Automated Stereo Retrieval of Smoke Plume Injection Heights and Retrieval of Smoke Plume Masks From AATSR and Their Assessment With CALIPSO and MISR

Daniel Fisher, Jan-Peter Muller, and Vladimir N. Yershov

***Abstract*— The longevity and dispersion of smoke and associated chemical constituents released from wildfire events are dependent on several factors, crucially including the height at which the smoke is injected into the atmosphere. The aim here is to provide improved emission data for the initialization of chemical transport models in order to better predict aerosol and trace gas dispersion following injection into the free atmosphere. A new stereo-matching algorithm, named M6, which can effectively resolve smoke plume injection heights (SPIH), is presented here. M6 is extensively validated against two alternative spaceborne earth observation SPIH data sources and demonstrates good agreement. Further, due to the spectral and dual-view configuration of the Advanced Along-Track Scanning Radiometer imaging system, it is possible to automatically differentiate smoke from other atmospheric features effectively—a feat, which currently no other algorithm can achieve. Additionally, as the M6 algorithm shares a heritage with the other M-series matchers, it is here compared against one of its predecessors, M4, which, for the determination of SPIH, M6 is shown to substantially outperform.**

***Index Terms*— Advanced Along-Track Scanning Radiometer (AATSR), injection height, smoke plume, stereo.**

I. Introduction

BIOMASS burning events in boreal forests generate significant amounts of important greenhouse gases, including CO_2, CO, and NO_x [1], [2]. Further, burning events may potentially convert these ecosystems from carbon sinks to net sources, in turn contributing to global warming [3]. Through the process of convection induced by heat and moisture released during fire events, vegetation fires can launch their emissions vertically. The vertical extent of the emissions is dependent upon the energy released by the fire, the moisture content of the fire stock, and the ambient meteorological conditions [4]. For the majority of observed fires, in a case study of fire events in North America, emissions appear to be trapped in the atmospheric boundary layer leading to localized deposition [2]. However, in sufficiently large fire events with large energy outputs or under strongly convective atmospheric conditions, the emissions can be injected into the free troposphere [2]. Under these conditions, the atmospheric lifetime of most trace gases and aerosols is substantially enhanced and their effects last longer thus affecting much larger areas [5]. In some cases, the smoke particles can persist in the free troposphere for several days to weeks and so contribute to the dispersal of greenhouse gases and other pollutants thousands of kilometers away from their sources [5].

To quantify the impacts of fire emissions on air quality and climate, chemical transport models (CTMs) are employed [6]. In order to generate effective predictions and better understanding, these CTMs must be initialized with reliable estimates of the vertical extent of the emissions, and also validated against plume dispersion measurements over time. Due to the lack of available *in situ* or satellite data, particularly at higher latitudes than visible to geostationary satellites, often rather arbitrary assumptions are used for initialization, such as fixed vertical injection levels [3]. This simplification is likely to lead to a reduction in the accuracy of the emission distribution outputs from such models. To improve CTM prediction accuracy, precise measurements of smoke plume injection heights (SPIH), which can be assumed as a proxy for all constituent emissions for fire events, were made for the European Space Agency (ESA) ALANIS Smoke Plumes (EASP) (http://www.alanis-smokeplumes.info) project. The EASP project aims to use this newly developed SPIH data source to potentially improve current large-scale dispersion forecasts of emissions from the Eurasian boreal forest ecosystem (latitudinal range 46°–78°, longitudinal range 0°–174°) from August 2008 to August 2011.

As the study region is vast, the use of earth observing (EO) satellite data is necessitated to provide suitable coverage. However, there are few suitable techniques for determining SPIH using EO. The Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) [7] carried onboard the joint National Aeronautics and Space Adminsitration and Centre Nationale d'Etudes Spatiales (NASA-CNES) satellite demonstrates one method; however, while the heights retrieved are very accurate, its spatial sampling is very limited (100-m footprints every 333-m along-track and around 500-km across-track) due to the lidar technology employed. A far more comprehensive coverage of SPIH is achieved in the Multiangle

Manuscript received July 16, 2012; revised December 11, 2012; accepted January 14, 2013. Date of publication March 26, 2013; date of current version December 12, 2013.

The authors are with the University College London, Dorking RH5 6NT, U.K. (e-mail: dnf@mssl.ucl.ac.uk; jpm@mssl.ucl.ac.uk; vny@mssl.ucl.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2013.2249073

Imaging SpectroRadiometer (MISR) Smoke Plumes project (http://misr.jpl.nasa.gov/getData/accessData/MisrMinxPlumes/) using the MISR [8] instrument, carried onboard the NASA Terra satellite, where the principles of stereo photogrammetry are applied. The method, developed at the Jet Propulsion Laboratory, the results of which are described in [2], relies on manual digitization of active smoke plume masks (SPM) and is therefore limited to those areas, which were studied to date. Here, we apply an automated stereo photogrammetric method to another stereo capable instrument, the advanced along-track scanning radiometer (AATSR).

AATSR [9], launched in 2002 onboard the ESA Envisat satellite, provides an excellent tool for the determination of SPIH. The instrument viewing geometry provides imagery at both nadir and in the forward direction (at 55° zenith angle), allowing for effective stereo reconstruction. Given AATSR, with a nominal nadir pixel resolution of 1 km, a swath of 512 km, and at least four orbits crossing the study region per day, a reasonable spatial sampling is achieved. Capable of imaging in stereo in seven channels, covering both the visible/near infrared, shortwave infrared and thermal (0.55, 0.65, 0.87, 1.6, 3.7, 11, and 12 μm), complete automation of the SPM process is possible.

A smoke plume "tuned" stereo-matching algorithm, referred to here as M6 due to a shared heritage with the other M-series matchers [10], [11], is developed for the EASP project. M6 utilizes novel normalization and matching techniques that perform effective SPIH determination in the study region and generate improved results, in terms of coverage and accuracy, when compared to a previous state-of-the art stereo-matching algorithm applied to AATSR data, e.g., M4 [11].

The next section describes the overall processing chain and output products developed for the EASP project. In Section III, a detailed algorithmic description of M6 is given, followed in Section IV by a description of the SPM method. Section V outlines the validation approach used for analyzing the SPIH outputs of M6, the results of which are given in Section VI. These results are discussed in Section VII and conclusions are drawn in Section VIII.

II. AATSR Processing Chain

A Java processing chain based upon the BEAM visualization toolkit (http://www.brockmann-consult.de/cms/web/beam/) is developed for the generation of the AATSR SPIH dataset for the EASP study region. The processing chain generates pixel-level precision SPIHs and, using AATSRs multispectral imaging capabilities, SPMs. The SPIHs are generated using the M6 matcher and the Mannstein camera model [11], [12]. The SPMs are then derived from the M6 output, the spectral variations between the input channels and a local digital elevation model (DEM) taken from the USGS GMTED2010 30 arc-seconds dataset [13]. The key stages of the entire processing chain can be summarized as follows.

1) Ingest Envisat AATSR N1 product and select the four relevant spectral channels (0.55 μm Forward and Nadir, 0.87 μm Nadir, 1.6 μm Nadir and 11 μm Nadir) as well as the relevant geo-referencing information, camera model, DEM, and co-registration correction coefficients [14].
2) Using the 11 μm channel generate a cloud mask using a thermal threshold. Dilate the cloud mask by one pixel, to ensure removal of cloud pixels at cloud edges, and apply to 0.55 μm channel prior to stereo processing.
3) Apply M6 to the cloud masked 0.55 μm channel views to produce a digital disparity model. Correct the disparities using the co-registration correction coefficients [14].
4) Use the Mannstein camera model to space intersect projected rays from the two views to find the point of closest perpendicular distance (minimum skewness) and provide elevation values above the reference ellipsoid for the disparity values to create a digital disparity and elevation model (DDEM).
5) Generate the SPM using the DDEM and a height threshold above a DEM to exclude any misclassifications. Apply the derived mask to the DDEM to generate the masked SPIH product.
6) The resultant masked SPIH product is written out in geoTIFF format using the geo-referencing information and then converted into NetCDFv4 (CF). In addition to the SPIH layer, the following additional layers are incorporated into the product: moderate resolution imaging spectroradiometer (MODIS) fire radiative energy of the nearest overpass; a false color composite browse product of the forward 0.55, 1.6, and 11 μm channels; and a red-cyan stereo anaglyph of the 0.55 μm (green) channel.

The total processing time to produce a geo-referenced SPIH output for a 512 $\times$ 512 pixel image subset is approximately 2 min on a single core of an Intel Core i5-processor running at 2.66 GHz. This processing chain is applied to the AATSR imagery for the defined study region and epoch, providing a new SPIH dataset for exploitation by the scientific community (accessible by application at http://www.alanis-smokeplumes.info).

III. M6 Stereo Matcher

Although M6 holds a shared heritage with the other M-series matchers [10], [11] in that it uses area-based techniques to solve the stereo reconstruction problem, it approaches the reconstruction with significant modifications, especially in the normalization procedure and the matching cost computation.

A. Normalization

Accounting for photometric variability (due for example to differences in aerosol scattering or aerosol or surface bi-directional reflectance) between stereo image pairs prior to disparity estimation is a standard pre-processing step for most area-based stereo-matching algorithms. M4 [11], for example, employs normalized cross correlation (NCC) where each image is modified so that it has local statistics comprised of a zero mean and a standard deviation equal to one. This increases robustness to changes in gain, bias, and local variability due to nonlambertian effects, between the stereo image

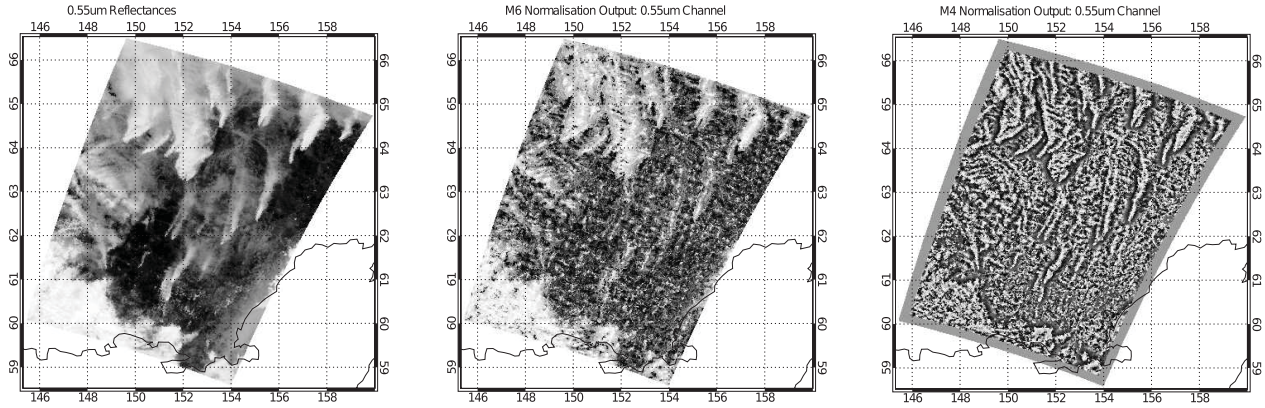


Fig. 1. Modified normalization outputs. (a) Subset of the 0.55 μm channel from AATSR orbit 43986. the lighter features are smoke plumes and clouds, while the darker features are the land surface. (b) M6 normalization of the subset in (a). Note the presence of texture in all regions of the image. (c) Result of the M4 normalization applied to the subset. Note the loss of image detail around plume edges as well as the loss of any retrievals around the image edge.

pair, which is vital for effective stereo matching of scenes in uncontrolled (i.e., natural) environments.

Normalization typically employs a band pass filter to suppress noise while preserving texture. In the case of NCC, this is achieved first through smoothing of the image with a Gaussian filter to provide the low-pass filtering and then subtracting this smoothed image from the original to provide a high-pass filtered residual. Normalization is then achieved through division by a local standard deviation within a small window. Due to the smoothing effect of the low-pass filter used in this method, blurring across discontinuities (changes in disparity) occurs during this normalization process. This blurring leads to a number of undesirable effects during the stereo matching of the image pair, including smoothed disparity fields and erroneous disparity estimation [15]. M6 aims to improve performance at discontinuities through an alternative normalization process.

Surrounding each pixel x in the reference image R, we define a neighborhood $\boldsymbol{N_x}$. For each $\boldsymbol{N_x}$, we find the subset of pixels ρ that are most similar to x, giving a reduced neighborhood $\boldsymbol{N_x^\rho}$. This is achieved through

$$\boldsymbol{N_x^\rho} = |\boldsymbol{x} - \boldsymbol{N_x}| < \boldsymbol{s} \tag{1}$$

where S is a threshold. We then use $\boldsymbol{N_x^\rho}$ to normalize x in place of $\boldsymbol{N_x}$. In this paper, N is of size 21 $\times$ 21 pixels, chosen to provide a large enough sample of pixels, while not increasing computation time significantly, and S is set to retain those pixels ρ whose absolute differences from x are within the 25% percentile.

If we let ξ define the pixels which comprise $\boldsymbol{N_x^\rho}$, i.e., $\boldsymbol{\xi} \in \boldsymbol{N_x^\rho}$, and we assume that these pixels are in the order sorted by value such that $\boldsymbol{\xi_0} \leq \boldsymbol{\xi_1} \leq \cdots \leq \boldsymbol{\xi_{n-1}}$, where n is the total number of pixels in $\boldsymbol{N_x^\rho}$, then we can calculate the median of the subset

$$\tilde{\boldsymbol{x}} = \left\{ \begin{array}{ll} \boldsymbol{\xi_k}\,, & \text{if } \boldsymbol{n} = 2\boldsymbol{k}-1 \\ (\boldsymbol{\xi_k} + \boldsymbol{\xi_{k+1}})^{0.5}\,, & \text{if } \boldsymbol{n} = 2\boldsymbol{k} \end{array} \right\}. \tag{2}$$

We can then compute the standard deviation of the subset $\boldsymbol{N_x^\rho}$ relative to its median $\tilde{\boldsymbol{x}}$

$$\boldsymbol{\sigma_x^\rho} = \sqrt{\frac{\sum_{i=0}^{n-1}(\boldsymbol{\xi_i} - \tilde{\boldsymbol{x}})^2}{\boldsymbol{n}} - \left(\frac{\sum_{i=0}^{n-1}(\boldsymbol{\xi_i} - \tilde{\boldsymbol{x}})}{\boldsymbol{n}}\right)^2}. \tag{3}$$

After using (2) and (3) to obtain the median and standard deviation for every pixel x in R, giving the medians of the reference image, $\boldsymbol{R_{\tilde{x}}}$, and the standard deviations of the reference image, $\boldsymbol{R_\sigma}$, we can perform the normalization

$$\boldsymbol{R}_{\text{norm}} = \frac{\boldsymbol{R} - \boldsymbol{R_{\tilde{x}}}}{\boldsymbol{R_\sigma} + \boldsymbol{\varepsilon}}. \tag{4}$$

The small positive number ϵ is set here to avoid zero divisions and to limit the amplification of image noise. Its default size is set equal to 10^{-3}. The same methods are also applied to normalize the comparison image

$$\boldsymbol{C}_{\text{norm}} = \frac{\boldsymbol{C} - \boldsymbol{C_{\tilde{x}}}}{\boldsymbol{C_\sigma} + \boldsymbol{\varepsilon}}. \tag{5}$$

This modified normalization algorithm leads to a significant reduction in smoothing at image discontinuities, allowing more reliable matching to be achieved in these challenging areas. The improved detail retention of M6 can be seen in comparison to the NCC normalization used in M4 shown in Fig. 1.

B. Matching

Let $\boldsymbol{i}$ and $\boldsymbol{j}$ define, respectively, the line and sample coordinates for all pixels $\boldsymbol{x}$ in the normalized reference and comparison images, R_{norm} and C_{norm}. We can then calculate the sum of absolute differences (SAD) for r different across track displacements u_r as follows:

$$\text{SAD}(\boldsymbol{u_r}) = \sum_{\boldsymbol{m}=0}^{\boldsymbol{p}} \sum_{\boldsymbol{n}=0}^{\boldsymbol{q}} \left| \boldsymbol{R}_{\text{norm}}(\boldsymbol{i_m}, \boldsymbol{j_n}) - \boldsymbol{C}_{\text{norm}}(\boldsymbol{i_m} - \boldsymbol{u_r}, \boldsymbol{j_n}) \right| \tag{6}$$

where p is the number of samples and q is the number of lines in the images. From (6), we find the displacement u_r, which minimizes the SAD over the entire image and use this to

displace the comparison image in the cross track direction such that $\boldsymbol{C}_{\text{norm}} = \boldsymbol{C}_{\text{norm}}(\boldsymbol{i}-\boldsymbol{u}_r, \boldsymbol{j})$. Once the images were aligned in the across track direction, we compute along track disparities, which give the height information. Given a suitably high across track wind speed (> 8 ms^{-1}), the across track disparities may not be consistent across the scene. Therefore, in the presence of high across track winds, we may not be locating the best matching pair of pixels between the images. We evaluated the matching outputs employing the above method and a pixel-by-pixel method against the validation datasets presented in Section V. Little difference between the outputs was noted, yet a significant boost in computer processing time was achieved by employing an entire scene shift, due to the need to evaluate each along track disparity only once, rather than υ_r times, where υ_r is the number of along track disparities.

The along track disparities are computed on a pixel-by-pixel basis rather than for the entire image as done in the cross track disparity estimation. M6 employs a modified SAD metric for its matching cost in this instance. The modification being that the sum is computed over a subset of pixels from the cost aggregation window, similar to variable window size algorithms [16]. To compute the subset of pixels over which to aggregate the costs, we apply (1) to each pixel in R_{norm}. Here, however, the indices of the pixels ξ of set $\boldsymbol{N}_{\boldsymbol{x}}^{\boldsymbol{\rho}}$ are used to define the size N neighborhood K_x over which the cost aggregation is performed, i.e., $\boldsymbol{K}_{\boldsymbol{x}}(\boldsymbol{\xi}) = 1$, and all indices, which are not of the set $\boldsymbol{N}_{\boldsymbol{x}}^{\boldsymbol{\rho}}$, are set to 0, i.e., $\psi \notin \boldsymbol{N}_{\boldsymbol{x}}^{\boldsymbol{\rho}}$, $\boldsymbol{K}_{\boldsymbol{x}}(\boldsymbol{\psi}) = 0$.

Once the pattern kernel is derived for x, the matching metric is evaluated as in

$$\text{SAD}(x, \upsilon_r) = \frac{\sum_{k=0}^{p}\sum_{l=0}^{q}\left|\begin{matrix}\boldsymbol{R}_{\text{norm}}(\boldsymbol{i_m}+\boldsymbol{k}-\boldsymbol{p}/2, \boldsymbol{j_n}+l-\boldsymbol{q}/2)- \\ \boldsymbol{C}_{\text{norm}}(\boldsymbol{i_m}+\boldsymbol{k}-\boldsymbol{p}/2, \boldsymbol{j_n}-\boldsymbol{v_r}+l-\boldsymbol{q}/2)\end{matrix}\right| \times \boldsymbol{K}_{\boldsymbol{x}}(k, l)}{\sum_{k=0}^{p}\sum_{l=0}^{q}\boldsymbol{K}_{\boldsymbol{x}}(k, l)} \tag{7}$$

where p is the number of samples and q is the number of lines in N. While working through the list of along track displacement vectors υ_r, the value of the metric and the associated displacement are stored for x. Then, to determine the final disparity, a spline interpolation routine is applied to the set of costs to find the subpixel disparity. The quoted accuracy of the algorithm is pixel level due to the pixel level accuracy of the co-registration correction coefficients [14]. The resultant final disparity map is then smoothed using a median filter. The algorithm output is shown in Fig. 2.

IV. Smoke Plume Masking

SPMs are required for the accurate delineation of smoke from other features such as clouds and the land surface, assuming that smoke when mixed sufficiently with cloud droplets will be difficult, if not impossible, to differentiate. Previous attempts at SPM generation have tended to focus upon radiometric threshold techniques. The thresholds are typically derived from statistical- or machine-learning methods [18], [19]. While oftentimes effective, these SPMs are often susceptible to misclassification due to the lack of a stable reflectance curve for smoke, with large overlaps existing between the spectral profile of smoke and non-smoke features [19]. Through a combination of stereo-derived heights, and visible channels and thermal channels, AATSR is capable of effectively classifying smoke and hence, automated smoke detection is achievable.

The SPM is generated in a number of steps as follows. First, a cloud mask is generated from the 11 μm channels using a radiometric or a statistically derived threshold

$$\boldsymbol{M}_c\ (\boldsymbol{x}) = \begin{cases} 0, & \text{if } (\boldsymbol{F}_{11}(\boldsymbol{x}) < \bar{\boldsymbol{F}} + \boldsymbol{F}_{\sigma}) \text{ or } (\boldsymbol{F}_{11}(\boldsymbol{x}) < 280 \text{ K}) \\ 1, & \text{otherwise} \end{cases} \tag{8}$$

where M_c is the cloud mask, F is the forward scene for the subscripted channel, $\bar{\boldsymbol{F}}$ is the image mean, and F_σ is the image standard deviation. This thresholding method only works effectively in the ALANIS study region during the fire season (April through September), as it relies upon the land surface being warmer than any detected cloud feature. Any snow-covered ground would also be flagged as cloud using this method. The second step is to buffer the cloud mask. This buffering is required to try to remove as many cloudy pixels as possible from the SPM. As stereo results can vary between channels, there is not always a direct cloud correspondence, so buffering removes the errors associated with this. The buffering is achieved through application of a boxcar averaging filter to the cloud mask. This effectively dilates the cloud mask leading to the required buffering effect. Once derived, the cloud mask is applied to the forward and nadir scenes prior to stereo matching to eliminate those pixels from the pattern matching. This is required as area-based methods, even with the modifications applied here, have a tendency to smooth features. Due to these smoothing effects, *a posteriori* application of the cloud mask rarely removes all cloud features.

The final SPM, M_{sp} is generated by removing land features by comparing the cloud masked 0.55 μm stereo-derived altitude result, SPIH, to a DEM. Any pixel whose elevation is in proximity to the DEM to within the algorithms' height precision ($\approx$1 km for AATSR), a, is flagged as land, and set to zero

$$\boldsymbol{M}_{sp}(\boldsymbol{x}) = \begin{cases} 1\ , & \text{if SPIH}(\boldsymbol{x}) - \text{DEM}(\boldsymbol{x}) > \boldsymbol{a} \\ 0\ , & \text{otherwise.} \end{cases} \tag{9}$$

The effectiveness of the SPIH mask can be seen in Fig. 3, note that no clouds contaminate the mask at all. This method does lead to occasional false-positives being detected, particularly over polluted populated areas such as the Benelux region. However, due to the unlikelihood of fires of sufficient energy to have SPIH of >1 km, it is easy to detect and screen these false-positives manually.

V. Validation Datasets and Methods Applied

A. CALIOP Validation Methodology

The CALIOP instrument is a two wavelength polarization sensitive lidar carried aboard the CALIPSO satellite in the NASA A-Train constellation. The CALIOP instrument, which is making measurements of clouds and aerosols since June 13, 2006, receives backscattered radiation in three channels,

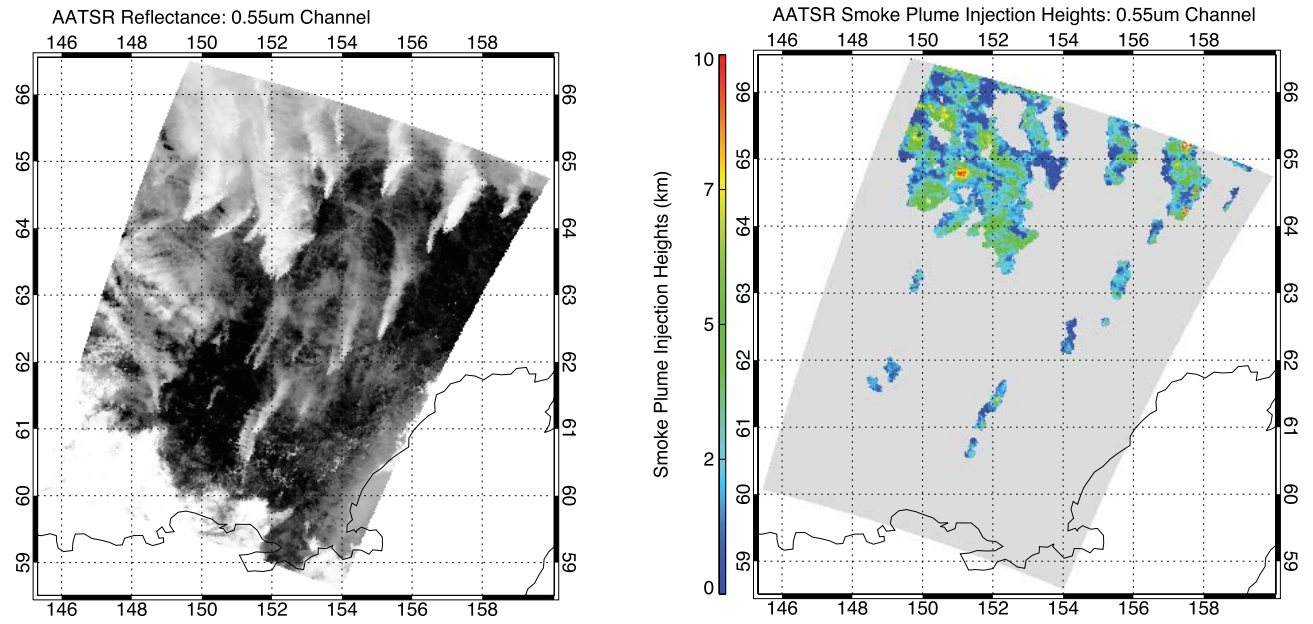


Fig. 2. (a) AATSR 0.55 μm forward reflectance for a large number of smoke plumes from the northwest of Kamchatka, AATSR orbit 43986 and (b) result of the SPIH and SPM algorithms (see Section IV for a description of the SPM algorithm) applied to the left image.

two at 532 nm with sensitivity to the backscattered intensity at orthogonal polarizations and one at 1064 nm. The lidar has a ground footprint on the ellipsoid of 100 m and pulses every 333 m along track. The vertical resolution is between 30 and 60 m depending on the altitude of the backscattering surface, with 30 m resolution achievable in the troposphere. The received level one (L1) data is processed into a number of level two (L2) products with numerous useful characteristics for intercomparison with stereo-derived SPIHs from AATSR.

Intercomparisons are performed using the 1-km cloud product (CAL_LID_L2_01kmCLay-ValStage1-V3-01), chosen for its similar resolution to the AATSR instrument, while still sensitive enough to detect smoke plume features. The lower limit for cloud detection at all horizontal scales is a backscattered signal of greater than 1×10^{-3} km^{-1}sr^{-1} (equivalent to an optical depth of 0.01 for cirrus clouds), which is also suitable for the detection of smoke plumes.

Due to the different orbital paths and equatorial overpass times (10:00 A.M. Envisat versus 13:30 CALIPSO), temporally collocated measurements between AATSR and CALIOP only occur with regularity at polar latitudes; as scene acquisitions descend toward the equator, the likelihood of finding collocated measurements decreases. Furthermore, the time difference between any possible collocations increases. The ALANIS study region enables the location of coincident orbital paths; however, the measurements are separated by approximately 2 h at the latitudes where most smoke plumes are present. Further the validation, of course, requires that smoke plumes are seen by both instruments; as the CALIOP instruments only samples 100 m footprints every 333 m, it still may not provide spatially coincident measurements of smoke plumes.

For 2008, all AATSR orbits in the study region were assessed visually for smoke plumes using the National Earth Observation Data Centre (NEODC) AATSR browse data set. For any scene, which was found to contain smoke, a second search was undertaken using the COVE tool (http://www.ceos-cove.org) to check for any co-incidences between the AATSR orbits and CALIPSO. Through this checking process, a total of three suitable scenes (29/7/10, 1/8/10, 2/8/10) where AATSR and CALIOP observed the same smoke plume were found from hundreds of individual orbits. The scenes are masked using the masking method outlined in Section IV prior to evaluation, to ensure that only M6s performance on smoke plumes is being assessed.

Due to the ~2-h time delay between the instrument overpasses, a correction for wind-induced shift in the location of measured AATSR SPIH pixels was tested to determine whether it might lead to more accurate results. The method involved finding the geographically closest AATSR pixel for each CALIOP measurement, using this pixel as an *a priori* location, and then back projecting to a position consistent with the CALIOP overpass using ECMWF meridional and zonal surface wind profiles. After preliminary evaluation, the wind correction led to a degradation of the intercomparison and therefore was not subsequently applied. The reasons for this are assumed to be due to the incompatibility between the coarse resolution of the 2.5° ECMWF wind grids and the mesoscale resolution of the AATSR observations to be corrected.

The geo-location accuracy of the AATSR instrument is not documented in the literature. However, ATSR-2, AATSRs predecessor, is documented to be good to ± 2 km geo-location accuracy [20], so the assumption made here is that AATSR shares a similar accuracy level. Therefore, one cannot be

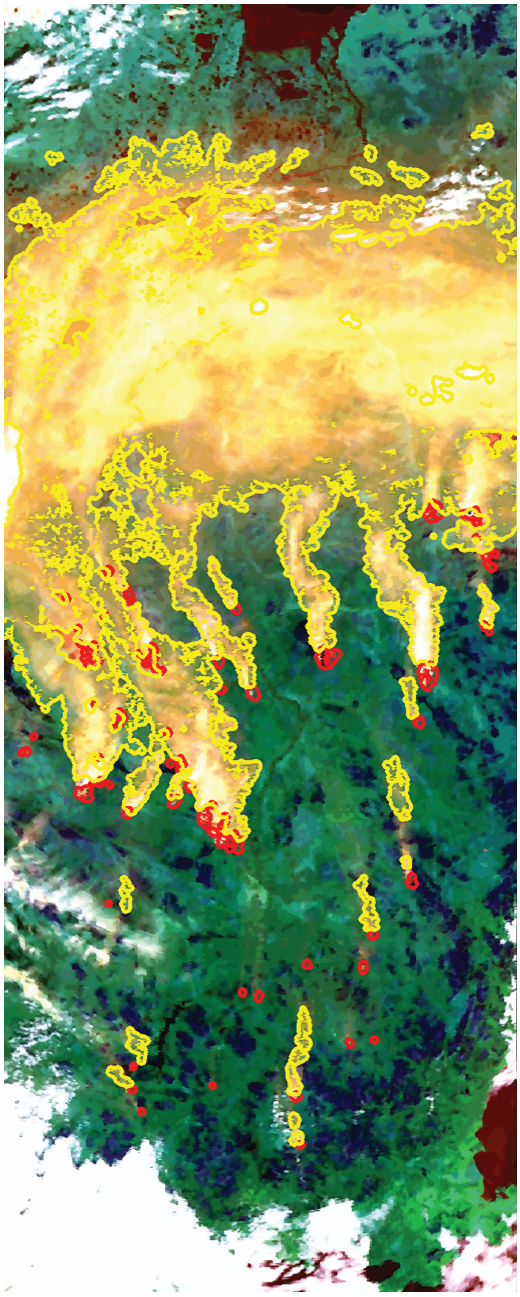


Fig. 3. SPM boundary of the SPIH output for a scene from AATSR (orbit: 43986) is shown here. The masked regions are defined as those within the yellow contours. The red dots are fires detected by the MODIS FIRMS algorithm [17].

confident that the collocated measurements are actually measuring the same point on the earth's surface. To account for this, a ±2 pixel bounding-box is used. During the intercomparison of a collocated measurement, the AATSR pixel within the bounding-box that returns the minimum difference from the CALIOP layer retrieved SPIH is retained as the collocated measurement, with the supposition that it is the most likely to be the pixel collocated with CALIOP. The effects of choosing the minimum SPIH difference from the bounding-box are further assessed in Section VII and are shown not to significantly influence the outcome of the results. Another caveat of AATSRs geo-location accuracy is that we

TABLE I
BOUNDING-BOX ASSESSMENT

Statistic	Minimum	Mean	Median
N	22885	22815	22804
RMSE	0.66	0.64	0.65
Bias	0.29	−0.03	−0.02
R**2	0.36	0.31	0.31

N is the number of points assessed, RMSE is the root mean square error, and R2 is the coefficient of determination.

cannot be certain that we are using the correct DEM height to correct elevation of the derived SPIH; in regions of rapidly varying terrain (> 250 m elevation change per 1 km), the accuracy of the retrieved SPIH will be reduced.

CALIOP is able to retrieve multiple layers of cloud and aerosols and is sensitive to very low aerosol optical depths equivalent ($\ll 0.01$). From this analysis, it is found that the top layer detected by CALIOP does not always correspond to the uppermost smoke layer observed by AATSR. This is due to the greater sensitivity of the lidar instrument compared to a passive imaging system. To account for this, all layers detected by CALIOP are compared with the SPIH pixels within the AATSR-bounding-box to ensure that the correct match-up layer is retrieved.

B. MISR Validation Methodology

The MISR instrument carried onboard the NASA Terra satellite, launched in December 1999, is a nine-camera (arranged between ±70°) radiometer with four spectral channels (0.45, 0.56, 0.67, and 0.87 μm) located in the visible and near-infrared bands. The instrument operates in two modes, local mode, and the general operational method, global mode. The global mode allows imaging at 275 m across all channels in the nadir camera but only in the red channel across all other imaging angles. All other channels image at 1100 m resolution, significantly reducing the data volume. The multiple camera arrangement allows for stereo cloud, smoke, and land height measurements to be made [10] achieving an accuracy of ∼500 m [21]. The Terra satellite has an equatorial overpass time of 10:30 A.M., hence being much closer in time to Envisat (10:00 A.M.). This allows for a more direct intercomparison when orbital swath coincidences occur and therefore potentially improved validation results.

Data from the MISR instrument has been used to generate SPIH datasets for a number of regions globally under the MISR Plume Height Project [2], [22]. The MISR SPIH from the MISR plume height project are manually digitized using the MISR INteractive eXplorer (MINX) tool, an IDL-based visualization and digitizing toolkit for making stereo measurements of clouds, smoke/volcanic plumes, dust clouds, and so forth, from MISR imagery. There are a number of algorithms within MINX; however, of most importance here to validation is the camera image-matching algorithm. The MINX stereo-matching algorithm uses Pearson's correlation coefficient to match an image patch extracted from the nadir camera to five other views (CF, BF, AF, AA, BA, and CA) at 275-m resolution [2], [22]. Once the patch has been matched in all six views, a minimum curvature surface is fitted to the

correlation matrix, and is used to interpolate the disparity to a subpixel location. Following matching, a wind speed correction algorithm is applied [2], [22] leading to a final height accuracy of approximately ±200 m.

The MISR plume height project includes a dataset for Siberia for the 2008 fire season (with plumes digitized in all orbits from March through until August 2008), which is coincident with the ESAP study region. The smoke plumes in the MISR SPIH dataset are collocated with AATSR orbits using the COVE tool, giving a total of nine orbits containing collocated smoke plumes, comprising thousands of individual measurements with overpass times differing by ∼30 min. The collocations are evaluated using the same methods, where applicable, as applied to the CALIOP analysis.

VI. Results

A. CALIOP Intercomparison Results

After finding the three orbital match-ups, all pixels which were intersected by the CALIOP footprints in the AATSR SPIH were extracted and compared. Outliers are defined as those pixels where the SPIH difference is greater than two standard deviations from the mean and removed from further analysis. See Fig. 4.

B. MISR M6 Intercomparison Results

A similar procedure for matching up pixels from the SPIHs from MISR with those from AATSR was applied, however, using a different criteria to eliminate outliers (discussed in Section VII). Further, as there are several MISR SPIH pixels for each AATSR pixel, the MISR data is resampled to a resolution of 1100 m, so that it is more similar to that of AATSR prior to statistical analysis. See Fig. 5.

C. MISR M4 Intercomparison Results

The methods are the same as those in Section V-B. See Fig. 6.

VII. Discussion

To ensure confidence in the results obtained in the intercomparison, the implications of using a ±2 pixel bounding-box to account for the poor geo-referencing of the AATSR instrument must be understood. In the analysis, the minimum difference between each independent SPIH measurement and the AATSR SPIH measurements from within the bounding-box is used as the collocated difference result. Using the minimum gives the optimum outcome; however, this result may be misleading due to noise effects typical in stereo-matching algorithms [23]. To assess whether the results are misleading, additional statistical measures (the median and mean differences) between the independent SPIH measurement and the SPIH measurements contained within the AATSR-bounding-box are generated and presented in Table I.

The results presented in Table I demonstrate that the median and mean show very similar statistical characteristics to the minimum. This indicates that the majority of the height differences within the bounding boxes are similar to the minimum difference height, giving confidence in using the minimum as the quality measure. This additional statistical analysis shows that the implications of using the minimum SPIH difference from the bounding-box are limited, and confidence in the analysis is justified.

The CALIOP intercomparison comprises a relatively small number of data points. This is due to the limited number of co-located measurements found, as would be expected at the study region latitudes and given the instrument characteristics. Following outlier removal, where an outlier is here defined as being $\geq$ 2 standard deviations from the mean, the intercompared points show mostly good agreement overall with an R^2 value of 0.5, which was found to be statistically significant to the 0.05 level using a two-tailed Student's t test. A slight high bias for CALIOP is evident, and is probably due to increased lidar sensitivity to aerosol particles, causing CALIOP to retrieve heights nearer to the top of the plume. Whereas, AATSR measures the height where the plume reaches a suitably high optical thickness to be detected in the passive sensor array.

The MISR analysis against M6 presented in Section VI-B provides a far more comprehensive intercomparison of SPIH, with the number of individual pixels intercompared numbering in the tens of thousands. Focusing solely on the results generated using the minimum difference shown in Fig. 4, it is found that the initial correlation statistics computed for all points show a rather poor R^2 score ∼0.3 following outlier removal, with anything $\geq$ 2 standard deviations from the mean classed as an outlier. Looking at the minimum plot in Fig. 5, this appears to be due to MISR detecting a large number of points as smoke plumes which AATSR detects as the land surface, i.e., $<$ 1 km. The reason for this is that MISR detects more smoke than AATSR, due to its higher resolution, higher quantization 14-b cf. 10-b for AATSR, and an increased number of observations of any plume feature (7 MISR camera angles are employed in MINX compared with 2 for AATSR); therefore it can resolve the smoke plumes, which AATSR cannot.

Those points that AATSR detects as the land surface, when excluded from the analysis as outliers, provide a better assessment of the quality of the M6 measurements. Here, we assume that any AATSR SPIH measurement, which is $<$ 1 km above the DEM is potentially the land surface and is therefore excluded from the statistical assessment. With this threshold applied, the statistics are greatly improved with an R^2 score of ∼0.69, at the 0.05 level using a two-tailed Student's t test.

In order to appreciate the improvement in efficacy and accuracy gained through application of M6, the M4 matcher output is also compared against the MISR data set and the results contrasted. M4, originally developed for cloud top height retrieval from ATSR2, performs poorly in comparison to M6 for smoke features. The coefficient of determination score, even with outlier removal, where similarly to the M6 analysis, outlier are classes anything $\geq$ 2 standard deviations from the mean, is not as good, at 0.28.

The poor results are due to a number of reasons: M4 is designed specifically to retrieve cloud top height and is effective at resolving larger image features, such as extensive unbroken cloud decks. The M4 algorithm applies Gaussian

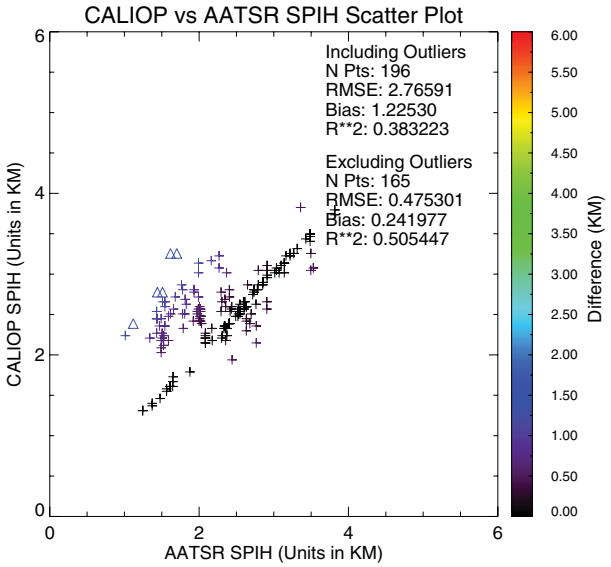


Fig. 4. Results for the intercomparison between AATSR and CALIOP using the methods described in Section V-A. Inset within the plot are two sets of statistics, one containing outliers and one where the outliers were removed, where an outlier is any result outside two standard deviations from the mean. Outliers are represented in the plot as triangles.

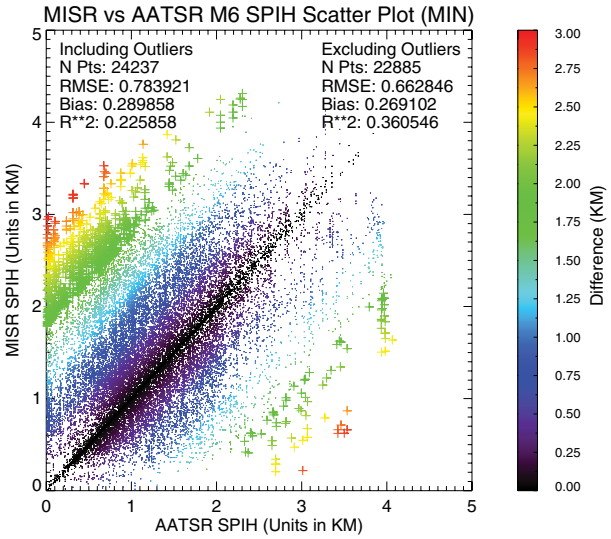


Fig. 5. intercomparison of MISR-AATSR for the M6 matcher. Multiple statistical measures are used in the evaluation; the reasoning for this is given in Section VII. Again outliers are classified as those points which are outside ±2 standard deviations from the mean, denoted in the scatter plot as crosses.

smoothing kernels of a large radius (10 pixels) during its normalization phase and Gaussian aggregation kernels during its matching phase (5 pixels). These kernels, while well suited to the extraction of extensive image features, such as clouds decks, suffer from poor performance in the presence of small image features and also at the boundaries or edges between distinct image features, which typically contain discontinuities, or a change in disparity. In the normalization phase, small image features can either be "washed out" from the image due to the averaging applied, or a halo effect can occur, where a region without texture is formed around the image feature [such effects can be seen in Fig. 1(c)]. The halo effect not only affects small image features, but also occurs at any boundary of two distinct intensity (pixel value) populations

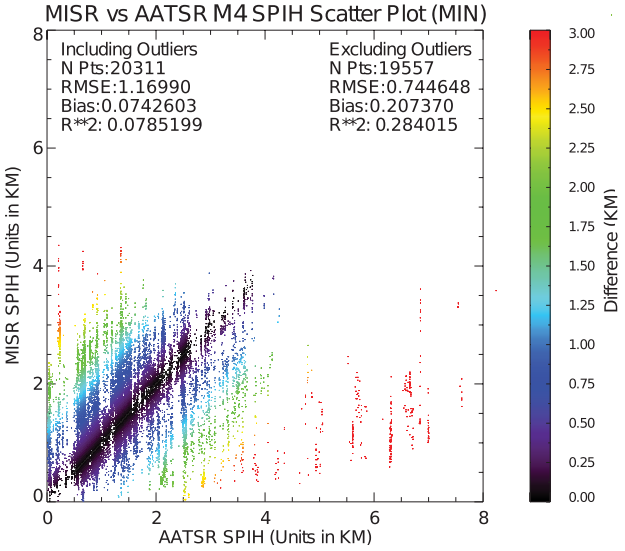


Fig. 6. Analysis of the MISR-AATSR intercomparison for the M4 matcher. The same outlier removal process as applied to M6 is used to ensure fair comparison between M6 and M4. The outliers are outside the axes limits and are not displayed. Note that the plot axes are different to those employed in Fig. 5 to better illustrate the SPIH determination problems associated with the M4 matcher, but the color applied is the same; everything in red is greater than 3 km in height difference.

(i.e., moving from a smoke plume feature to a water cloud feature). The halo effect is caused by the inability of the NCC algorithm applied in M4 to effectively normalize bimodal distributions. NCC is reliant upon the intensity distribution from the normalization window being Gaussian; when two distinct intensity populations exist, as is common when moving from a feature at one disparity to another, the Gaussian requirement is not fulfilled and the mean and standard deviation used in the normalization process do not effectively describe the data within the normalization window. This leads to a loss of texture at discontinuities and the aforementioned halo effect at feature edges.

For smoke plumes, extent is dependent on the magnitude of the fire and the local meteorological conditions. For example, looking at the plumes in Figs. 1–3, we can see that some plumes have combined to form an extensive plume cloud, while a number of other plumes are smaller and distinct in nature. M4 is able to effectively capture the larger plume features, as they are similar to extensive cloud decks in form; however, due to the aforementioned reasoning, the smaller plume features are lost and poor performance is shown at interfaces between smoke plumes and water clouds, the impacts of which are evident in Fig. 6. In Fig. 6, there are a large number of erroneous height estimations where the AATSR M4 SPIH outputs are more than 4 km above the MISR outputs. These erroneous height estimations are caused by the halo regions present at discontinuities in the M4 normalization, as within the halo there is limited texture for the matching algorithm to evaluate, therefore the matching is ambiguous and consensus as to which is the optimum match cannot be reached, leading to invalid height estimations being returned. This significantly reduces the quality of the M4 output, and dominates in the presence of smaller features, as is often the case for smoke plumes, which are frequently observed to be saturated by

edges and therefore discontinuities and intensity changes. The M6 normalization does not employ the smoothing strategies applied in M4, but rather separates the intensity populations so that only pixels from a similar intensity distribution are used in the normalization process. Therefore, it is able to perform more effectively at discontinuities, enabling detection of more of the smaller smoke plumes in a more robust way than M4, as is evident in the analysis in Fig. 5.

VIII. Conclusion

A JAVA-based processing chain for the automated derivation of SPIH and SPMs from AATSR was described. The processing chain was implemented at Mullard Space Science Laboratory on a multithreaded Linux cluster with the output products being delivered to our partners on the ESA-Noveltis smoke plumes project (http://www.noveltis.fr/alanis/) in order to determine their potential in improving the predictive efficacy of CTMs. The final product comprised a masked SPIH NetCDF derived from the 0.55 μm AATSR channel, as well as other useful ancillary data, on a daily timescale for the EASP study area for yearly epochs of April through September for four years (2008–2011).

The outputs from the processing chain were evaluated against SPIH measurements from the independent instruments, MISR and CALIOP for the whole of the fire season in 2008. The outputs of the validation showed excellent agreement between the AATSR results and the two alternative measurement datasets. Further, M6 was shown to significantly outperform M4, a stereo-matching algorithm designed to retrieve cloud top height, further demonstrating M6's effectiveness in the determination of SPIH compared to more traditional stereo-matching algorithms. Given boundary layer heights, it is possible to employ the SPIHs with SPMs applied to determine how much material is injected into the free troposphere. The SPIHs/SPMs could then be used to assess the impact of these new datasets on CTMs. However, this is not the subject of this paper.

Acknowledgment

The authors would like to thank the NEODC, the MISR plumes project, and the CALIPSO science teams for providing validation data. They would also like to thank NCEO for their partial support of NERC Ph.D. studentship number NER/S/C/2006/14369 as well as ESA for ESA ALANIS contract (ESRIN Contract N°42000230S5/1O/I-LG).

References

[1] P. J. Crutzen, L. E. Heidt, J. P. Krasnec, W. H. Pollock, and W. Seiler, "Biomass burning as a source of atmospheric gases CO, H_2, N_2O, NO, CH_3Cl and COS," *Nature*, vol. 282, pp. 253–256, Nov. 1979.

[2] M. V. Martin, J. A. Logan, R. A. Kahn, F.-Y. Leung, D. L. Nelson, and D. J. Diner, "Smoke injection heights from fires in North America: Analysis of 5 years of satellite observations," *Atmos. Chem. Phys.*, vol. 10, no. 4, pp. 1491–1510, 2010.

[3] S. Turquety, J. A. Logan, D. J. Jacob, R. C. Hudman, F. Y. Leung, C. L. Heald, R. M. Yantosca, S. Wu, L. K. Emmons, D. P. Edwards, and G. W. Sachse, "Inventory of boreal fire emissions for North America in 2004: Importance of peat burning and pyroconvective injection," *J. Geophys. Res.*, vol. 112, pp. 1–13, Apr. 2007.

[4] B. Langmann, B. Duncan, C. Textor, J. Trentmann, and G. R. van der Werf, "Vegetation fire emissions and their impact on air pollution and climate," *Atmos. Environ.*, vol. 43, no. 1, pp. 107–116, Jan. 2009.

[5] P. R. Colarco, M. R. Schoeberl, B. G. Doddridge, L. T. Marufu, O. Torres, and E. J. Welton, "Transport of smoke from Canadian forest fires to the surface near Washington, DC: Injection height, entrainment, and optical properties," *J. Geophys. Res., Atmos.*, vol. 109, no. D6, Mar. 2004.

[6] M. Krol, S. Houweling, B. Bregman, M. van den Broek, A. Segers, P. van Velthoven, W. Peters, F. Dentener, and P. Bergamaschi, "The two-way nested global chemistry-transport zoom model TM5: Algorithm and applications," *Atmos. Chem. Phys.*, vol. 5, no. 2, pp. 417–432, 2005.

[7] D. M. Winker, M. A. Vaughan, A. Omar, Y. Hu, K. A. Powell, Z. Liu, W. H. Hunt, and S. A. Young, "Overview of the CALIPSO mission and CALIOP data processing algorithms," *J. Atmos. Ocean. Technol.*, vol. 26, no. 11, pp. 2310–2323, Nov. 2009.

[8] D. Diner, T. Ackerman, A. Braverman, C. Bruegge, M. Chopping, E. Clothiaux, R. Davies, L. Di Girolamo, R. Kahn, Y. Knyazikhin, Y. Liu, R. Marchand, J. Martonchik, J. Muller, A. Nolin, B. Pinty, M. Verstraete, D. Wu, M. Garay, O. Kalashnikova, A. Davis, E. Davis, and R. Chipman, "Ten years of MISR observations from Terra: Looking back, ahead, and in between," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2010, pp. 1297–1299.

[9] D. Llewellyn-Jones, M. Edwards, C. Mutlow, A. Birks, I. Barton, and H. Tait, "AATSR: Global-change and surface-temperature measurements from ENVISAT," *ESA Bull-Eur. Space*, vol. 105, pp. 11–21, Feb. 2001.

[10] J. Muller, A. Mandanayake, C. Moroney, R. Davies, D. Diner, and S. Paradise, "MISR stereoscopic image matchers: Techniques and results," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 7, pp. 1547–1559, Jul. 2002.

[11] J.-P. Muller, M.-A. Denis, R. D. Dundas, K. L. Mitchell, C. Naud, and H. Mannstein, "Stereo cloud-top heights and cloud fraction retrieval from ATSR-2," *Int. J. Remote Sens.*, vol. 28, no. 9, pp. 1921–1938, Jun. 2007.

[12] M. A. Denis, J. P. Muller, and H. Mannstein, "ATSR-2 camera models for the automated stereo photogrammetric retrieval of cloud-top heights-initial assessments," *Int. J. Remote Sens.*, vol. 28, no. 9, pp. 1939–1955, 2007.

[13] J. J. Danielson and D. B. Gesch, "Global multi-resolution terrain elevation data 2010 (GMTED2010)," U.S. Geological Survey, Reston, VA, USA, Tech. Rep. 2011-1073, 2010.

[14] D. Fisher and J.-P. Muller, "Global warping coefficients for improving ATSR co-registration," *Remote Sens. Lett.*, vol. 4, no. 2, pp. 151–160, Feb. 2012.

[15] H. Hirschmuller, P. Innocent, and J. Garibaldi, "Real-time correlation-based stereo vision with reduced border errors," *Int. J. Comput. Vis.*, vol. 47, nos. 1–3, pp. 229–246, Apr. 2002.

[16] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *Int. J. Comput. Vis.*, vol. 47, no. 1, pp. 7–42, May 2002.

[17] D. K. Davies, S. Ilavajhala, M. M. Wong, and C. O. Justice, "Fire information for resource management system: Archiving and distributing MODIS active fire data," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 1, pp. 72–79, Jan. 2009.

[18] D. Mazzoni, J. A. Logan, D. Diner, R. Kahn, L. Tong, and Q. Li, "A data-mining approach to associating MISR smoke plume heights with MODIS fire measurements," *Remote Sens. Environ.*, vol. 107, no. 1, pp. 138–148, Mar. 2007.

[19] Y. Xie, J. J. Qu, X. Xiong, X. Hao, N. Che, and W. Sommers, "Smoke plume detection in the eastern United States using MODIS," *Int. J. Remote Sens.*, vol. 28, no. 10, pp. 2367–2374, Jul. 2007.

[20] C. J. Merchant, J. J. Simpson, and A. R. Harris, "A cross-calibration of GMS-5 thermal channels against ATSR-2," *Remote Sens. Environ.*, vol. 84, no. 2, pp. 268–282, Feb. 2003.

[21] C. Moroney, R. Davies, and J. P. Muller, "Operational retrieval of cloud-top heights using MISR data," *IEEE Trans. Geosci. Rem. Sens.*, vol. 40, no. 7, pp. 1532–1540, Jul. 2002.

[22] S. Scollo, R. Kahn, D. Nelson, M. Coltelli, D. J. Diner, M. J. Garay, and V. J. Realmuto, "MISR observations of Etna volcanic plumes," *J. Geophys. Res.*, vol. 117, no. D6, pp. 1–33, Jul. 2012.

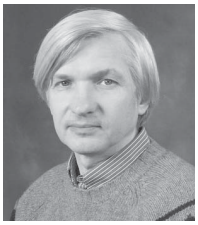
[23] H. Hirschmüller and D. Scharstein, "Evaluation of stereo matching costs on images with radiometric differences," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 9, pp. 1582–1599, Sep. 2009.

Daniel Fisher received the B.Sc. degree (Hons.) in geography and the M.Sc. degree in remote sensing from University College London (UCL), London, U.K., in 2006 and 2009, respectively. He is currently pursuing the Ph.D. degree in atmospheric remote sensing with Mullard Space Science Laboratory, UCL.

Jan-Peter Muller received the B.S. degree (Hons.) in physics from Sheffield University, Sheffield, U.K., in 1976, the M.S. and D.I.C. degrees in atmospheric physics and dynamics from Imperial College, London, U.K., in 1977, and the Ph.D. degree in planetary meteorology from University College London (UCL), London, U.K., in 1982.

He is currently the Head of the Imaging Group and a Professor of image understanding and remote sensing with UCL, where he has been a member of the faculty since 1984. He is currently a Principal Investigator of the ESA GlobalAlbedo Project, generating an unique 15-year record of earth's land surface albedo from European and U.S. sensors, and the Principal Investigator on the creation of the unique ESA ADAM spectral BRDF database of the Earth. He has been a member of the Multiangle Imaging Spectroradiometer and Moderate Resolution Imaging Spectroradiometer Science Teams since 1990. His current research interests include practical applications of machine vision to remote sensing of the Earth and planets, with a special focus on operational applications of stereo photogrammetry and other 3-D imaging technologies for applications in climate modeling, weather forecasting, and planetary exploration.

Dr. Muller is the Chair of the U.K. Aurora Advisory Committee, the Committee on Earth Observation Satellites WG on Cal/Val Terrain mapping Sub-Group, and the International Society for Photogrammetry and Remote Sensing IV/3 Working Group on Global DEM interoperability.



Vladimir N. Yershov received the M.Sci. degree in astronomy from Leningrad State University, Leningrad, Russia, in 1979, and the Ph.D. degree in physics and mathematics from the Central Astronomical (Pulkovo) Observatory, USSR Academy of Sciences, Leningrad, in 1986.

He worked in the field of fundamental astronomy and astronomical instrumentation with the Central Astronomical (Pulkovo) Observatory, USSR Academy of Sciences. In 2001, he moved to Mullard Space Science Laboratory, University College London, London, U.K., where he continues working on data processing for space astronomy missions, such as XMM Newton and Swift.