the Met Office's automated algorithm which is used to detect volcanic ash. In addition to BTD's, the Met Office algorithm uses parameters known as beta ratios to further improve our

volcanic ash detection beyond what is achieved with BTD tests alone. Beta ratios are parameters devised by Mike Pavolonis [*Pavolonis*, 2010] which are a function of the ratios of the emissivities in two channels and utilise model data in order to account for some other variables that can result in the thresholds for the BTD tests being exceeded.

The two beta ratios used are $\beta(8.60, 10.45)$ and $\beta(12.35, 10.45)$, which are defined as:

$$\beta(\lambda_1, \lambda_2) = \frac{\ln(1 - \varepsilon(\lambda_1))}{\ln(1 - \varepsilon(\lambda_2))} \quad (1)$$

where λ_1 is the peak wavelength of the first channel, λ_2 is the peak wavelength of the second channel and ε is an effective emissivity for each channel. The effective emissivities used to calculate the beta ratios are given by:

$$\varepsilon(\lambda) = \frac{R_{obs}(\lambda) - R_{clr}(\lambda)}{R_{ovc}(\lambda) - R_{clr}(\lambda)} \quad (2)$$

where $R_{obs}(\lambda)$ is the measured radiance for that wavelength, $R_{clr}(\lambda)$ is the RTTOV clear sky radiance and $R_{ovc}(\lambda)$ is the RTTOV overcast radiance, whilst assuming an emission height which corresponds to an NWP model temperature of BT(10.8) – 5 K. This height is chosen as it should be a reasonable estimate for the optically thicker ash in the regions nearer the source [1]. The use of RTTOV modelled variables will lead to inaccuracies in the calculated emissivities and subsequently calculated beta ratios. Further information on the current method of volcanic ash detection used at the Met Office (at the time of writing) can be found in both [*Francis et al*, 2012] and SA Tech Memo n.110 [*Saint*, 2020].

R = BTD IR12.0μm - IR10.8μm (Min = -4 K, Max = +1 K)

High Red values indicate ash

G = BTD IR10.8μm - IR8.7μm (Min = 0 K, Max = +15 K)

Varies depending on ash size (smaller=greener) and SO2 content.

B = BT IR10.8 μm (Min = 261 K, Max = 289 K)

Low Blue = cold (i.e. It is high up)

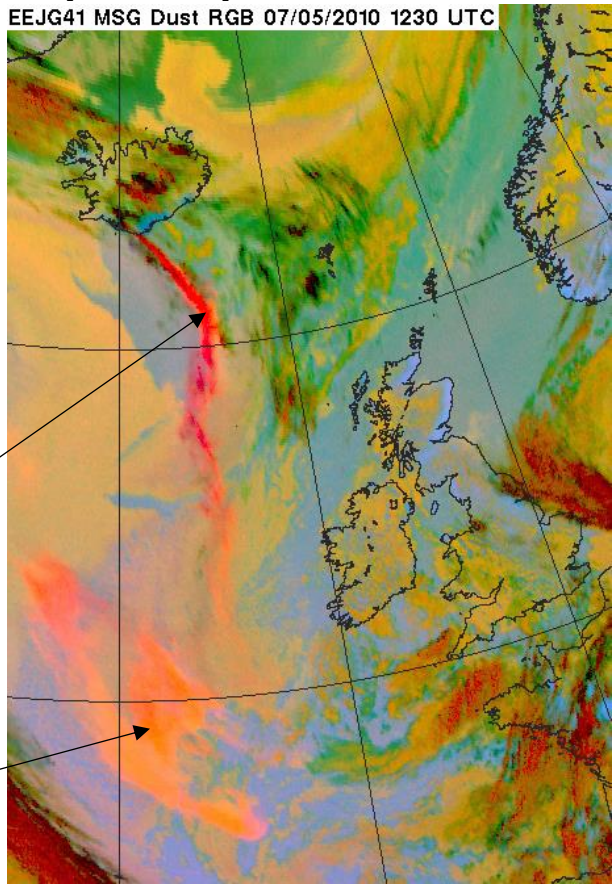

Figure 2 *Graphic demonstrating how a Dust RGB can be used to identify ash.*

3. Detection improvements

The detection of volcanic ash can be improved in two ways. These are by either increasing the number of pixels correctly marked as ash or by reducing the number of pixels incorrectly marked as ash. In general, it is difficult to increase correct detections without also increasing incorrect detections, as a pixel containing ash could have the same brightness temperatures in the 3 channels used in the detection as a pixel not containing ash. I have made two main changes to the detection method that take advantage of factors that impact the prior probability of both a correct or false detection of ash.

In order to help determine whether the 'marginal' pixels (pixels with BTD's that do not allow confident detection of ash) contain ash, I created a spatial filter designed to increase the likelihood of ash being detected if nearby pixels have detected ash. For practical reasons this is applied in a two-step process, first increasing the likelihood of pixels being marked as ash, then pixels which don't have significant detection of ash nearby are marked as not ash if they don't pass stricter BTD and beta ratio thresholds. The BTD and beta ratio tests for both the initial test, which now marks more pixels as ash than the previous one step process, and the retest, are both set based on qualitative testing across several cases. The aim of the testing was to increase the number of pixels correctly detected by more than the number of new pixels incorrectly marked as ash. The increased detection due to this additional spatial filtering can be seen comparing Figures 14(b) and 14(c), in particular over the UK.

I have also introduced location-based filters designed to reduce false detections from common issues that can lead to BTD signals indicative of ash. One source of BTD signals similar to ash is dust. This is because dust also results in a strongly negative BTD(10.8-12.0). Significant dust is usually only seen over arid regions of Earth and as such I have introduced stricter BTD and beta ratio thresholds in arid regions to reduce the chances of pixels being marked as ash. As with the spatial filter, the thresholds were defined based on subjective testing across several cases. Unfortunately, there was no test data with ash over the relevant regions, so the aim of the testing was to remove false detections from all but the strongest dust events. These significant dust events are most similar to a volcanic ash plume and as such it is hoped that if these are still detected then significant ash would also be detected. However, a more detailed assessment of this would certainly be beneficial should suitable test data become available. In these regions it is also common for surface emissivity effects to lead to pixels being incorrectly marked as ash, and these are also reduced by the new location-based thresholds. The impact of this change can be seen with the significant reduction of false detection visible in imagery produced with my experimental algorithm in Figure 15(c) compared to our current operational algorithm used in Figure 15(b). In addition to altered thresholds over arid surfaces, I also made small changes to the thresholds at high zenith angle to counter false detections due to increased path lengths and at low latitudes to counter false detections from optically thick clouds with high cloud tops. It should be noted that all of the location filters reduce the chance of real ash being detected over these regions and are therefore only applied in areas where false detections are very common.

The thresholds used in the new algorithm are defined in the following sections. They are set based on extensive qualitative testing of a variety of archived datasets containing ash. As with the spatial filter, all thresholds are set with the aim of increasing the number of pixels correctly detected by more than the number of new pixels incorrectly marked as ash, or reducing the number of pixels incorrectly marked as ash by more than the number of correctly detected ash pixels lost. As a result of the limited data available containing ash, the

archived datasets will not be representative of all conditions. This issue is more significant for particular satellites such as GOES-16 and GOES-17 where no data containing ash was available. Extending the archived datasets when future eruptions occur or with simulated imagery could help with this issue. The testing included experimentation with a range of beta spaces for both the initial test and the second test following the spatial filter where appropriate. The BTD_Cutoff and clearskycutoff variables were also optimised for MSG and Himawari. The weightings applied to the confidence levels before the spatial filter, which are discussed further in section 3.4, were also optimised for Himawari and MSG. As this resulted in the same weightings for both satellites, and due to the lack of volcanic test data, these were assumed to be optimum for GOES and MODIS. Further testing on these satellites if test data became available would likely result in some improvement. In all cases the confidence thresholds were not modified as I did not find any evidence that the optimisation of these thresholds could be improved.

Prior to becoming the new operational method, the experimental settings were run on the development servers for a period of weeks to ensure no issues. This resulted in some minor modifications to counter false detection seen in the GOES and MODIS imagery. In both cases, the BTD_Cutoff was reduced and significantly so in the case of GOES. These changes were made as prior to the changes the new method would often result in more false detection than the previous operational method. The changes were made with the aim of ensuring that the amount of false detection in the new method was always less than or similar to the old operational method. It should be noted that with no archived datasets containing ash for GOES I was unable to test the impact of these changes on correct detection. It would be worth assessing this in future should datasets containing ash become available or perhaps with the production of simulated imagery for these satellites. It is not immediately clear why the instruments on board the GOES satellites require a significantly more negative BTD_Cutoff than for the instrument on Himawari given the similarities of the instruments. However, it's possible that minor differences in the instruments combined with the different meteorology and topography could result in false detection occurring at more negative BTD's for the GOES satellites.

3.1. Updated confidence assignment

In addition to the introduction of the spatial and location-based filters, I also made some modifications to the confidence assignment when marking pixels as ash. There were two main aims of this change: First, I wanted to utilise the information in the 3-channel BTD more in the assignment of ash confidence; Second, with the addition of the spatial filter it had become more important to ensure that the confidence value assigned to a pixel reflected the real confidence the pixel was ash. Previously all pixels marked as ash were treated the same regardless of the confidence level applied, but with the addition of the spatial filter the specific confidence applied was now important in the calculation of the weighted average (discussed further in Section 3.4). The changes apply to confidence levels 1-6 and a full comparison to the previous confidence assignment method can be made by viewing the previous method described in detail in [*Saint, 2020*].

First, I will define the Fortran variables and functions that are used in the assignment of confidence that volcanic ash is present in a given pixel.

BTD2 = BTD(10.8-12.0) = 2 channel BTD

CT(i)= confidence_thresh(i) : This 4 element vector is used to define the relevant 2 channel BTD's that should be used in the assignment of volcanic ash confidence (as defined in Table 1)

BTD_Cutoff = This is now essentially a 5th element of the above array and defines the most positive 2 channel BTD where ash may be detected for each satellite.

BTD3 = BTD(10.8 -12.0 + 10.8 - 8.7) = 3 channel BTD

BTD3Thresh = 1.5 K = This defines the 3 channel BTD used in the confidence assignment. Where ash is present BTD's are more likely to be less than this value.

Clearskycutoff = This is the threshold used in determining whether a pixel has been incorrectly marked as ash due to a surface effect. Its value is not dependent on the satellite, but is dependent on which location-based filter is being considered and is therefore defined later in Tables 2 and 3. The value is compared to the difference between the 10.8 μm radiance observed at the satellite and the simulated clear sky radiance as calculated by RTTOV.

SurfaceEffect = This is a logical which is true when the difference described above is greater than the relevant clearskycutoff value. In effect, that means pixels are marked as clear sky pixels when the radiance observed at the satellite is not sufficiently colder than the RTTOV clear sky radiance.

Satellite	CT(1) (K)	CT(2) (K)	CT(3) (K)	CT(4) (K)	BTD_Cutoff (K)
MSG	-2.0	-1.5	-1.0	-0.5	-0.1
Himawari	-1.99	-1.38	-0.78	-0.17	0.5
GOES-16	-2.06	-1.47	-0.88	-0.29	-0.29
GOES-17	-2.06	-1.47	-0.88	-0.29	-0.29
AQUA-MODIS	-1.40	-1.07	-0.73	-0.39	-0.39
TERRA-MODIS	-1.39	-1.06	-0.73	-0.39	-0.39

Table 1. Two channel BTD for all the satellites and instruments using the 3-channel method of detection (some polar satellites use a simpler 2-channel BTD method).

For pixels outside the northern hemisphere arid surface filter region (which is defined later in section 3.3.1), confidence 7 is assigned independent of beta ratios and the surface effect test. These pixels are considered to have a strong enough BTD signal that it is not necessary to filter these detections with beta ratio tests:

- BTD2 <= CT(1) → VA_Confidence = 7

Pixels which don't have a BTD2 signal clearly indicating ash must also be within a favourable region of beta space to be marked as ash. These beta spaces are defined by the following tests with the values of aa, bb and cc dependent on the geographical region (as defined in Table 3):

- $\beta(12.0,10.8) < aa * \beta(8.7,10.8)^2 + bb * \beta(8.7,10.8) + cc - 0.4$ → beta_mask_conservative

- $\beta(12.0,10.8) < aa * \beta(8.7,10.8)^2 + bb * \beta(8.7,10.8) + cc$ → beta_mask_liberal

By varying aa, bb and cc it is straightforward to alter the strictness of the beta tests for different regions. In order to counter significant false detection in arid regions in the northern hemisphere from dust and surface effects it was necessary to test these pixels with strict

beta and clearskycutoff tests in addition to the simple 2 channel BTD. As a result, the confidence 7 pixels in these regions are assigned alongside the confidence 1-6 pixels in other regions in the following tests. For each of these tests the SurfaceEffect logical must be false:

- NHAridSurfaceFilter and BTD2 <= CT(1) and beta_mask_liberal
  → VA_Confidence= 7
- CT(1) < BTD2 <= CT(3) and BTD3 <= BTD3Thresh and beta_mask_conservative)
  → VA_Confidence= 6
- CT(1) < BTD2 <= CT(3) and BTD3 <= BTD3Thresh and beta_mask_liberal)
  → VA_Confidence= 5
- CT(1) < BTD2 <= CT(4) and BTD3 >= BTD3Thresh and beta_mask_conservative)
  → VA_Confidence= 4
- CT(3) < BTD2 <= BTD_Cutoff and BTD3 <= BTD3Thresh and
  beta_mask_conservative) → VA_Confidence= 3
- CT(1) < BTD2 <= CT(3) and BTD3 >= BTD3Thresh and beta_mask_liberal)
  → VA_Confidence= 2
- CT(3) < BTD2 <= CT(4) and beta_mask_liberal) → VA_Confidence= 1

3.2. Updated beta space

Table 2 and Figure 3 show how the new acceptable beta space for detection compares to the previous operational beta space. In order for a pixel to be marked as volcanic ash it must be below and to the left of the relevant line depending on which confidence is being assigned as detailed in section 3.1. Here it can be seen that the new acceptable beta space is larger allowing for greater detection of volcanic ash. The main reason this is possible is that the spatial filter will remove many of the new false detections allowed by this increased acceptable beta space whilst keeping some of the new correct detection allowed through. It is worth noting that in the previous operational method confidence 1 detections were assigned with a different beta space than the three lines in Figure 3. However, this beta space only included a very small region outside of the ‘operational liberal’ region and was entirely within the new ‘experimental liberal’ region so does not alter the key overall point that the new acceptable beta space is larger. In particular, the new beta space allows for more detection in the top left region of the graph, where β(8.7,10.8) is relatively low and β(12.0,10.8) is high. Qualitative testing indicated this space allowed for the best compromise of increasing correct detection whilst minimising false detection.

Region	aa	bb	cc
Experimental liberal	-0.4	-0.4	2.5
Experimental conservative	-0.4	-0.4	2.1
Operational liberal	0	-1.0	2.1
Operational	0	-1.0	1.9
Operational conservative	0	-1.0	1.7

Table 2. Variables defining the acceptable beta spaces for detection for the previous operational method and new experimental method.

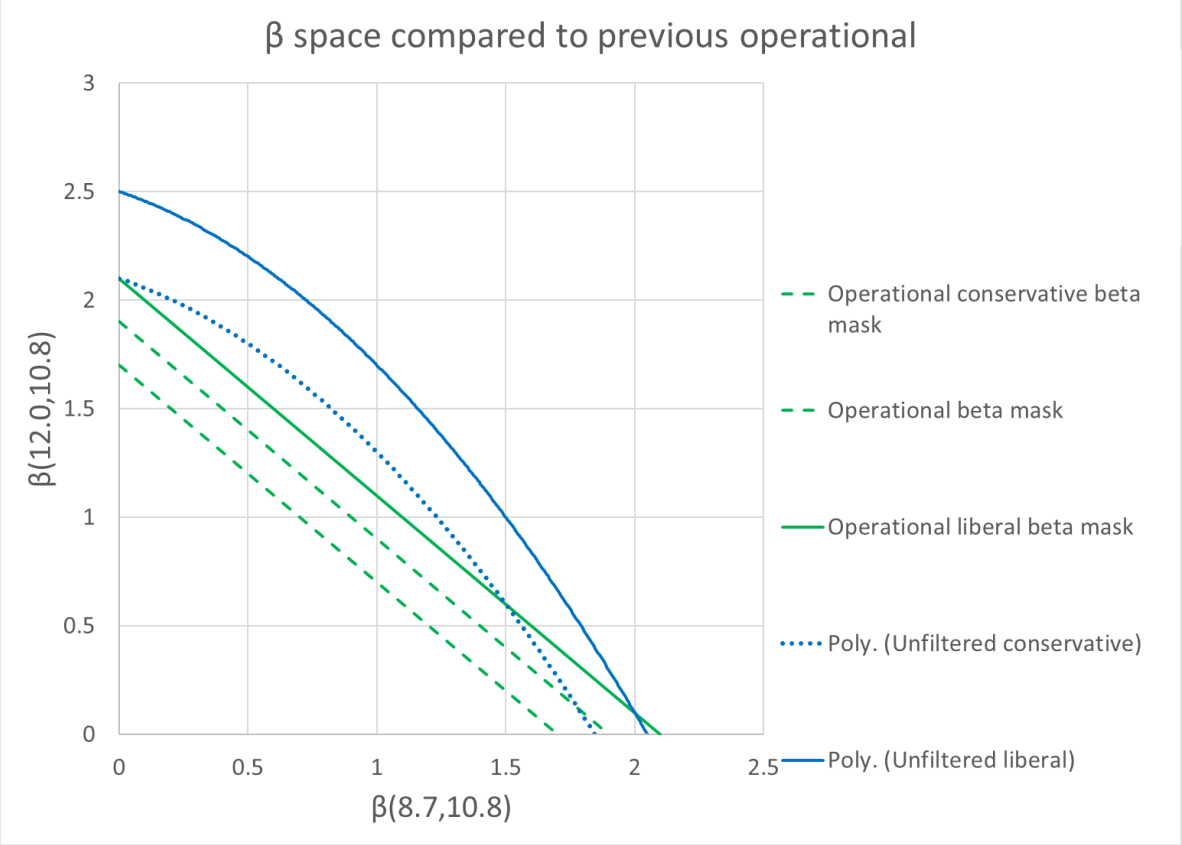

Figure 3 *Comparison of the operational beta space (green lines) with the new experimental beta space in unfiltered regions (blue lines).*

3.3. Regional filters

Much of the false detection seen in current operational imagery is limited to specific geographical regions and is the result of known issues for the detection method. In order to reduce this false detection, I have introduced a selection of simple geographical filters that reduce the chances of detection in areas where false detection is most common. This is achieved primarily by reducing the beta space where ash can be detected in addition to reductions to the clearskycutoff variable in arid regions. The specific changes are detailed in Table 3 and the changes to the beta space are also displayed in Figure 4. The specific regions which the geographical filters act on are shown in Figure 5 and the motivations behind the relative strictness of the beta space for each region are described in the following sections. In any case where the requirements for more than one regional filter are met, the thresholds will be set at the values of the strictest filter.

Region	aa	bb	cc	Clearskycutoff (K)
Unfiltered	-0.4	-0.4	2.5	-5.0
Low latitude	-0.9	0	2.3	-5.0
High Zenith angle	-1.0	0	2.3	-5.0
High Zenith angle (MODIS)	-1.0	0	1.5	-5.0
SH Arid Surface	-1.0	0	1.6	-6.0
NH Arid Surface	-1.0	0	1.3	-25.0

Table 3. Variables that are varied in the different regional filters. aa, bb and cc define the beta space for each filter.

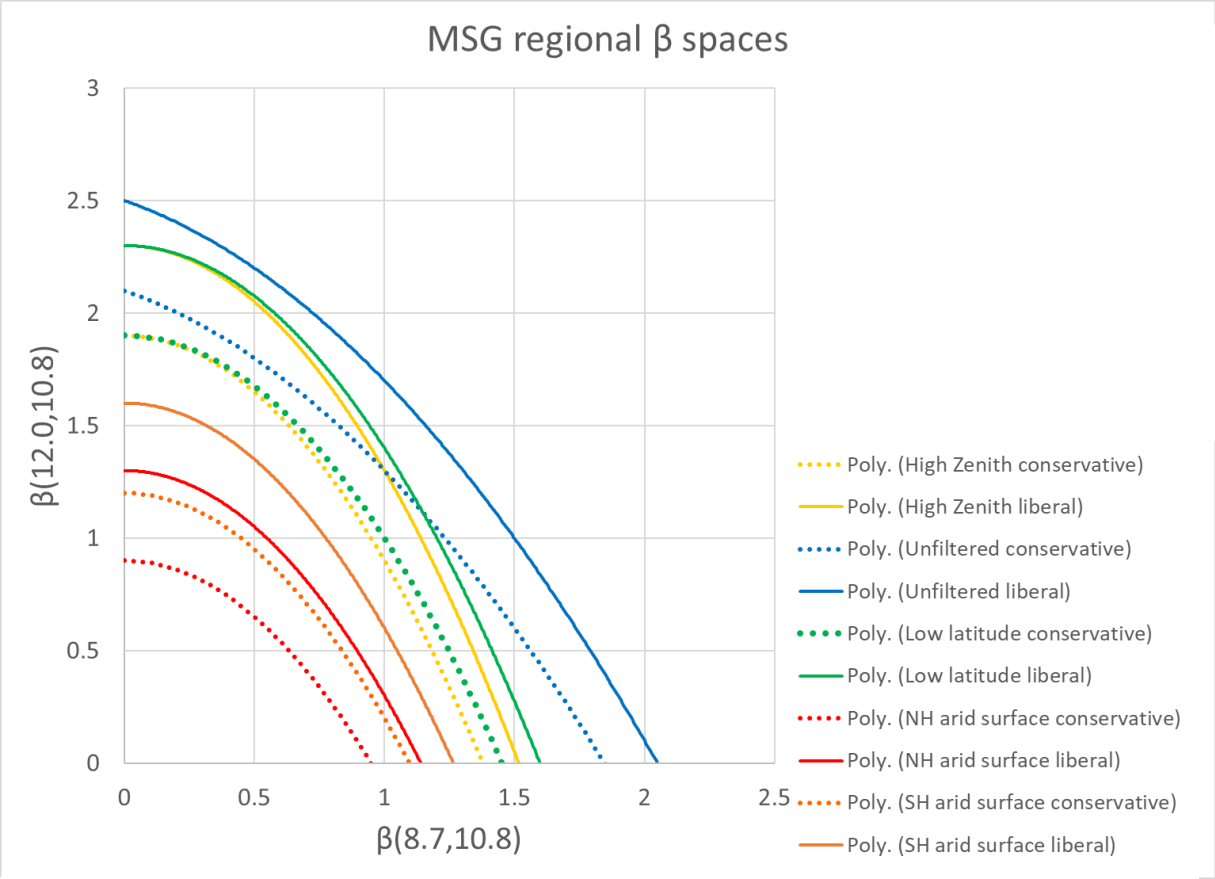

Figure 4 *Graph with all of the regional beta spaces for MSG in the new experimental method: unfiltered regions (blue), Low latitude (green), High zenith (yellow), SH arid surface (orange), NH arid surface (red)*

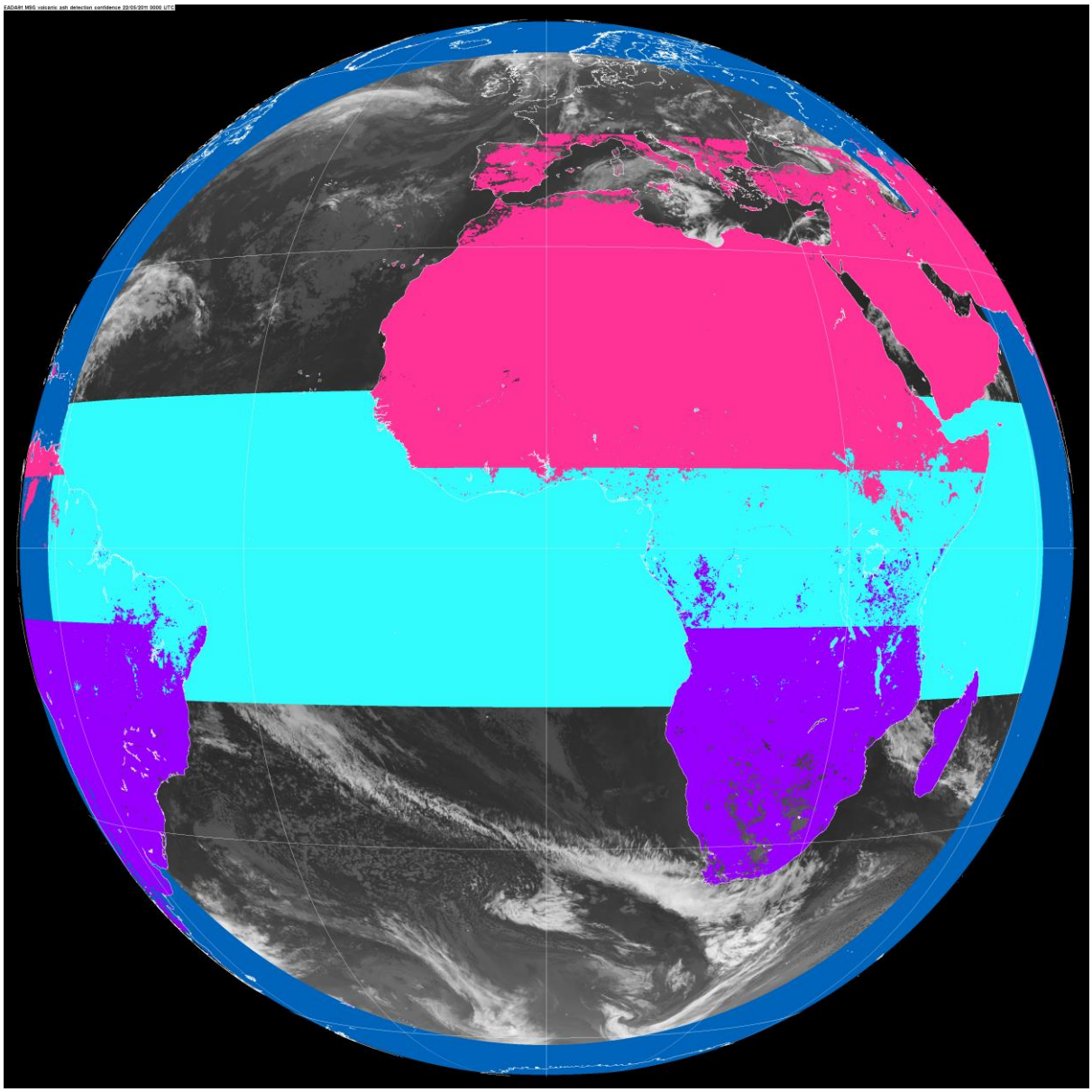

Figure 5 *Image showing the regions where the regional filters act: NH arid surfaces (pink), SH arid surfaces (purple), high zenith angle (dark blue), low latitude (light blue).*

3.3.1 NH Arid surface filter

The northern hemisphere arid surface filter is the strictest of the regional filters, as shown in Table 3 and Figures 4 and 6. The specific regions covered by this filter are defined based on the IGBP surface types defined in SPS. IGBP surface types 9 and 16 are covered by this filter across the northern hemisphere, and surface types 2, 7, 8, 10, 19, 20 and 22 are covered by this filter for latitudes between 7.5° and 45°. In the current operational method, the most significant false detection is seen in these regions, in particular over the Sahara Desert. This is due to a combination of the frequency of dust events in the region and surface emissivity effects. Dust is an issue for the ash detection as it produces a very similar spectral signal to ash. Although it is usually closer to the surface than ash, it would be difficult to consistently mark pixels as dust if they have a relatively warm single channel brightness temperature without losing low altitude or optically thin ash plumes from the

detection. The cause of the remaining false detection seen in these regions is likely a combination of variation in the surface emissivity with channel wavelength and inaccurate model surface temperatures.

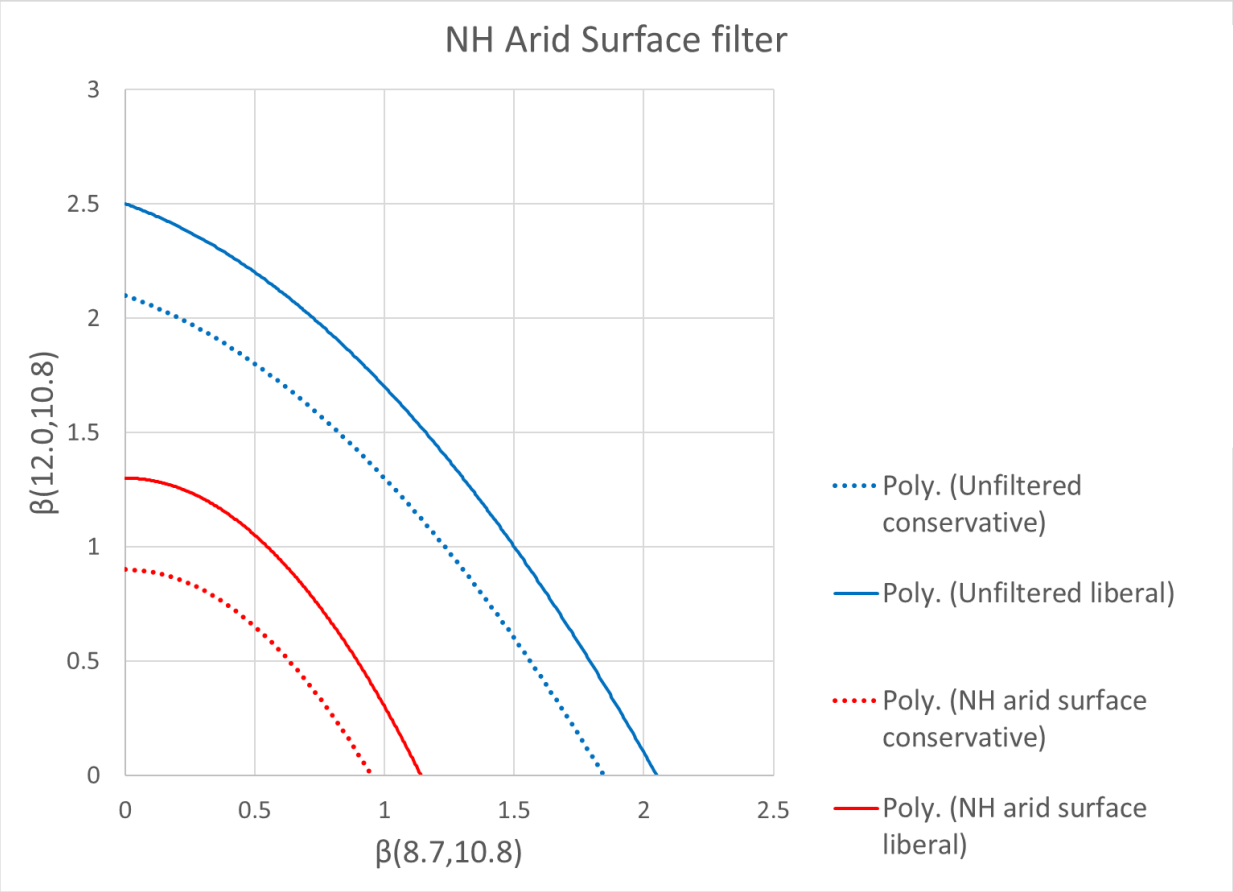

Figure 6 *Comparison of the beta space for unfiltered regions (blue) with the beta space for the NH arid surface regions (red)*

In arid regions the surface emissivity in the 12.0 µm channel is greater than the 10.8 µm channel [*Ashpole and Washington*, 2012]. Therefore, ignoring atmospheric effects, the brightness temperature seen at the satellite is warmer in the 12.0 µm channel than the 10.8 µm channel resulting in the negative BTD(10.8-12.0) used to detect ash. However, for ash to be detected the beta ratio thresholds must also be passed and these should account for the impact of the surface emissivity effects. The beta ratios are defined in equations 1 and 2. Ideally the different surface emissivities in each channel, and their effect on the modelled brightness temperatures, would be accounted for within the RTTOV clear sky and overcast radiances, and therefore not impact the emissivities being calculated for the atmospheric absorber we are trying to identify. However, there is a known issue that the model surface temperature can be quite inaccurate in arid regions, this is usually an underestimation especially at dawn, dusk and overnight. This inaccuracy in the modelled surface temperature will affect the modelled brightness temperature in each channel differently depending on the emissivity in the channels. As the surface emissivity is higher in the 12.0 µm channel than the 10.8 µm channel, BT(12.0) will be more significantly underestimated compared to the BT(10.8) in proportion to the error in the surface temperature in the model (in the case of an underestimated model temperature). As a result the modelled BTD(10.8-12.0) will be less negative than the real BTD(10.8-12.0), and therefore the model will need to increase the

emissivity of the atmospheric absorber in the 10.8 μm channel compared to the 12.0 μm channel, which will increase the chances of the beta thresholds indicative of ash being passed. It is possible that in the future a more accurate model radiative skin temperature may reduce the frequency and magnitude of the issue and allow for weaker filters in these regions.

The false detection in these regions is significantly reduced through the use of a much stricter beta space and also a greatly reduced clearskycutoff. It should be noted that these filters will also significantly reduce the chance of ash being detected in these regions. This is a compromise that we consider worthwhile for a number of reasons. Firstly, false detection is far more common in these areas than real detections. Real ash detection is also much more likely to be outside of these regions as there aren't many active volcanoes in these regions. As a result, in the vast majority of cases this change will objectively improve the detection. In addition, most of these surfaces are outside the London VAAC's area of responsibility and therefore detecting ash in these regions is not our highest priority. Furthermore, these changes should not make the detection of significant ash impossible. Despite the changes, significant dust events are usually still at least partially detected by the new method so it is reasonable to expect that the same would also be true for significant ash. However, it would be worth investigating this more thoroughly in the future should ash travel over the relevant regions or perhaps by producing simulated ash plumes over this region.

An example where significant false detection from both surface emissivity effects and significant dust events can be seen is shown in Figure 7. The significant dust events can be identified through the small pink coloured regions with sharp boundaries visible over the Sahara Desert in Figure 7(a). There are other pink regions that blend more smoothly into the background which are likely caused by surface emissivity effects. Figure 7(b) shows that these two effects result in significant false detection of ash over Africa and Arabia (alongside the correct detection of the Calbuco ash plume around South America), using the previous operational method. This false detection is significantly reduced by the new experimental settings (Figure 7(c)), with only small regions of false detection remaining. Although in this case the remaining false detection is not useful, it is indicative that significant ash could still be observed in these regions despite the stricter thresholds used in the detection. This is because only dust producing the strongest signal in the BTD(10.8-12.0) is still detected, and this is likely a similar signal to what the presence of a significant amount of ash would produce.

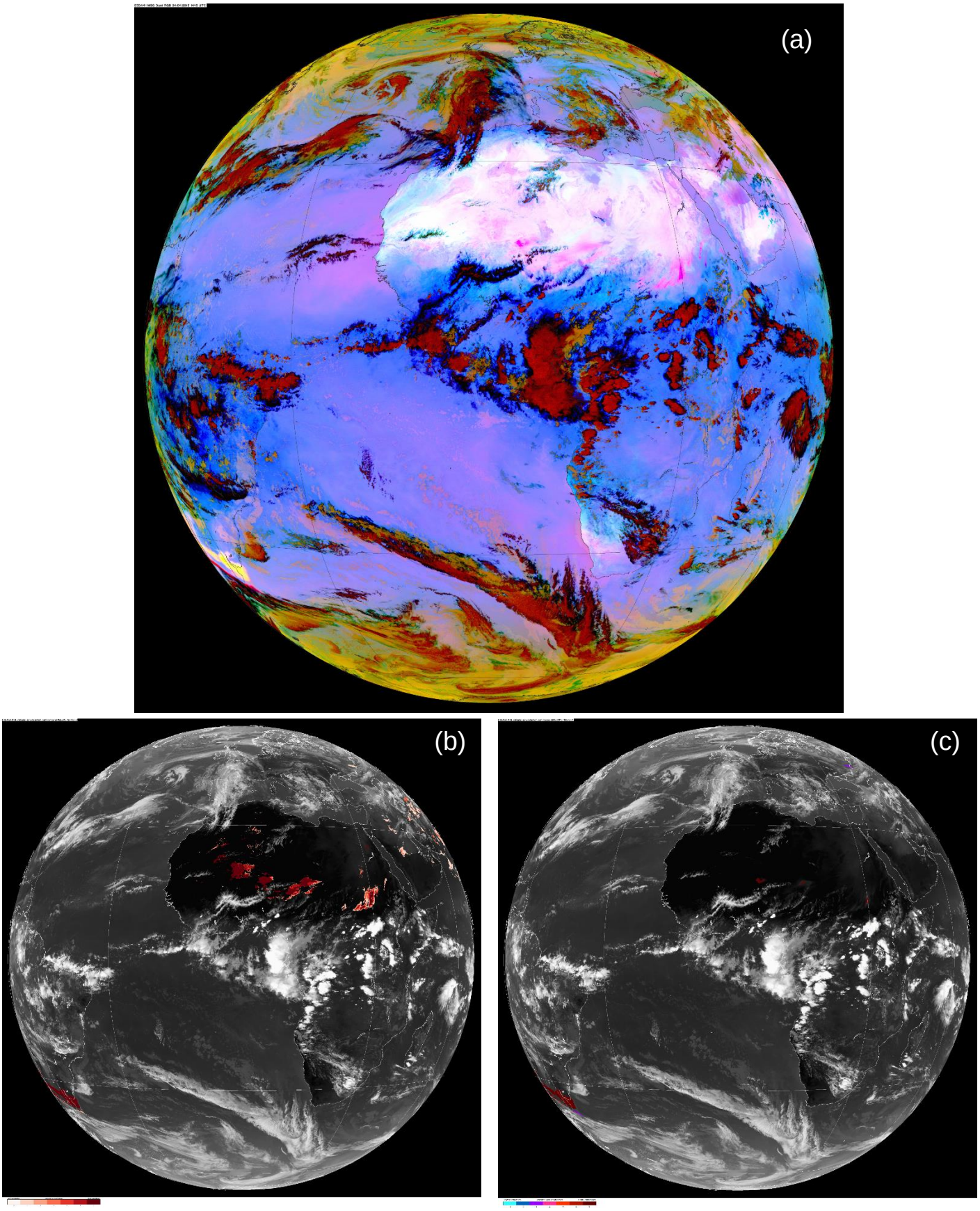

Figure 7 *Dust RGB (a), old operational volcanic ash confidence imagery (b), new experimental volcanic ash confidence imagery (c) at 1445 (24/04) following the 2015 Calbuco eruption.*

3.3.2 SH Arid surface filter

The next strictest filter (for geostationary satellites) is the southern hemisphere arid surface filter, for which the acceptable beta space in these regions is displayed in Figure 8. This does not need to be as strict as the northern hemisphere arid surface filter as both the dust and surface emissivity issues in these regions result in strong BTD signals indicative of ash much less frequently. This is probably due to the smaller size of the deserts in the southern hemisphere. As with the northern hemisphere arid surface filter, the regions covered by this filter are IGBP surface types 9 and 16 in the southern hemisphere, and surface types 2, 7, 8, 10, 19, 20 and 22 for latitudes between -7.5° and -45°.

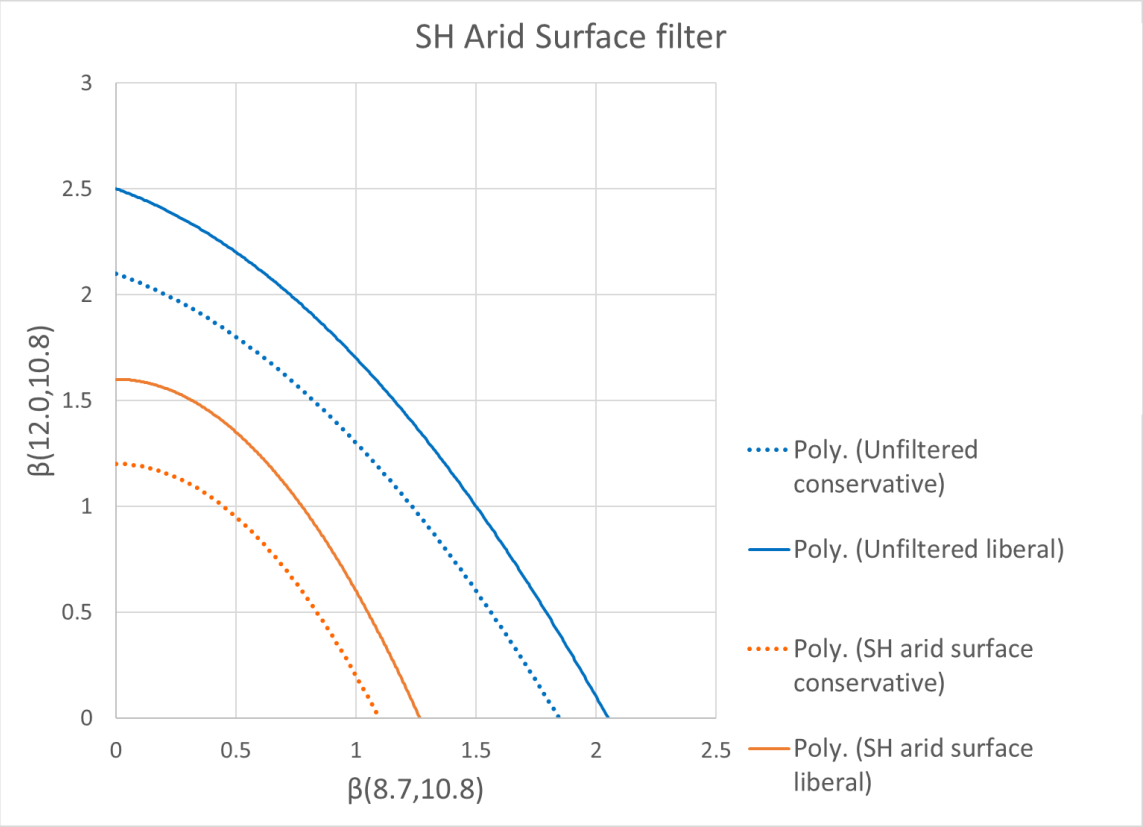

Figure 8 *Comparison of the beta space for unfiltered regions (blue) with the beta space for the SH arid surface regions (orange).*

3.3.3 High zenith angle filter

The high zenith angle filter uses thresholds only marginally stricter than the thresholds used in the unfiltered regions. This is partially because the beta space shown in Figure 9 was sufficient to remove the majority of false detection from these regions during the testing of the settings. However, it is also because detecting ash around Iceland is the primary responsibility of the London VAAC, and this region is at relatively high zenith angle. The satellite zenith angle above which the beta space in Figure 9 is used is set at 70 degrees for MSG and Himawari. This means that only a small region is filtered with these settings, but this could still be significant if ash were to initially spread north or west of Iceland. Following testing on the autosat development server, it was decided that reducing the satellite zenith angle above which the below filtering is used to 65 degrees would be beneficial in removing false detection for both GOES satellites. For the MODIS satellites, the zenith angle above which the relevant filter is active is 62.5 degrees, and more detail on this is provided in section 3.3.5.

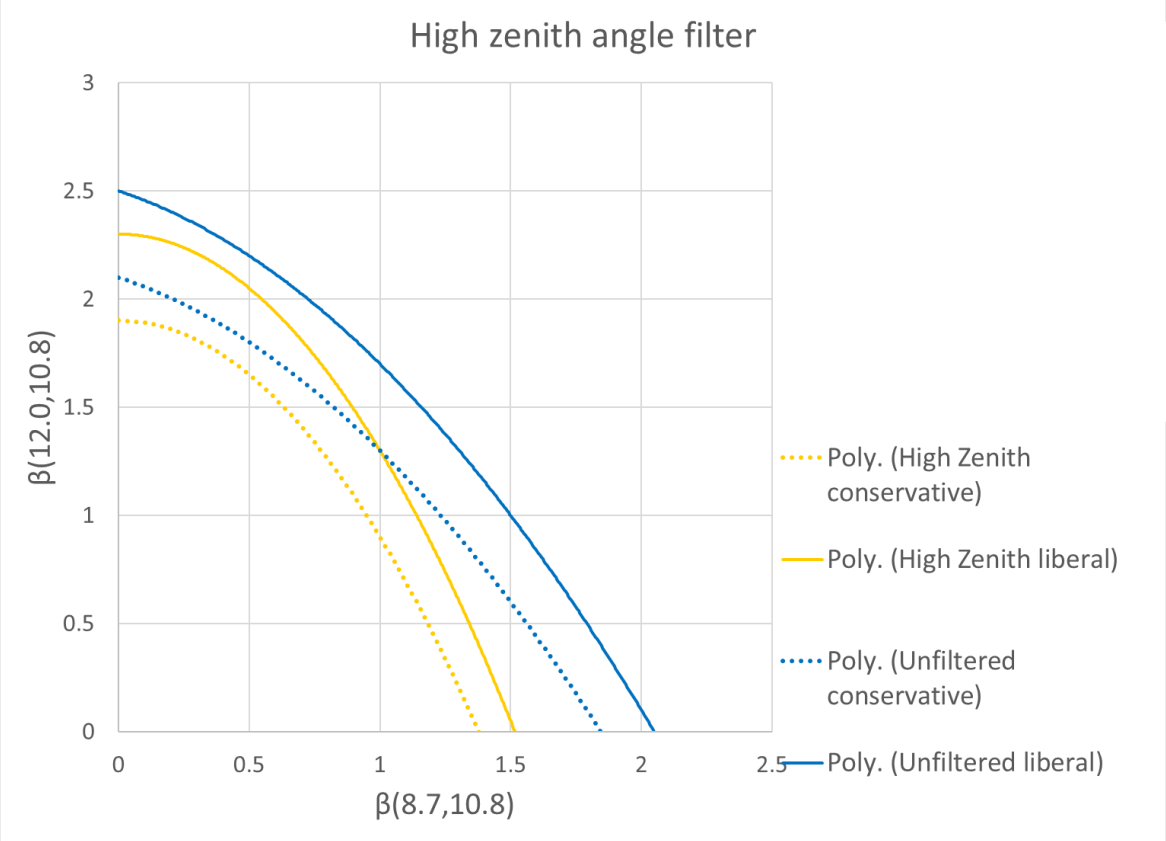

Figure 9 *Comparison of the beta space for unfiltered regions (blue) with the beta space for the high zenith angle regions (yellow).*

The reason a high zenith angle filter is required is that for pixels towards the edge of the satellites field of view, the path length through the atmosphere from the surface is increased. This means that BTD's of larger magnitude are more common, and it is therefore more likely that the thresholds used to mark pixels as ash will be exceeded. Figure 10 is a simple diagram demonstrating this effect.

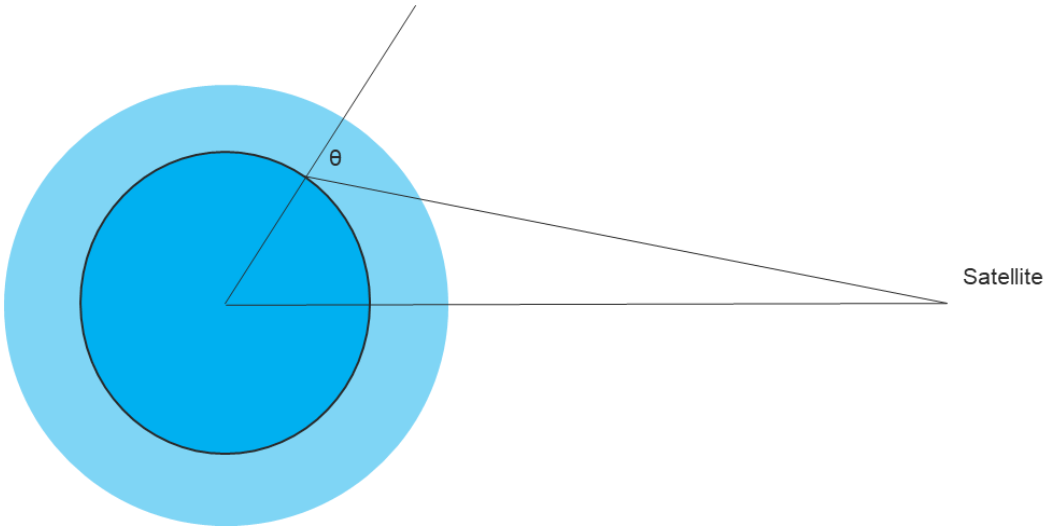

Figure 10 *Graphical representation demonstrating how the satellite zenith angle increasing towards the edge of the field of view results in an increased path length through the atmosphere.*

3.3.4 Low latitude filter

The low latitude filter is the least strict of all the regional filters. It was introduced in order to counter false detection seen over optically thick clouds with high altitude cloud tops. These are most common in the tropics and as a result this weaker filter was introduced in regions of low latitude. The acceptable beta space in Figure 11 is therefore used in regions with latitude less than ±15°.

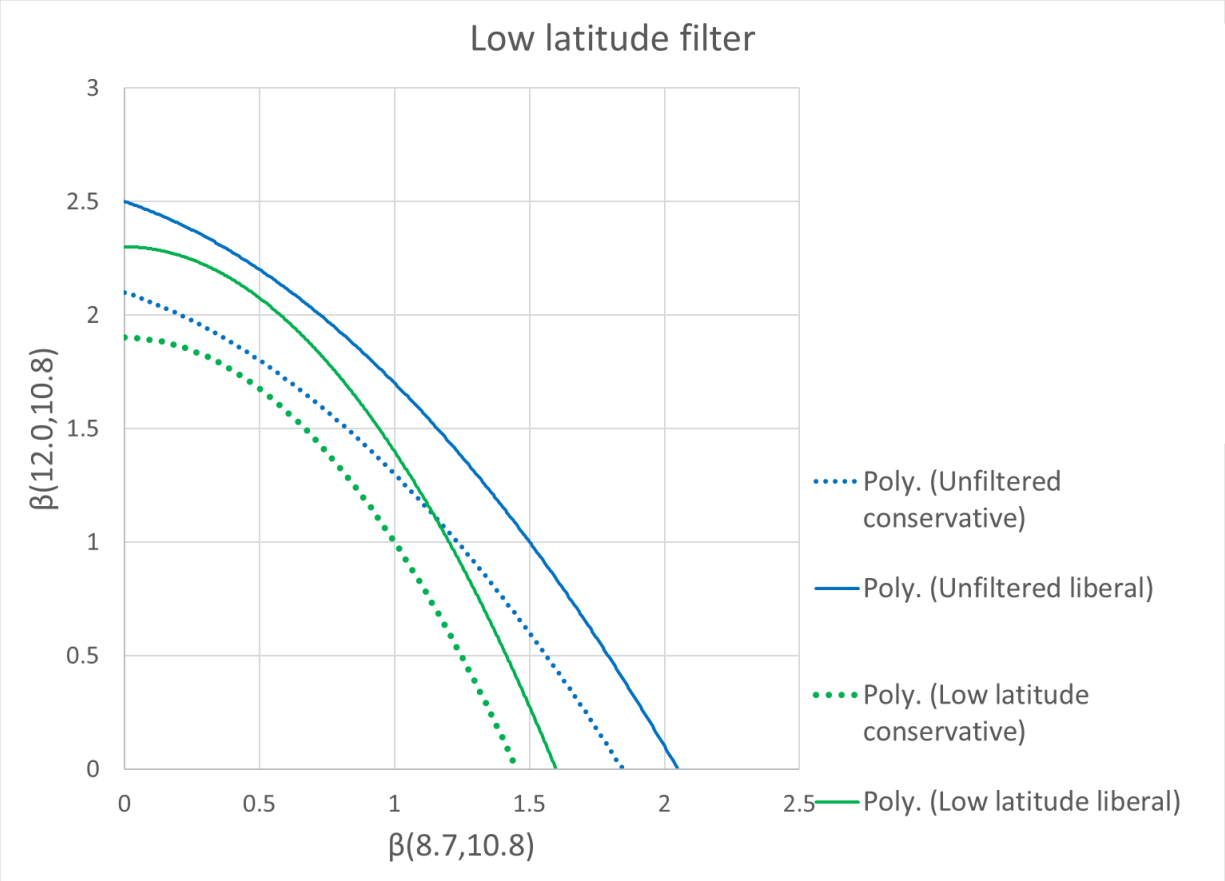

Figure 11 *Comparison of the beta space for unfiltered regions (blue) with the beta space for the low latitude regions (green).*

3.3.5 MODIS High zenith angle filter

The high zenith angle filter had to be modified for MODIS, which isn't too surprising as the instrument is aboard polar satellites rather than a geostationary satellite. As a result, the edge of the MODIS disk is still looking at the Earth, rather than space which is the case for the geostationary satellites. Therefore, the spatial filter, described in more detail in section 3.4, does not act on the edge 5 pixels of the MODIS field of view. This means that false detection is much more likely in these regions, especially at the areas of the MODIS field of view which are at high zenith angle (The MODIS scanning method means that the zenith angle only varies significantly in the cross-track direction). Additionally, the edge of the MODIS field of view is at ~66° so, in order to more strictly filter the higher satellite zenith angle pixels that are not filtered with the spatial filter, the high zenith angle for MODIS is set to 62.5°. The combination of the lack of a spatial filter and the high zenith angle effect means the strict filter seen in Table 3 and Figure 12 is required to remove the majority of false detection. The same beta space region is used for this regional filter in the spatial retest as there is no benefit to having a stricter filter. This is because the first test needs to remove the false detection for the edge pixels where the spatial filter will not act anyway. It is also worth noting that it is possible increased false detection may be seen at the edges of the MODIS field of view that are not at high zenith angle, as these pixels do not have a strict beta space and the spatial filter will not act on them. However, this effect was not significant in testing, so it wasn't considered worthwhile to code a more intelligent method to deal with the edge pixels.

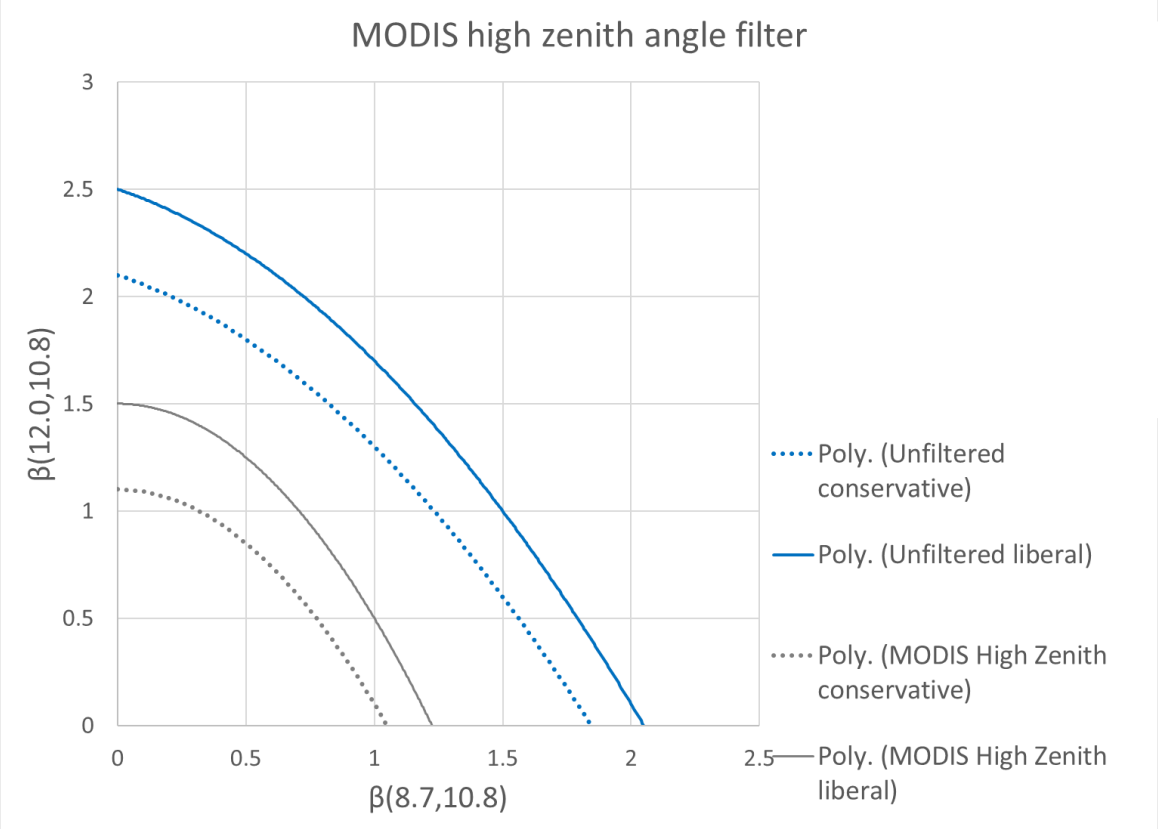

Figure 12 *Comparison of the beta space for unfiltered regions (blue) with the beta space for the high zenith angle regions for MODIS (grey).*

3.4 Spatial filter

The spatial filter was introduced with the intent of increasing the detection of ash which produces a weak BTD signal nearby to ash producing a stronger BTD signal. This is because it is more likely for ash to be present in a pixel when the neighbouring pixels are indicating the confident detection of ash. However, it is practically easier to implement a system where the likelihood of detecting ash is increased across all pixels and these pixels are then subsequently filtered based on the amount of ash detected in neighbouring pixels. As a result, the spatial filter works by considering the volcanic ash confidence output from the first tests in the pixels within an 11x11 box (square kernel) surrounding the relevant pixel. Where the mean volcanic ash confidence within this box does not exceed a given value the pixels are then retested with stricter thresholds. The stricter thresholds used include stricter beta spaces (particularly in regions that were not previously tested with the regional filters) and a reduction in each of the confidence threshold vector elements of 0.6 K. The beta spaces and clearskycutoff used in each region of the spatial filter are given in Table 4. The average volcanic ash confidence in the 11x11 box is weighted due to the significantly greater likelihood of false detection of confidence 4 and below. As a result, a factor of 3 weighting is applied to confidence 5 and greater, and the required weighted average to not retest is 3.0. This was found by consideration of the best balance of maximising correct detection whilst minimising false detection for ash cases from MSG and Himawari. It is also worth noting that the average only includes pixels that are looking over Earth to ensure that RMDI values on pixels looking at space don't cause issues in the averaging. A similar modification was also made to the existing median filter to ensure this issue didn't occur.

Region	aa	bb	cc	Clearskycutoff (K)
Unfiltered	-0.4	-0.4	1.9	-5.0
Low latitude	-0.9	0	1.9	-5.0
High Zenith angle	-1.0	0	1.9	-5.0
High Zenith angle (MODIS)	-1.0	0	1.5	-5.0
SH Arid Surface	-1.0	0	1.6	-10.0
NH Arid Surface	-1.0	0	1.3	-25.0

Table 4. Variables that are varied in the different regional filters as part of the spatial filter retests. aa, bb and cc define the beta space for each filter.

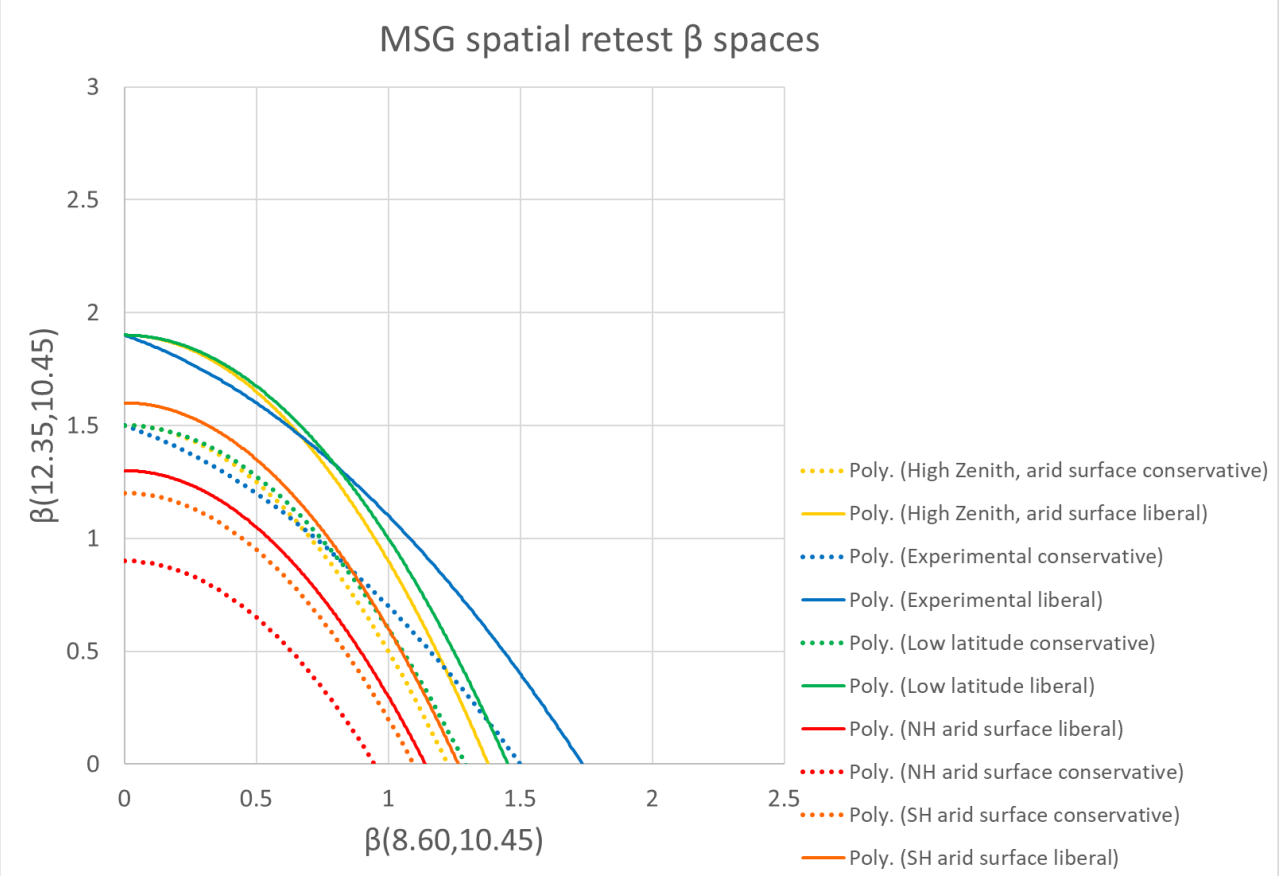

Figure 13 *Comparison of the beta space for the spatial filter retest for: unfiltered regions (blue), low latitude regions (green), high zenith angle regions (yellow), SH arid surface regions (orange) and NH arid surface regions (red).*

4. Verification of improvements

The verification of the changes to the volcanic ash detection was a multi-step process. Initially as the changes were being developed, there was ongoing subjective verification by visually comparing the effect of any changes on the amount of correct and false detection. Once the settings were finalised an extensive visual verification was completed against the operational method to ensure that the changes were in general an improvement. Examples of the results of the visual verification can be seen in sections 4.1.1, 4.2 and 4.3. Then in order to produce a statistical measure of the improvement from the new method, comparisons of the current operational method and the new experimental method against 'expert detections' were completed. These 'expert detections' are images where the pixels containing ash have been subjectively marked by an expert based on the Dust RGB imagery.

4.1. MSG verification

The verification for MSG was both visual and statistical. Section 4.1.1 contains two examples demonstrating the type of improvements that the experimental detection method can produce. Due to the limited nature of existing 'expert detections', comparisons to these could only be completed for MSG and these are detailed in section 4.1.2.

4.1.1 Visual Verification

The most straightforward method for comparing the performance of two detection methods is to simply visually compare the pixels marked as ash and make a subjective judgement on which is performing better (i.e which has correctly detected the most ash and has the least false detection). The main advantage of this method is that it can be done quickly for many cases without the requirement to complete lots of preparatory work. Whereas to complete the statistical analysis in section 4.1.2 across a larger selection of cases, a large amount of work would be required to produce more ‘expert detections’. It is also possible to identify more specific changes in the quality of the detection that may not show up in a simple statistical measure. As a result, a visual verification of the settings was completed for many cases from multiple eruptions and across a large time period, which encompassed as many conditions as were available in the test data. For MSG images from the Eyjafjallajökull, Grimsvotn, Puyehue and Calbuco eruptions were used. Examples from this verification can be seen in Figures 14 and 15. In Figure 14, the impacts of the spatial filter are clear with significant increases in correct detection of ash, particularly over the UK, as well as some regions where narrow regions of ash have reduced detection between the UK and Iceland. The most notable feature in Figure 15 is the reduction in false detection due to the regional filters, but there are also small increases in detection of the ash plume.

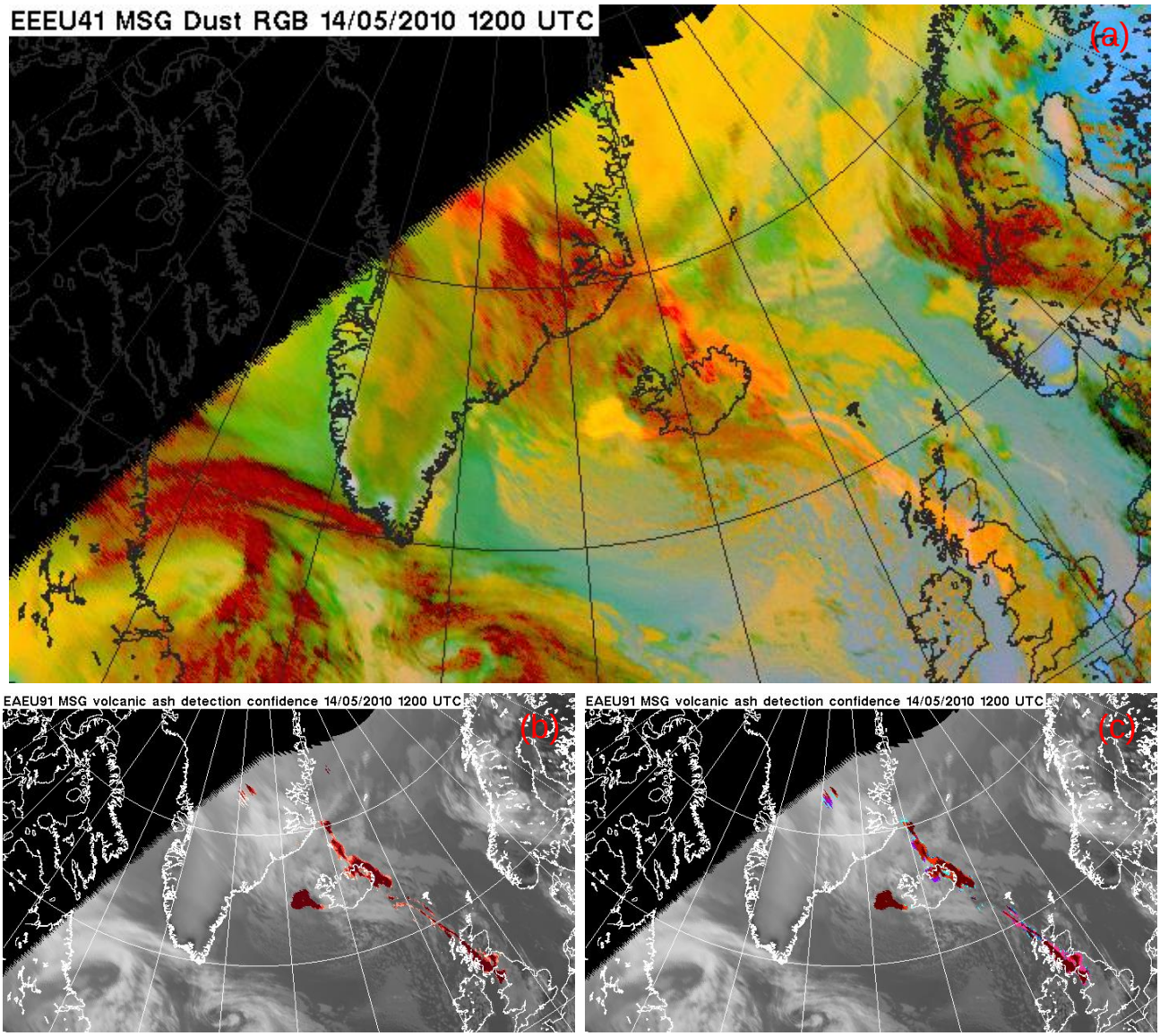

Figure 14 *Dust RGB (a), old operational volcanic ash confidence imagery (b), new experimental volcanic ash confidence imagery (c) at 1200 (14/05) following the 2010 Eyjafjallajökull eruption.*

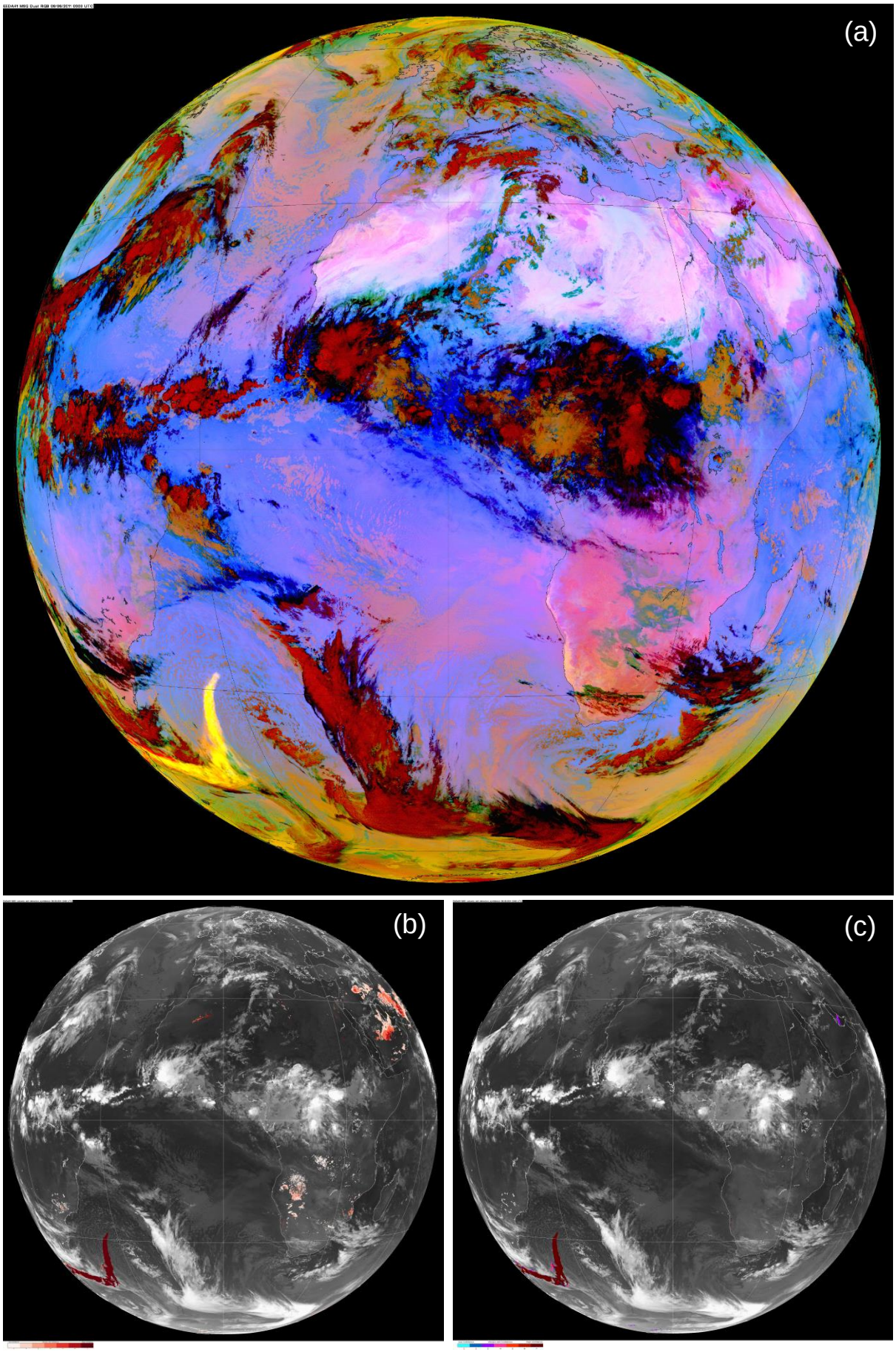

Figure 15 *Dust RGB (a), old operational volcanic ash confidence imagery (b), new experimental volcanic ash confidence imagery (c) at 0300 (06/06) following the 2011 Puyehue eruption.*

4.1.2 Objective verification against expert detections

In order to statistically verify the improvement from the experimental settings over our current operational detection, objective comparisons to 'expert' detections, which were treated as ‘truth’, were completed. These images, where pixels have been marked as containing ash by a person, are still considered the most accurate method for identifying ash contaminated pixels. The result of this was that, over the 18 test cases, the experimental detection resulted in an 11% increase in correct detection and an 85% reduction in false detection. The 18 test cases were all from MSG and contained 1 case from Puyehue, 7 cases from Grimsvotn and 10 cases from Eyjafjallajökull. Four of the expert detections were produced by Mike Cooke and 14 were produced at NOAA. The methodology behind the expert detections produced by Mike and those produced at NOAA is different. The expert detections produced by Mike contain pixels that have been individually marked as ash meaning they are highly likely to contain ash. Whereas in the NOAA expert images a larger region is enclosed including all areas with a high likelihood of ash and some neighbouring areas which might be ash free. As a result, a perfect score for these detections will be much less and has significant variance from as low as ~1% to as high as ~50%. In this case this difference is not too important as the relative score for both the operational method and new experimental method will be unaffected, but it highlights an important point that statistical measures of an ash detection algorithm are highly dependent on the verification dataset.

I have included examples of both a NOAA expert detection (Figure 16(a)) and one provided by Mike Cooke (Figure 17(a)), as well as the verification scores produced for both the operational and experimental detection for each (Table 5). The time and location of all 18 cases used to calculate the average figures is given in Appendix A along with their verification scores. The correct detection percentage refers to the percentage of pixels identified as ash by the expert detection that are also marked as ash by the automated algorithm. The pixels marked as ash are shown as white pixels in Figure 17(a) and are all the pixels within the green line in Figure 16(a). The false detection percentage is the percentage of total pixels in the image not marked as ash by the expert detection that are marked as ash by the automated algorithm. The percentage changes in these numbers are then calculated for the correct detection and false detection change columns.

Time	Operational correct detection (%)	Experimental correct detection (%)	**Correct detection change (%)**	Operational false detection (%)	Experimental false detection (%)	**False detection change (%)**
0400 (08/05/2010)	22.3	25.5	+14.3	0.408	0.097	-76.2
1230 (07/05/2010)	90.7	95.4	+5.2	0.273	0.030	-89.0

Table 5. Verification scores for the expert detection examples in Figures 16 and 17.

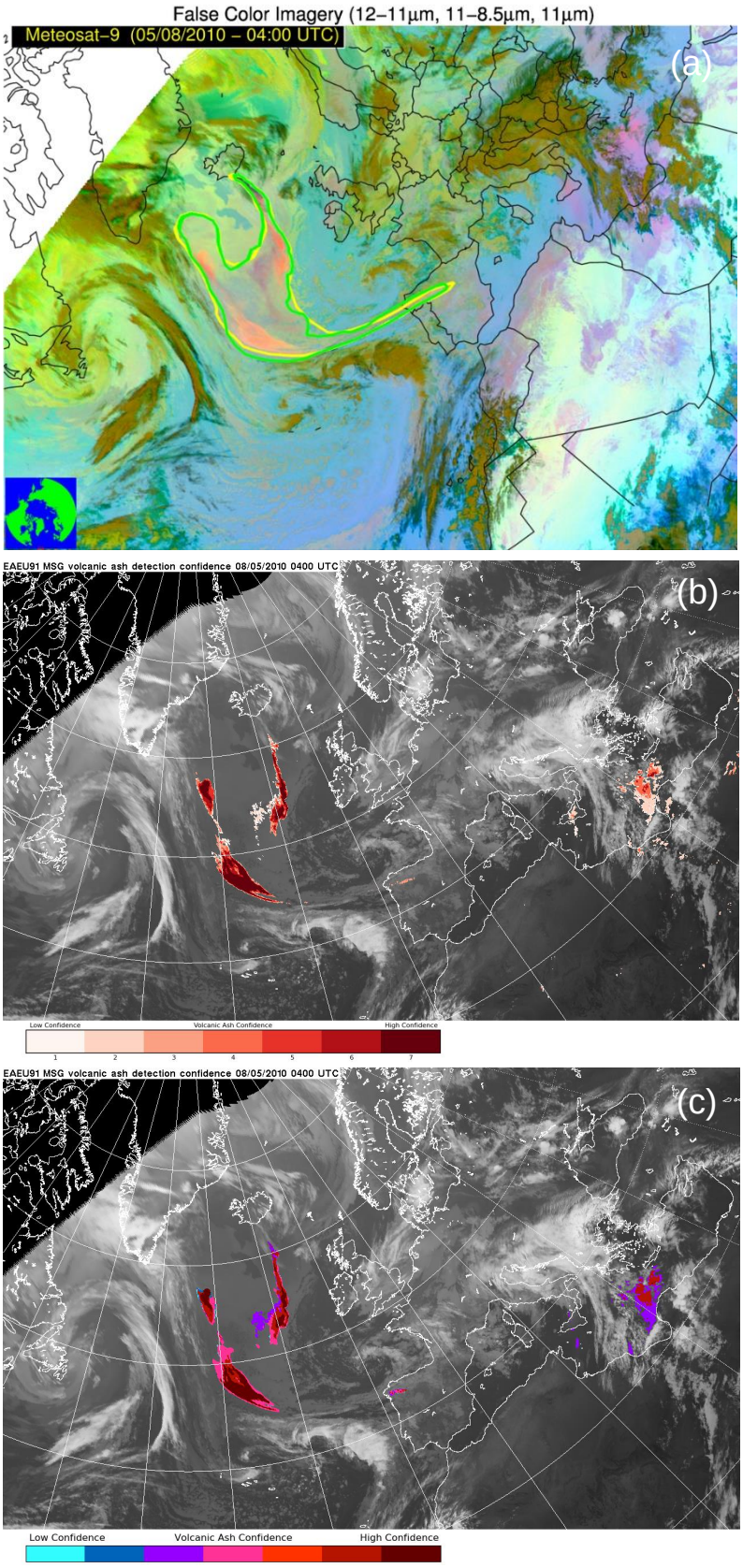

Figure 16 *Dust RGB with NOAA expert detection in green and yellow (a), old operational volcanic ash confidence imagery (b), new experimental volcanic ash confidence imagery (c) at 0400 (08/05) following the 2010 Eyjafjallajökull eruption.*

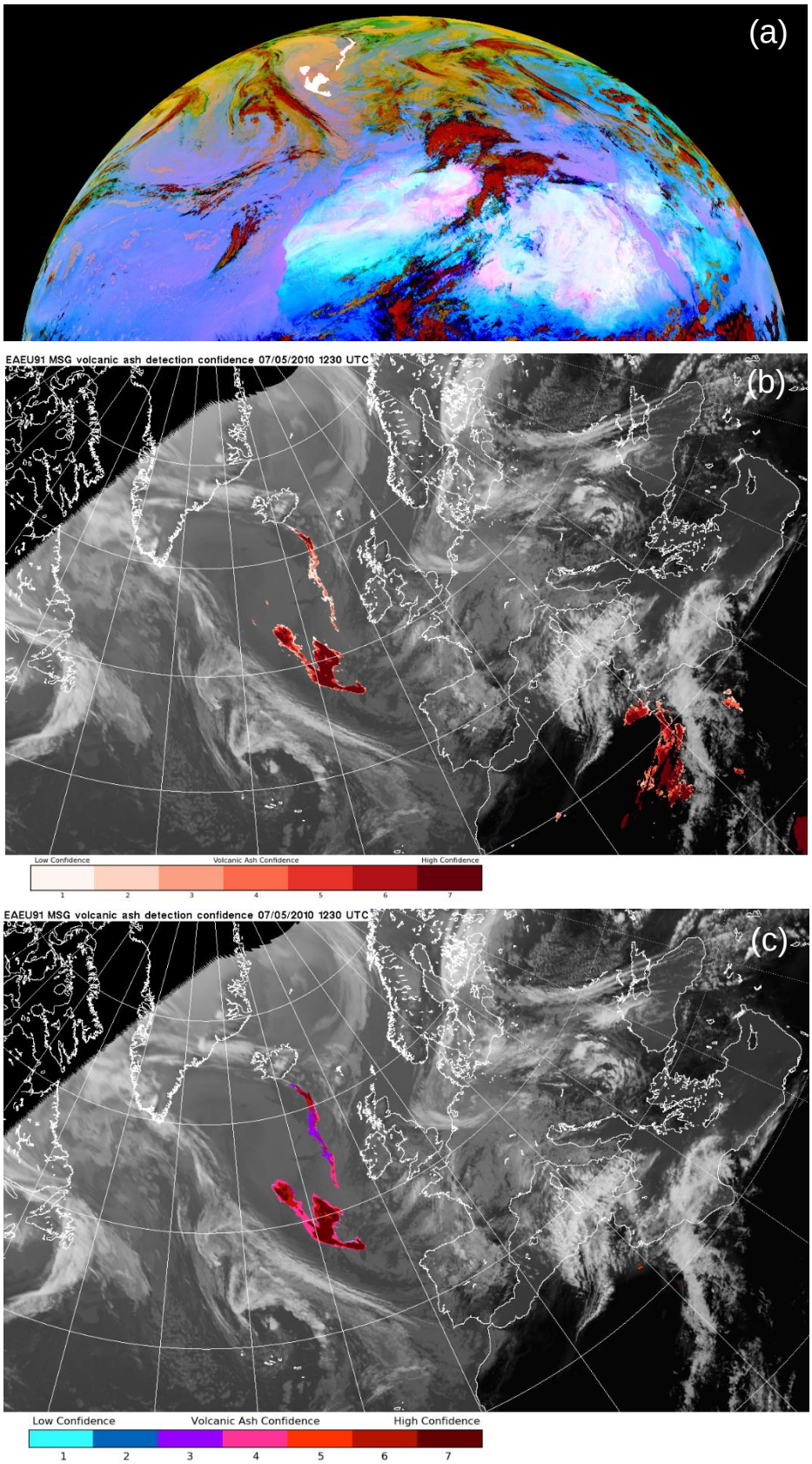

Figure 17 *Dust RGB with Mike Cooke's expert detection overlaid in white (a), old operational volcanic ash confidence imagery (b), new experimental volcanic ash confidence imagery (c) at 1230 (07/05) following the 2010 Eyjafjallajökull eruption.*

4.2. Himawari verification

For Himawari a similar process of visual verification was used as the one described in section 4.1.1 for MSG. In this case imagery used was primarily from the Raikoke eruption, although the November 2021 eruption of Karymsky produced some useful additional test data that allowed for some threshold optimisation for Himawari. An example of a test case used extensively for Raikoke along with the detection produced by the old operational and new experimental algorithms is shown in Figure 18.

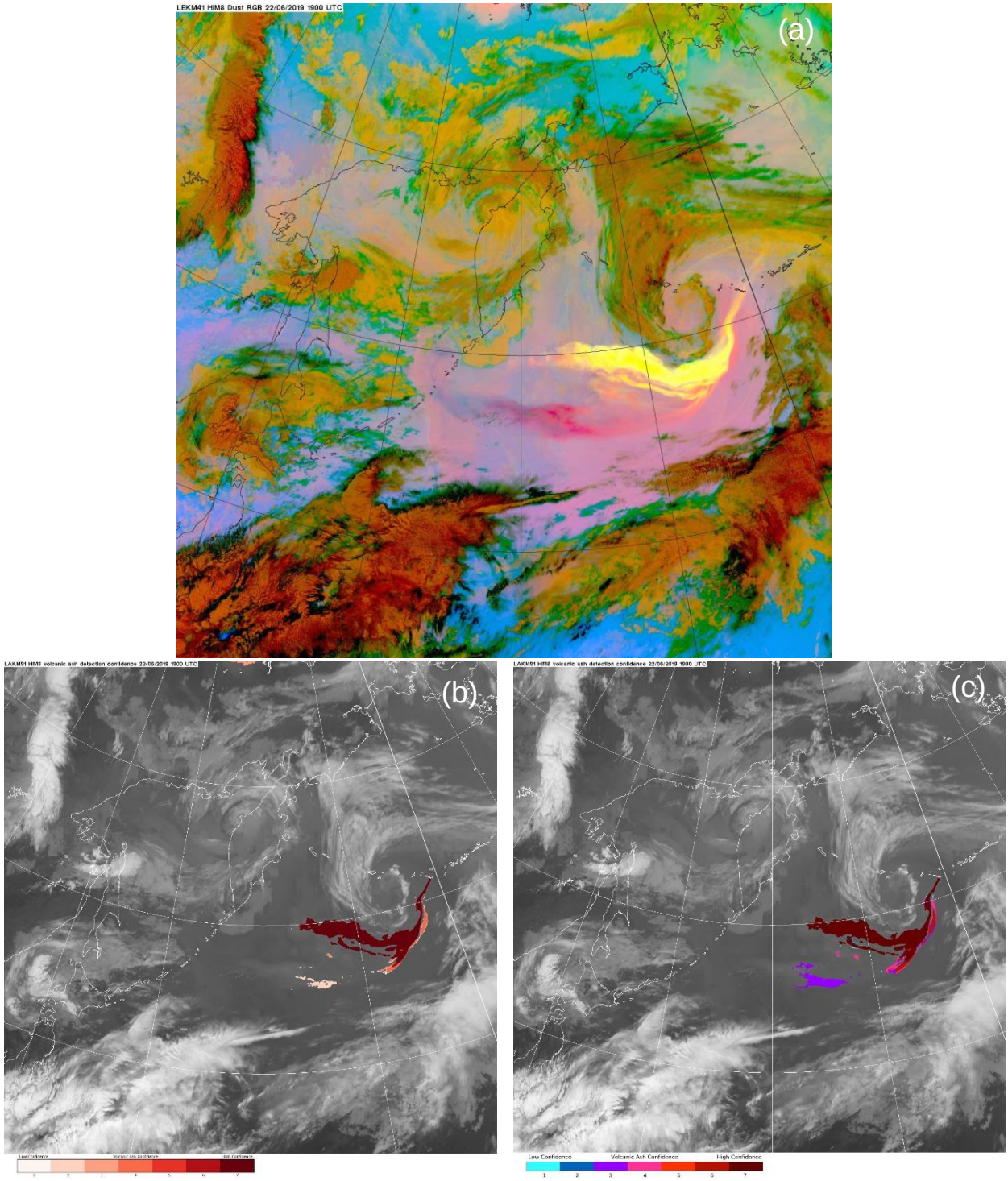

Figure 18 *Dust RGB (a), old operational volcanic ash confidence imagery (b), new experimental volcanic ash confidence imagery (c) at 1900 (22/06) following the 2019 Raikoke eruption.*

4.3. MODIS verification

The MODIS verification was a briefer assessment comparing the old operational and new experimental method with the 4 cases from Eyjafjallajökull and Grimsvotn that were available. This assessment led to the modification of the MODIS high zenith angle filter to counter significant false detection seen at the very edge of the MODIS field of view along the high satellite zenith angle sides. An example of MODIS imagery from the Eyjafjallajökull eruption along with the output imagery from the old operational and new experimental methods are shown in Figure 19.

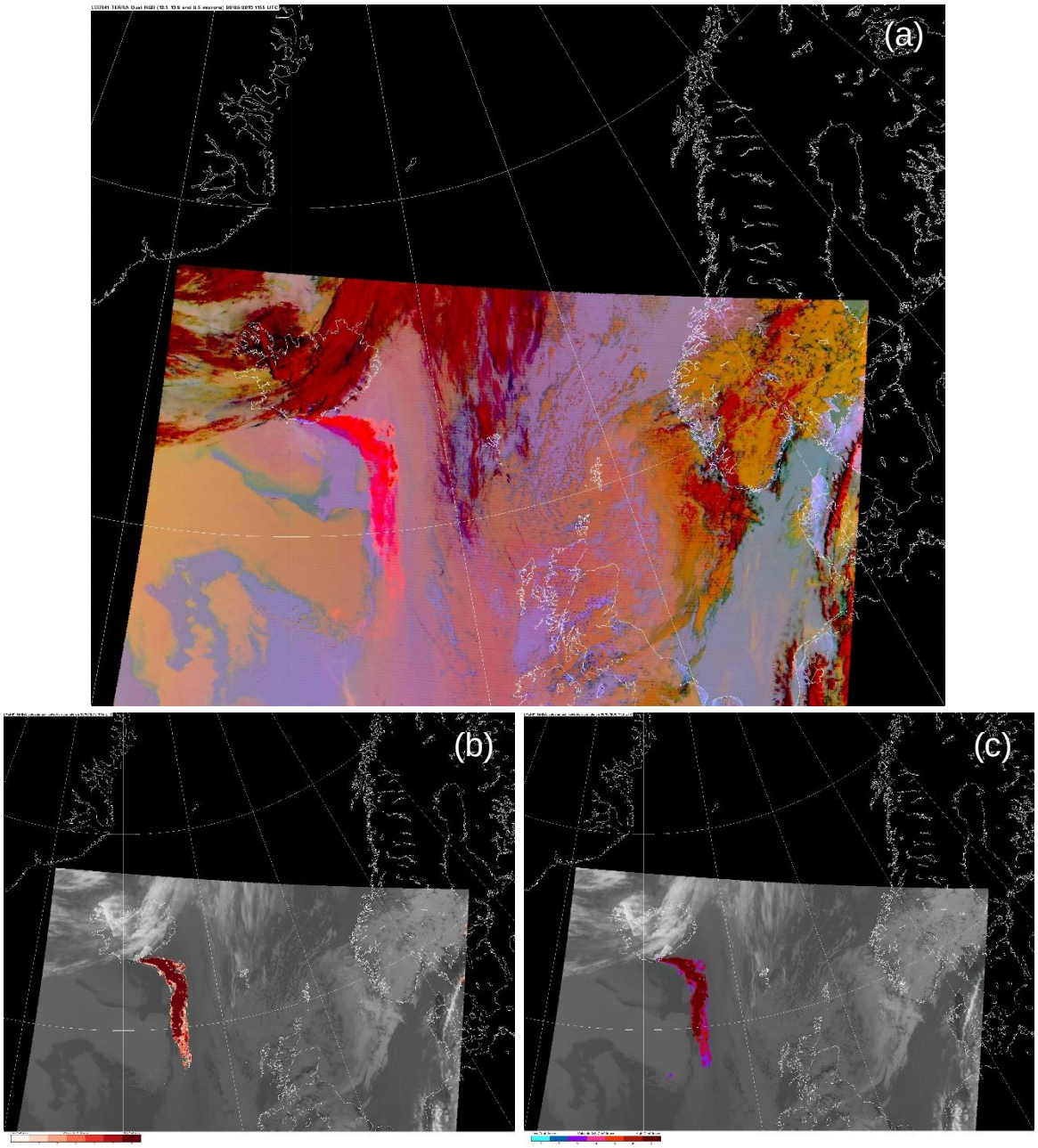

Figure 19 *Dust RGB (a), old operational volcanic ash confidence imagery (b), new experimental volcanic ash confidence imagery (c) at 1155 (06/05) following the 2010 Eyjafjallajökull eruption.*

4.4. False detection assessment

After completing extensive testing, the new experimental algorithm was run on the development server for a few months in order to ensure it was not causing any unexpected issues and to compare the performance to the operational method. As volcanic eruptions are fairly infrequent, only an assessment of false detection could be completed in this way. Initially, this revealed a minor issue for polar satellites that use the simplified 2-channel BTD alone. This was resolved and adjustments were made to both MODIS and GOES satellites to counter increased false detection seen in some areas. This was primarily caused by the increases to the BTD_Cutoff. Then regular checks of the imagery produced by each algorithm were completed using the 'Product comparison' tool. Ideally, assessing the change in false detection from the old method to the new method would be completed in a more quantitative manner. Unfortunately, in the old code there was no count of the total number of pixels detected (which in most cases will be a count of the falsely detected pixels as real volcanic ash is rare). This meant that although this count now exists it wasn't possible to make a comparison as the number of pixels of false detection from the old method wasn't available. As the new code does have this count, future changes can have the change in false detection assessed statistically in a more wide-ranging manner than is available with the 'expert detections'. However, the 'Product comparison' tool did allow for a straightforward comparison of the false detection from the operational method and the new experimental method and produced a few noteworthy conclusions:

- MSG false detection was significantly reduced overall as expected. In particular, reductions of false detection over filtered arid regions were significant. However, some false detection is still present regularly. This false detection is primarily from meteorological cloud tops and was present (and more significant) in the previous operational method as well. In general, this false detection is more common in the tropics, but it is also present on high cloud tops in the midlatitudes (e.g on frontal systems). Occasionally this false detection is significantly reduced by the new experimental algorithm when the false detection present in the old operational method is very speckly and therefore similar detection is likely to be filtered by the spatial filter in the new experimental method.
- Himawari false detection is also significantly reduced. Primarily this is over arid regions of Asia and at high zenith angle. There were very rare occasions when new false detection was present in the new experimental method, although there was still a decrease in false detection across the whole disc. Overall there was significant improvement from the new experimental algorithm, and it is likely that any changes to reduce the rare occurrences of new false detection would result in decreased capability to correctly detect ash, given that these settings were optimised with cases containing real ash.
- For GOES the impact on false detection of the new experimental algorithm is more mixed, although still an improvement overall. As with Himawari and MSG, false detection is reduced over arid regions and at high zenith angle. However, the false detection present at high zenith angle for both GOES satellites was already more significant with the old operational method than for Himawari and MSG. Therefore, it is perhaps not surprising that relatively significant false detection at high zenith angle remains in the imagery produced by the new experimental method. In future, it may be worth more strictly filtering pixels as zenith angle increases (within the high zenith angle region that is filtered in the new experimental method). However, the optimum

way to do this would be with a smoothly varying scale that should be straightforward to produce with a probability-based method. Therefore, as a future change should allow for a more optimum method of dealing with this issue and as GOES imagery is not an operational priority, I have not experimented with the addition of further high zenith angle filtering for GOES. Small regions of new false detection are more common in the GOES imagery than the Himawari imagery. As with the Himawari imagery, overall false detection is still reduced in these cases, but ideally optimisation of the GOES settings could be completed to decrease the frequency of these cases. However, without real ash test data available it is difficult to accurately assess what impact modification of certain thresholds would have on correct detection of ash. Therefore, as the overall impact on false detection of the new experimental method is positive, I have not made any further changes.

- False detection was already fairly common on MODIS imagery with the old operational method. Overall, this has been reduced in the new experimental method imagery. Much of this reduction comes from the very strict high zenith angle filter added for MODIS. As explained in section 3.3.5, a more intelligent method of countering this issue for the spatial filter on MODIS imagery would be ideal, particularly as these contribute to the required strictness of the high zenith angle filter. Over the testing period there was little evidence of significant false detection at the edges of the MODIS field of view which are not at high zenith angle. There is a small reduction in false detection when not at high zenith angle with the more speckly false detection seen in the old operational imagery often removed.
- The change in colour scale for the confidence imagery means that lower confidence detection is now much easier to see, particularly over cloud tops. This means that in some areas it may appear that the new experimental method results in more false detection over cloud tops than was seen before with the old operational method. This is not the case, although occasionally new false detection appeared in these areas, much more false detection present in the old operational method was removed. An additional benefit of the colour scale change is that it is easier to identify the confidence assigned to a given pixel which is useful for developmental purposes.

5. Summary and possible future work

5.1. Summary

In summary, the new experimental method has been shown to be an improvement over the old operational method and will therefore replace this as the operational method. Reductions in false detection are seen across all of the relevant satellites. These are most significant for the Himawari and MSG satellites, which is not surprising given that these satellites provided the vast majority of the test data used to optimise the new algorithm. Most of the false detection reduction is as a result of the new location-based filters, in particular the arid surface filters. The reduction of this false detection may not be hugely important to InTEM as it is very possible InTEM would already be treating these false detections as not useful information (due to their distance from the source). However, they still result in a measurable improvement in the algorithms ability to accurately identify regions with and without ash. Some further reduction in false detection is achieved through the use of the spatial filter removing some speckled regions of false detection. Again, the impact on InTEM's performance of this change may not be significant, as the re-gridding process may effectively remove some of this ash detection before it is seen by InTEM. However, it is likely

that some of the removal of false detection will be beneficial, given the reduction in false detection shown in the expert detection analysis of 85%. Additionally, this allows users of the imagery to have more confidence that apparent detections of ash will be correct. The spatial filter also increases correct detection in some areas where ash producing a marginal spectral signal is close to ash with a clearer spectral signal. This effect is clearest in the expert detection analysis which showed an 11% increase in correct detections, although the increase is also visible in some imagery.

Despite the improvements there are some regions where the new algorithm does not perform better than the old algorithm. These include narrow plumes of ash producing a weak spectral signal, certain regions of GOES imagery, and the edge of the MODIS field of view. Possible work to overcome some of these issues and produce further improvements in the detection algorithms performance are discussed in section 5.2.

5.2. Possible Future Work

Here ideas for further work to improve the detection algorithm are briefly detailed:

- Further GOES testing if examples with ash occur. This will allow for an optimisation to assess whether some of the remaining false detection can be removed without significantly reducing the potential to correctly detect ash.
- More elegant solution to the MODIS edge pixel problem. Ideally the spatial filter would act all the way to the edge of the field of view to allow the high zenith angle filter to be less strict.
- Move to a probability-based method rather than the regional filters and spatial filters. This will allow for a more flexible approach to the filtering. For example, the arid surface filter can be stricter in areas where false detection is much more probable. Additionally, the high zenith angle filter can effectively become increasingly strict as the zenith angle increases. It also allows for a probability of ash just from the spectral signal to be produced, which may be useful in the event of an eruption in a region where false detection is more common.
- Production of a standard suite of ‘expert detections’, including data from satellites other than MSG. This could allow for a standard assessment of any changes to the detection, and potentially allow for a statistical method to optimise certain thresholds within the algorithm. This is important to ensure that any future changes can be justified objectively rather than relying on subjective assessments during the development process.

6. References

Ashpole, I., and R. Washington (2012), An automated dust detection using SEVIRI: A multiyear climatology of summertime dustiness in the central and western Sahara, J. Geophys. Res., 117, D08202, doi:10.1029/2011JD016845.

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.

Mackie, S. and Watson, M.: Probabilistic Detection of Volcanic Ash Using a Bayesian Approach, J. Geophys. Res. Atmos., 119, https://doi.org/10.1002/2013JD021077, 2014.

Pavolonis, M. J. (2010), Advances in extracting cloud composition information from spaceborne infrared radiances: A robust alternative to brightness temperatures. Part I: Theory, J. Appl. Meteorol. Climatol., 49, 1992–2012, doi:10.1175/2010JAMC2433.1.

Pavolonis, M. J., Sieglaff, J., & Cintineo, J. (2015b). Spectrally enhanced cloud Objects—A generalized framework for automated detection of volcanic ash and dust clouds using passive satellite measurements: 2. Cloud object analysis and global application. Journal of Geophysical Research: Atmospheres,120, 7842–7870. https://doi.org/10.1002/2014JD022969

Piontek, D.; Bugliaro, L.; Schmidl, M.; Zhou, D.K.; Voigt, C. The New Volcanic Ash Satellite Retrieval VACOS Using MSG/SEVIRI and Artificial Neural Networks: 1. Development. Remote Sens. 2021, 13, 3112. https://doi.org/10.3390/rs13163112

Saint, C., Raikoke Eruption: Volcanic ash plume tracking, detection and retrieval, Met Office (2020), Satellite Applications Technical Memo n.110.

7. Appendices

Time	Operational correct detection (%)	Experimental correct detection (%)	Correct detection change (%)	Operational false detection (%)	Experimental false detection (%)	False detection change (%)
Eyjafjallajökull (1900) 06/05/2010	79.5	82.7	+4.0	0.107	0.006	-94.8
Eyjafjallajökull (1200) 07/05/2010	90.7	95.4	+5.3	0.273	0.030	-89.1
Grimsvotn (1000) 23/05/2011	97.7	97.7	+0.0	0.214	0.031	-85.4
Puyehue (0900) 06/06/2011	84.1	87.9	+4.6	0.337	0.099	-70.7
Eyjafjallajökull (0400) 08/05/2010	22.3	25.5	+14.1	0.408	0.097	-76.1
Eyjafjallajökull (1400) 11/05/2010	2.16	2.78	+28.8	0.198	0.012	-93.7
Eyjafjallajökull (0330) 12/05/2010	1.40	2.26	+61.2	0.667	0.034	-94.9
Eyjafjallajökull (1300) 13/05/2010	3.27	3.28	+0.3	0.130	0.005	-96.1
Eyjafjallajökull (1230) 14/05/2010	26.8	31.9	+19.0	0.187	0.011	-94.0
Eyjafjallajökull (0300) 16/05/2010	12.8	14.8	+16.0	0.212	0.028	-86.7
Eyjafjallajökull (1130) 17/05/2010	11.0	12.7	+15.5	0.358	0.002	-99.4
Eyjafjallajökull (0730) 18/05/2010	4.23	4.20	-0.7	0.421	0.109	-74.0
Grimsvotn (0000) 22/05/2011	0	0	0.0	0.295	0.059	-80.0
Grimsvotn (0200) 22/05/2011	0	0	0.0	0.299	0.060	-79.9
Grimsvotn (0400) 22/05/2011	0.400	2.40	+500.0	0.359	0.069	-80.7
Grimsvotn (0600) 22/05/2011	0.504	1.20	+138.9	0.228	0.048	-79.2
Grimsvotn (0800) 22/05/2011	0.939	1.54	+64.3	0.136	0.033	-75.5
Grimsvotn (1300) 22/05/2011	3.07	6.08	+98.3	0.171	0.033	-80.4

Appendix A. Verification scores for all 18 expert detections used to produce statistics.