

AUTOMATIC DETECTION OF CHANGES IN VOLCANIC ACTIVITY USING GROUND
BASED NEAR-INFRARED CAMERAS TO MONITOR THERMAL INCANDESCENCE

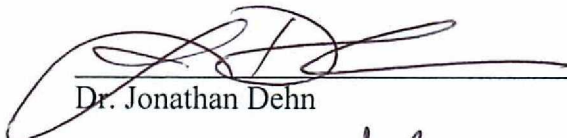
By

Martin C. Harrild

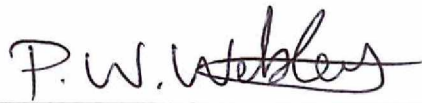
RECOMMENDED:



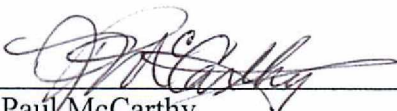
Dr. Franz Meyer



Dr. Jonathan Dehn




Dr. Peter Webley
Advisory Committee Chair

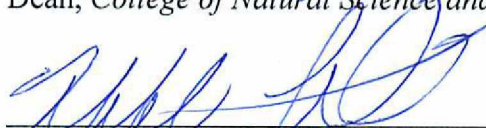


Dr. Paul McCarthy
Chair, *Department of Geosciences*

APPROVED:



Dr. Paul Layer
Dean, *College of Natural Science and Mathematics*



Dr. Michael Castellini
Dean of the *Graduate School*



Date

AUTOMATIC DETECTION OF CHANGES IN VOLCANIC ACTIVITY USING GROUND
BASED NEAR-INFRARED CAMERAS TO MONITOR THERMAL INCANDESCENCE

By

Martin C. Harrild, B.S.

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

in

Geophysics

University of Alaska Fairbanks

May 2017

APPROVED:

Dr. Peter Webley, Committee Chair

Dr. Jonathan Dehn, Committee Member

Dr. Franz Meyer, Committee Member

Dr. Paul McCarthy, Chair

Department of Geosciences

Dr. Paul Layer, Dean

College of Natural Science and Mathematics

Dr. Michael Castellini, *Dean of the Graduate School*

Abstract

An increase in thermal activity is a common precursor of volcanic eruptions and, if identified, can be used to advise local observatories to disseminate the appropriate advanced warnings. As continuously operating near-infrared (NIR) cameras are becoming more readily available at active volcanoes around the world, this investigation explores the use of identifying changes in pixel brightness in webcam imagery resulting from increased thermal incandescence. A fast, efficient, and fully automated Python algorithm has been developed with a primary focus on effective volcano monitoring and reducing overall financial costs. The algorithm includes three important tests (statistical analysis, edge detection, and Gaussian mixture model) to identify changes in activity in near-real time. The developed algorithm can be installed locally with a webcam or at a central location, with no need for additional costs. This algorithm approach was preliminarily tested on data from a permanently installed thermal infrared camera at Stromboli volcano, with a successful detection rate of 75.34%. The algorithm based methodology was further developed and applied to freely available online webcam imagery from Shiveluch volcano, with an overall accuracy of 96.0%, and a critical success index (CSI) of 76.7%.

Further refinements to the algorithm were made to reduce the false alarm rate (FAR) and number of missed events, and applied to four additional image datasets at Shiveluch, Fuego, Popocatepetl, and Stromboli. The algorithm successfully identified two large eruptions at Shiveluch, between 40 minutes and 2.5 hours prior to other satellite remote sensing methods, correctly identified the beginning of a large eruption at Fuego, which corresponded with local seismic data, and successfully identified a 90-minute window of increased activity leading to a large paroxysm event at Popocatepetl, which was described by the local observatory as having 'little to no warning'. The algorithm underperformed at Stromboli as the images here were captured in the thermal infrared (TIR) instead of the NIR, identifying the need for further improvements to ensure the algorithm performs correctly across multiple datasets. Overall, the algorithm developed here identifies thermally incandescent activity from increases in image pixel brightness remarkably well, and would complement existing volcano observatory monitoring tools, especially in remote or financially restricted locations as the equipment and coding language used here are extremely cheap compared to many other monitoring methods.

Table of Contents

	Page
Abstract	iii
Table of Contents	v
List of Figures	ix
List of Tables	xiii
List of Acronyms	xv
Acknowledgements	xix
Section 1: Introduction	1
1.1 Background	1
1.2 Volcano Monitoring	3
1.3 Research Aim	3
1.4 Volcanoes of Interest	4
1.4.1 Stromboli Volcano, Italy	5
1.4.2 Shiveluch Volcano, Russia	8
1.4.3 Fuego Volcano, Guatemala	10
1.4.2 Popocatépetl Volcano, Mexico	11
1.5 Overview	13
Section 2: Spectral Response Experiments	15
2.1 Experimental Set Up	16
2.2 Equipment Information	17
2.2.1 Omega BB-4A Infrared Calibration Source	19
2.2.2 EverFocus ED350HQ Camera	20
2.2.3 CM711 1/3" CCD CCTV Camera	20
2.2.4 Starview 4SV558CW48 Camera	21

2.2.5 Swann SWPRO-535 Camera	21
2.2.6 Zosi ZSZR04BA Camera	22
2.2.7 Logitech HD Pro C920 Camera	22
2.2.8 Apexis J012 IP Camera	23
2.2.9 SBIG ST-i Monochrome Astronomical Camera	23
2.2.10 FLIR S40	24
2.2.11 FLIR T620	24
2.3 Preliminary Laboratory Testing	26
2.3.1 Blackbody Temperature Stability Tests	26
2.3.2 Image Capture Sequence	28
2.3.3 Camera Testing	28
2.3.4 Atmospheric Conditions	31
2.3.5 Conclusions	34
2.4 Primary Laboratory Testing	35
2.4.1 Dome Cam	36
2.4.2 Silver Cam	36
2.4.3 Starview Cam	37
2.4.4 Swann Cam	39
2.4.5 Zosi Cam	39
2.4.6 Logitech Cam	41
2.4.7 Conclusions	43
2.5 Field Campaign Deployment (Stromboli Volcano)	45
2.5.1 Location	45
2.5.2 Equipment Used	45
2.5.3 Results	46

2.5.4 Limitations	53
Section 3: Automatic Detection Algorithm Development.....	55
3.1 Introduction.....	55
3.2 Preliminary Investigation: Stromboli Volcano.....	56
3.2.1 Image Acquisition.....	56
3.2.2 Manual Observations	57
3.2.3 Automatic Algorithm Analysis.....	57
3.2.4 Automatic Eruption Activity Detection.....	60
3.3 Primary Investigation: Shiveluch volcano.....	61
3.3.1 Image Acquisition.....	61
3.3.2 Algorithm Development	61
3.3.3 Algorithm Application.....	73
3.4 Conclusions.....	80
Section 4: Automatic Detection Algorithm Testing.....	83
4.1 Introduction.....	83
4.2 Algorithm Improvements.....	84
4.2.1 Grayscale Image Classification.....	84
4.2.2 Duplicate Image Classification.....	85
4.2.3 Average Background Image	86
4.2.4 Gaussian Mixture Model Test.....	87
4.2.5 Log-Difference Image.....	87
4.3 Algorithm Application.....	88
4.3.1 Shiveluch Volcano, Russia	88
4.3.2 Fuego Volcano, Guatemala.....	94
4.3.3 Popocatepetl Volcano, Mexico.....	99

4.3.4 Stromboli Volcano, Italy.....	104
4.4 Conclusions.....	109
4.4.1 Volcano Tests.....	109
4.4.2 Image Tests.....	112
4.4.3 Future Work.....	112
Section 5: Conclusions.....	115
5.1 Summary of Findings.....	115
5.2 Method Development.....	116
5.3 Method Application.....	118
5.3.1 Shiveluch Volcano.....	119
5.3.2 Fuego Volcano.....	120
5.3.3 Popocatepetl Volcano.....	120
5.3.4 Stromboli Volcano.....	120
5.4 Future Work.....	122
Section 6: References.....	125

List of Figures

	Page
Figure 1.1. View of Stromboli volcano, Italy, including an inset map of its location in the Tyrrhenian Sea.....	5
Figure 1.2. Illustrative representation of Type 1, 2a, and 2b eruption characteristics at Stromboli volcano.....	7
Figure 1.3. View of Shiveluch volcano, Russia, including an inset map showing its location on the Kamchatka Peninsula in eastern Russia.....	8
Figure 1.4. View of Fuego volcano, Guatemala, including an inset map showing its location in Central America.....	10
Figure 1.5. View of Popocatepetl volcano, Mexico, including an inset map showing its location in Central America.....	12
Figure 2.1. Planck’s law representing the relationship between spectral radiance and wavelength at blackbody temperatures of 200, 400, 600, 800, and 1,000°C.....	16
Figure 2.2. Laboratory schematic showing the layout and observation distances for the spectral response experiments.....	17
Figure 2.3. Images of six cameras acquired for the spectral response experiments.....	18
Figure 2.4. Images of four additional cameras used during this investigation.....	19
Figure 2.5. Preliminary results from the Dome cam spectral response experiments.....	32
Figure 2.6. Preliminary results from the Silver cam spectral response experiments.....	33
Figure 2.7. Side by side image comparison from the Dome and Silver cameras at three different temperature during spectral response testing.....	34
Figure 2.8. Spectral response experiment results at one, two, four, six, and eight-meter distances showing the maximum pixel brightness for A) Dome cam, B) Silver cam, and C) Starview cam..	38

Figure 2.9. Spectral response experiment results at one, two, four, six, and eight-meter distances showing the maximum pixel brightness for A) Swann cam and B) Zosi cam.....	40
Figure 2.10. Spectral response experiment results at one, two, four, six, and eight-meter distances showing the maximum pixel brightness for the Logitech cam at different exposure settings.....	42
Figure 2.11. Summary of all results from the spectral response experiments showing each camera's initial detection and saturation temperatures at all five distances.....	44
Figure 2.12. Coincidental image sequence of an event at Stromboli volcano on June 29, 2013 from Location 1.....	48
Figure 2.13. Coincidental images of an event at Stromboli volcano on June 29, 2013 at 20:23:28 UTC from Location 2.....	50
Figure 2.14. An eruption sequence captured by the Ethernet cam at Stromboli volcano from the summit at Location 3 on June 23, 2013.....	52
Figure 3.1. Detected events at Stromboli volcano from manual analysis and automatic image processing between March 30 and April 9, 2014.....	59
Figure 3.2. Automatically identified events at Stromboli volcano from March 30 to August 31, 2014.....	60
Figure 3.3. Overview flow diagram of HAW algorithm analysis process.....	62
Figure 3.4. Predefined region of interest to remove erroneous sources of data using a binary system.....	65
Figure 3.5. Example images demonstrating the effect of varying Gaussian kernel sizes (σ) from 1 to 5) used to smooth noise in the image before the Canny edge detection process.....	67
Figure 3.6. Example images of the edge detection process.....	68
Figure 3.7. Results from the Gaussian Mixture Model test applied to a thermally incandescent image.....	72
Figure 4.1. Automatic volcanic activity detection algorithm results from Shiveluch volcano.....	91

Figure 4.2. Example images where the algorithm falsely detected activity or failed to detect activity at Shiveluch volcano.....93

Figure 4.3. Automatic volcanic activity detection algorithm results from Fuego volcano.....95

Figure 4.4. A seismogram from Fuego volcano courtesy of INSIVUMEH on March 1, 2015 showing increased activity and initial stages of a new paroxysm event.....96

Figure 4.5. Example images where the algorithm falsely detected activity or failed to detect activity at Fuego volcano.....98

Figure 4.6. Automatic volcanic activity detection algorithm results from Popocatepetl volcano.100

Figure 4.7. Example images demonstrating the increase in thermally incandescent activity in fifteen minute intervals leading up to a large paroxysm event on April 18, 2016.....101

Figure 4.8. Example images where the algorithm falsely detected activity of failed to detect activity at Popocatepetl volcano.....103

Figure 4.9. Sequence of images showing an individual event at Stromboli volcano that is falsely identified as five unique events by the HAW algorithm on March 30, 2014.....108

List of Tables

	Page
Table 2.1 Summary table of all camera information and specifications used in this investigation..	25
Table 2.2 Recorded blackbody temperature stability timings.....	27
Table 2.3 Dome cam spectral response experiment results.....	30
Table 2.4 Silver cam spectral response experiment results.....	31
Table 2.5 Atmospheric temperature (°C) and relative humidity (%) as recorded throughout the preliminary spectral response experiments.....	31
Table 2.6 Summary of primary spectral response experiment results from all cameras.....	41
Table 3.1 Manual and automatic analysis results from an eleven-day period of eruptions at Stromboli volcano.....	59
Table 3.2 Confusion matrix results from the comparison of a manual analysis and the automatic algorithm analysis for NIR webcam images at Shiveluch volcano from 2015.....	74
Table 3.3 Statistical classification confusion matrix results from the HAW algorithm analysis of NIR webcam images at Shiveluch volcano from 2015.....	76
Table 3.4 Statistical classification confusion matrix results from individual tests in the HAW algorithm analysis of NIR webcam images at Shiveluch volcano from 2015.....	77
Table 4.1 Confusion matrix results from the comparison of a manual analysis and the HAW algorithm analysis for thermal camera webcam images at Stromboli volcano from March 30 to April 7, 2014.....	105
Table 4.2 Statistical classification results from the HAW algorithm analysis for thermal camera webcam images at Stromboli volcano from March 30 to April 1, 2014.....	107

List of Acronyms

Above Sea Level (ASL)

Advanced Band Imager (ABI)

Advanced Himawari Imager (AHI)

Advanced Very High Resolution Radiometer (AVHRR)

Analog-to-Digital Converter (ADC)

Application Program Interface (API)

Before Present (BP)

Central Processing Unit (CPU)

Centro Nacional de Prevención de Desastres (CENAPRED)

Charge-Coupled Device (CCD)

Closed-Circuit Television (CCTV)

College of Natural Science and Mathematics (CNSM)

Complementary Metal Oxide Semiconductor (CMOS)

Critical Success Index (CSI)

Events Per Day (EPD)

False Alarm Rate (FAR)

False Negative (FN)

False Positive (FP)

Field Of View (FOV)

Forward Looking Infrared Radiometer (FLIR)

Frames Per Second (FPS)

Gaussian Mixture Model (GMM)

Geophysical Institute Graduate Student Association (GIGSA)

Geostationary Operational Environmental Satellite (GOES)

Global Positioning System (GPS)

Graphical User Interface (GUI)

Graphics Processing Unit (GPU)

Harrild and Webley (HAW)

Innovative Technology and Education (ITE)

Instituto Nacional de Sismología, Vulcanología, Meteorología E Hidrología (INSIVUMEH)

Istituto Nazionale di Geofisica e Vulcanologia' (INGV)

JavaScript Object Notation (JSON)

Kamchatka Volcanic Eruption Response Team (KVERT)

Kiloannum, thousand years ago (ka)

Koninklike Luchtvaart Maatschappij (KLM)

Light-Emitting Diode (LED)

Logitech Webcam Software (LWS)

Megapixel (MP)

Michigan Technological University (MTU)

Micrometers or microns (μm)

National Aeronautics and Space Administration (NASA)

National Oceanic and Atmospheric (NOAA)

Near-Infrared (NIR)

Northeast (NE)

Probability Of Detection (POD)

Probability of Event (P_E)

Probability of False Alarm (P_{FA})

Probability Final (P_F)

Radio Corporation of America (RCA)

Random-Access Memory (RAM)

Red, Green, Blue (RGB)

Region Of Interest (ROI)

Santa Barbara Instrument Group (SBIG)

Southwest (SW)

Success Rate (SR)

Television Lines (TVL)

Thermal Infrared (TIR)

Trans-Mexican Volcanic Belt (TMVB)

True Negative (TN)

True Positive (TP)

Undergraduate Research and Scholarly Activity (URSA)

Universal Serial Bus (USB)

University of Alaska Fairbanks (UAF)

Volcanic Explosivity Index (VEI)

Volcano Observatory Notice for Aviation (VONA)

World Organization of Volcano Observatories (WOVO)

West Ridge Research Building (WRRB)

Acknowledgements

This work was primarily funded by a National Aeronautics and Space Administration (NASA) collaborative research grant no. NNX14AQ96G as a sub-contract from the University of Pittsburgh (Principal investigator: Professor Michael Ramsey).

Funding for this investigation was also provided by the University of Alaska Fairbanks (UAF) Graduate School in the form of a Thesis Completion Fellowship.

Additional financial support to present results from this investigation at scientific conferences has been provided by the College of Natural Science and Mathematics (CNSM) and the Geophysical Institute Graduate Student Association (GIGSA).

Further financial support to purchase additional near-infrared (NIR) cameras was provided by the Undergraduate Research and Scholarly Activity (URSA) Innovative Technology and Education (ITE) award.

I would primarily like to thank my advisor, Dr. Peter Webley, who has provided a substantial amount of help and support throughout my time at UAF, without whom I would not be in the position I am today.

Additionally, I would like to thank my other committee members, Dr. Jonathan Dehn and Dr. Franz Meyer, for all the help they have provided me throughout my thesis, and Donavan Kienenberger for invaluable technological assistance when desperately needed.

I would also like to specifically thank my parents who have always believed in me and supported my time in Alaska, and for reading through my entire thesis several times to provide constructive edits.

Finally, I would like to thank all my other friends and family members who have supported me over the years and kept pushing me during tough times towards my final graduation goal.

Section 1: Introduction

1.1 Background

Across the globe, approximately 1,300 volcanoes have erupted throughout the Holocene, while almost 600 have erupted within the last 500 years (Siebert et al., 2011). Although volcanism on Earth predominantly occurs on the sea floor, with submarine volcanism accounting for an estimated 83% of all eruptions (Crisp, 1984), the remaining subaerial volcanoes have significant potential to adversely affect millions of people around the world. An estimated 20 million people live within five kilometers of an active volcano, which has erupted at least once in the last 500 years and could be significantly impacted by small to moderate sized eruptions, rated a two or three on the Volcanic Explosivity Index (VEI) (see Newhall and Self, 1982 for VEI description) (Siebert et al., 2011). There have been approximately 280,000 fatalities reported in historic records relating to volcanic eruptions from 533 individual events, with almost 100,000 occurring in the 20th century (Auker et al., 2013). Although this number is relatively low compared to other natural disasters such as flooding, where the same number of deaths occurred in only the last decade of the 20th century (Jonkman, 2005), the risk will increase with time given the rate of population growth and rapid urbanization (Small and Naumann, 2001). Most of these fatalities (74%) occurred from only 13 of the listed eruptions, but were predominantly located in developing countries where regular volcanic monitoring may be limited or absent altogether (Auker et al., 2013).

Loss of life isn't the only significant impact volcanic eruptions can have on population centers and local infrastructure. Aviation safety and the impact of volcanic ash on daily operations has become the significant focus of volcanic monitoring efforts in recent years. From 1970 to 2010, almost 400 separate ash clouds reached 6,000 m (~20,000 ft) above sea level (ASL) or higher across the North Pacific (Webley et al., 2012). Globally, many ash clouds reached altitudes greater than 9,000 m (~30,000 ft) ASL, averaging approximately 20 per year (Miller and Casadevall, 2000), which can significantly impact aviation safety. Prior to 2004, 105 aviation ash cloud encounters were documented (Guffanti et al., 2004), while on the ground, 101 airports in 28 countries were affected from 171 unique eruptions (Guffanti et al., 2009). Fortunately, most aviation ash cloud encounters have had minimal impacts on aircraft, but two famous examples

where all four engines failed are detailed here. In 1982, an aircraft (British Airways Flight 9) encountered an ash cloud from Galunggung volcano in Indonesia, resulting in a loss of power to all four engines and a terrifying glide for 13 minutes. Fortunately, engine power was restored at an altitude of 4,100 m (13,500 ft) ASL, shortly before the pilot needed to decide on attempting a risky ditching in the ocean (Hanstrum and Watson, 1983). The second encounter occurred in 1989 when an aircraft (Koninklike Luchtvaart Maatschappij (KLM) Flight 867) encountered an ash cloud from Redoubt volcano, Alaska, USA as the aircraft descended into Anchorage airport. Again, all four engines lost power and the plane rapidly descended toward the local mountain range until several engines were restarted allowing for a successful landing. However, severe damage to the exterior of the plane and the replacement of all four engines resulted in \$80 million worth of damage (Casadevall, 1994), \$157 million today after accounting for inflation.

Major aircraft encounters with ash clouds are fortunately rare, but the economic impact of volcanic eruptions is increasing with time as the population increases and globalization expands rapidly. Miller and Casadevall (2000) estimate that approximately \$250 million of damage to aviation engines, aircraft, and avionics was caused by volcanic eruptions between 1980 and 1998. Most famously, the April 2010 eruption of Eyjafjallajökull volcano in Iceland caused an ash cloud to drift over European airspace, resulting in the closure of large areas of this airspace for eight days (Gudmundsson et al., 2012). This resulted in the cancellation of approximately 110,000 flights, leaving an estimated 10 million passengers stranded, and the financial impact to the aviation industry is estimated at \$1.8 billion, with a total cost of approximately \$5 billion once all other economic impacts are taken into consideration (Parker, 2015).

Considering all the adverse effects volcanic eruptions can have on human life, aviation, and the economy, monitoring as many volcanoes as possible is paramount to attempt to mitigate the impacts of future eruptions.

1.2 Volcano Monitoring

As volcanoes have the potential to cause significant adverse effects, the need to monitor changes in behavior is paramount to effectively mitigating these impacts. If monitored successfully, appropriate warnings and vital information can be promptly disseminated to necessary individuals who could be affected, potentially saving lives or limiting financial impacts.

The earliest example of systematic monitoring at a volcano observatory is the ‘Osservatorio Vesuviano’ in Italy, completed in 1845, which now constitutes the Naples branch of the ‘Istituto Nazionale di Geofisica e Vulcanologia’ (INGV) (Sparks et al., 2012). Today, the World Organization of Volcano Observatories (WOVO, <http://www.wovo.org/>) has 80 members, with the ability to monitor a significant proportion of global volcanic activity.

Several methods exist to monitor this activity including, but not limited to, satellite based remote sensing; ultraviolet (e.g. Thomas and Prata, 2011), infrared (e.g. Webley et al., 2009a), microwave (e.g. Meyer et al., 2015), ground based remote sensing; ultraviolet (e.g. Nadeau et al., 2011), visible (e.g. Carling et al., 2015), infrared (e.g. Spampinato et al., 2011), global position system (GPS) and tiltmeters (e.g. Larson et al., 2010), seismic (e.g. Lyons and Waite, 2011), and infrasound (e.g. Arámbula-Mendoza et al., 2013). Individually, these methods are powerful monitoring tools and when combined together to form a complete monitoring network, large volcanic eruptions, in areas where monitoring is commonplace, are rarely unforeseen.

1.3 Research Aim

Despite a global effort to effectively monitor volcanic activity, where countries or local states are responsible to develop monitoring capabilities for volcanoes in their assigned area (Romero, 2004), availability of finances and geographical location often results in inadequate monitoring for volcanoes. Many modern techniques are expensive, sometimes tens of thousands of dollars to build a ground network or to purchase a single instrument, and are rarely deployed to volcanoes in remote locations where they may potentially be destroyed and are difficult to replace. Aspinnall et al. (2011) determined that of the 441 active volcanoes in 16 developing countries, 384

of them have rudimentary or no monitoring at all, including 65 volcanoes identified as ‘high risk’ to large local populations. Satellite remote sensing (e.g. Webley et al., 2009a) can attempt to fill this void given the large spatial distribution capabilities, but is rarely applied in real-time monitoring in these locations.

The methods proposed here aim to monitor changes in volcanic activity using ground based remote sensing, specifically imagery from webcams and other permanently installed cameras aimed at volcanoes. As telemetered webcam imagery is relatively inexpensive to install compared to other monitoring methods, it is often one of the most cost effective data types available to volcanic observatories, which is usually only qualitatively used in their daily data analysis and provided to the public through the observatory’s website. Unfortunately, this type of data is extremely underutilized and often classified as ‘dark data’ (Heidorn, 2008), meaning observatories collect, process, and store this information daily, but fail to use it for other purposes. Patrick et al. (2010) demonstrates a similar approach to the methods proposed here by developing an algorithm to identify and quantify volcanic incandescence in webcam imagery from Kilauea volcano, Hawai’i. Nighttime images are analyzed and combined to create a composite image of thermally incandescent activity, which is manually analyzed and reported in the daily monitoring logs. However, automatically detecting these increases in activity in near real-time is still absent.

This investigation aims to determine several objectives, listed below:

- Determine how specific cameras respond to increased light levels from thermal incandescence.
- Determine approximate eruption temperatures from pixel brightness using collocated cameras.
- Develop an automatic detection algorithm for webcam imagery to identify thermal incandescence, with a focus on reduced financial cost and ease of deployment.
- Automatically detection volcanic eruptions by analyzing webcam imagery.

1.4 Volcanoes of Interest

Several volcanoes were chosen for this investigation that represent a variety of geolocations and eruptive styles, ranging from the sub-polar regions of the North Pacific to the tropical regions of Central America to the islands of the Mediterranean and volcanologically, from Strombolian to

sub-Plinian eruptive styles. Each particular chosen volcano has a webcam capturing images of eruptions at relatively frequent intervals that is freely available online. This has allowed the collection of a significant dataset of images spanning several years for each volcano.

1.4.1 Stromboli Volcano, Italy

Stromboli volcano is the northernmost volcanic island of the Aeolian archipelago, located in the Tyrrhenian Sea, southern Italy, resulting from the subduction of the African plate below the Eurasian plate throughout the Quaternary (Gasparini et al., 1982). The island is the subaerial section of a larger volcanic edifice, rising approximately 2,000 m (6,562 ft) from the sea floor, reaching an elevation of approximately 924 m (3,032 ft) ASL (Rosi et al., 2000).



Figure 1.1. View of Stromboli volcano, Italy, including an inset map of its location in the Tyrrhenian Sea. Markers identify the currently active craters, the location of the permanently installed thermal infrared camera, and three field campaign locations. The oblique view is courtesy of Google Earth™.

The volcano itself is over 200,000 years old (Gillot and Keller, 1993) and has experienced several flank collapses during its growth, most notably the Upper Vancori (~13 ka) and Neo-Stromboli ($5,600 \pm 3,300$ ka) (Arrighi et al., 2004). It has suffered two other, albeit less destructive, flank collapses (Pizzo Sopra la Fossa and Sciara del Fuoco) of unknown age, although these events occurred after 5,000 years before present (BP) (Tibaldi, 2001). Due to an exceptional and almost continuous persistence in activity throughout the last 2,500 years, Stromboli has earned a reputation as the ‘Lighthouse of the Mediterranean’ (Bullard, 1980; Francis, 1993).

Current activity at Stromboli occurs from three main craters (~750 m ASL) within the Sciara del Fuoco, a horseshoe-shaped depression on the northwest flank of the cone (Rosi et al., 2000) which collapsed less than 5,000 ka, involving approximately $1,800 \text{ m}^3$ of rock (Kokelaar and Romagnoli, 1995) (Figure 1.1). Strombolian activity is named after the activity at this particular volcano and consists of periodic low-energy eruptions that usually last for less than 20 seconds, occurring approximately every 10 to 20 minutes. This results in incandescent scoriae, ash, and magma lumps to be ejected to heights of several tens of meters around the rims of the summit craters (Speranza et al., 2008; Taddeucci et al., 2015).

Several studies have attempted to quantify these rates and styles of eruptions at Stromboli over the years to determine their behavioral patterns. Blackburn et al. (1976) were one of the first to identify eruption gas velocities here, with 31 m s^{-1} for one vent, and 56 m s^{-1} for another. Ripepe et al. (1993) used high-speed infrared film to capture eruptions and quantify the mass distributions of material ejected compared to recorded seismic signals. Hort et al. (2003) used Doppler radar to estimate the dynamics of the eruptions at Stromboli and found comparatively similar results to the previous studies listed here. Interestingly, Hort et al. (2003) noticed the length of eruptions at Stromboli increased by a factor of two after a rain storm passed through the area, suggesting the water significantly affects the eruption styles here.

A large paroxysm event at Stromboli occurred on April 3, 2003 (Bertagnini et al., 2011), which could have potentially changed the eruption dynamics seen here. With 70% of data recorded after this event, Patrick et al. (2007) conducted an investigation to determine the different styles of eruptions seen at Stromboli and to specifically define several explosion characteristics. Using a Forward Looking Infrared Radiometer (FLIR) thermal infrared (TIR) camera, Patrick et al. (2007) imaged approximately 350 unique events over four years at 30 frames per second (FPS), allowing

two distinct types of eruptions to be identified (Figure 1.2). Patrick et al. (2007) state the origin of the ash in Type 2 eruptions results from backfilled material in the crater, caused by previously erupted material falling back into the vent or material breaking from the crater rim during an explosion and ‘choking’ the vent. This ash may detrimentally affect observations of an eruption by obscuring the view of the event.

This ‘normal’ activity is occasionally interrupted by periods of quiescence, the effusion of lava flows, or larger eruptive events known as ‘paroxysms’, which usually occur without accompanying lava flows, of which twelve have been observed during the last century (Barberi et al., 1993). The most recent period of ‘abnormal’ activity occurred throughout the summer of 2014, where a significant increase in the frequency of explosions was recorded in June, leading to the onset of an effusive eruption on August 6, 2014. This activity, combined with persistent explosive events and variable rates of degassing from the craters, continued for several weeks before waning and officially ceasing on November 17, 2014 (Rizzo et al., 2015).

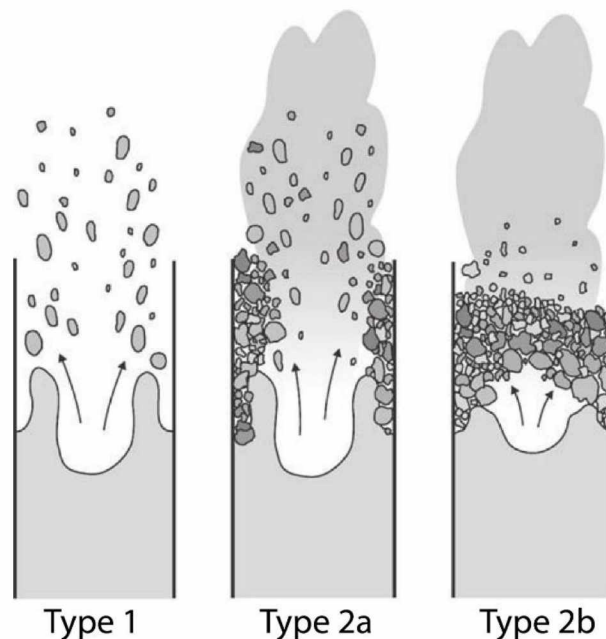


Figure 1.2. Illustrative representation of Type 1, 2a, and 2b eruption characteristics at Stromboli volcano. Type 1 eruptions are dominated by coarse ballistic particles with little to no ash, while Type 2 eruptions are dominated by an optically-thick, ash-rich plume. Type 2 eruptions are further subcategorized into Type 2a (contains a large number of ballistic particles with high gas thrust velocities ($> 15 \text{ m s}^{-1}$)) and Type 2b (contains few, if any, ballistic particles with more buoyant velocities ($< 15 \text{ m s}^{-1}$)). Taken from Patrick et al. (2007).

1.4.2 Shiveluch Volcano, Russia

Shiveluch is the northernmost active volcano on the Kamchatka Peninsula, situated in eastern Russia, in the Central Kamchatka Depression at the supposed conjunction zone of the Aleutian and Kuril-Kamchatka island arcs (Ponomareva et al., 1998). The volcano edifice consists of two main parts; ‘Old Shiveluch’, a horseshoe shaped caldera, and ‘New Shiveluch’, the currently active volcanic dome (Figure 1.3).



Figure 1.3. View of Shiveluch volcano, Russia, including an inset map showing its location on the Kamchatka Peninsula in eastern Russia. Markers identify Young and Old Shiveluch volcanoes, and show the location of the permanently installed webcam. The oblique view is courtesy of Google Earth™.

‘Old Shiveluch’ consists of the ruins of a large stratovolcano of Pleistocene age which reached over 4,000 m (13,123 ft) ASL before its destruction (~10 ka, after the last glaciation), but has a current height of 3,335 m (10,942 ft) ASL (Belousov et al., 1999). The exact date of this collapse is contested among researchers, as Melekestsev et al. (1991) suggest the event occurred before the last glacial maximum (23 - 24 ka), whereas Pevzner et al. (2013) suggest the collapse could have occurred during the glaciation (15 - 16 ka).

‘Young Shiveluch’ is composed of multiple extrusive domes at a height of 2,800 m (9,186 ft) ASL, surrounded by short but thick lava flows. Typical activity of this volcano throughout the Holocene has been characterized by Plinian eruptions, alternating with periods of dome growth (Ponomareva et al., 1998). Written records of the activity at Shiveluch date back to 1739 with reports of eruptions in 1739, 1790, 1793, and between 1790 and 1810, but this information lacks credibility and cannot be confirmed (Gorshkov and Dubik, 1970). More recent and reliable records began in 1854 when a large ash rich eruption was recorded, but Ponomareva et al. (1998) question the accuracy of that recording too. Later eruptions were recorded in 1879 - 1883, 1897 - 1898, 1905, 1928 - 1929, and 1944 - 1950, resulting in extrusive dome growth and ash fall. The most recent Plinian eruption at Shiveluch occurred on November 12, 1964, where an avalanche deposit resulting from a sector collapse covered an approximate 100 km² area (Gorshkov and Dubik, 1970).

Since this most recent Plinian eruption, approximately 0.3 km³ of magma has been discharged during three main eruptive phases; 1) 1980 - 1981, 2) 1993 - 1995, and 3) 2001 - 2004 (Dirksen et al., 2006). A current additional phase of dome extrusion commenced in 2006, accompanied by minor explosive activity (van Manen et al., 2012), and continues at the time of writing (March 2017).

1.4.3 Fuego Volcano, Guatemala

Fuego volcano (Figure 1.4), meaning ‘fire volcano’ in Spanish, is located in the western highlands of Guatemala, situated in the second section (of eight) of the Central American volcanic front (Carr et al., 2007). It is a basaltic-andesite stratovolcano that reaches a height of 3,800 m (12,467 ft) ASL and is the southernmost of four vents that form the north-south trending Fuego-Acatenango stratovolcano complex (Lyons and Waite, 2011).

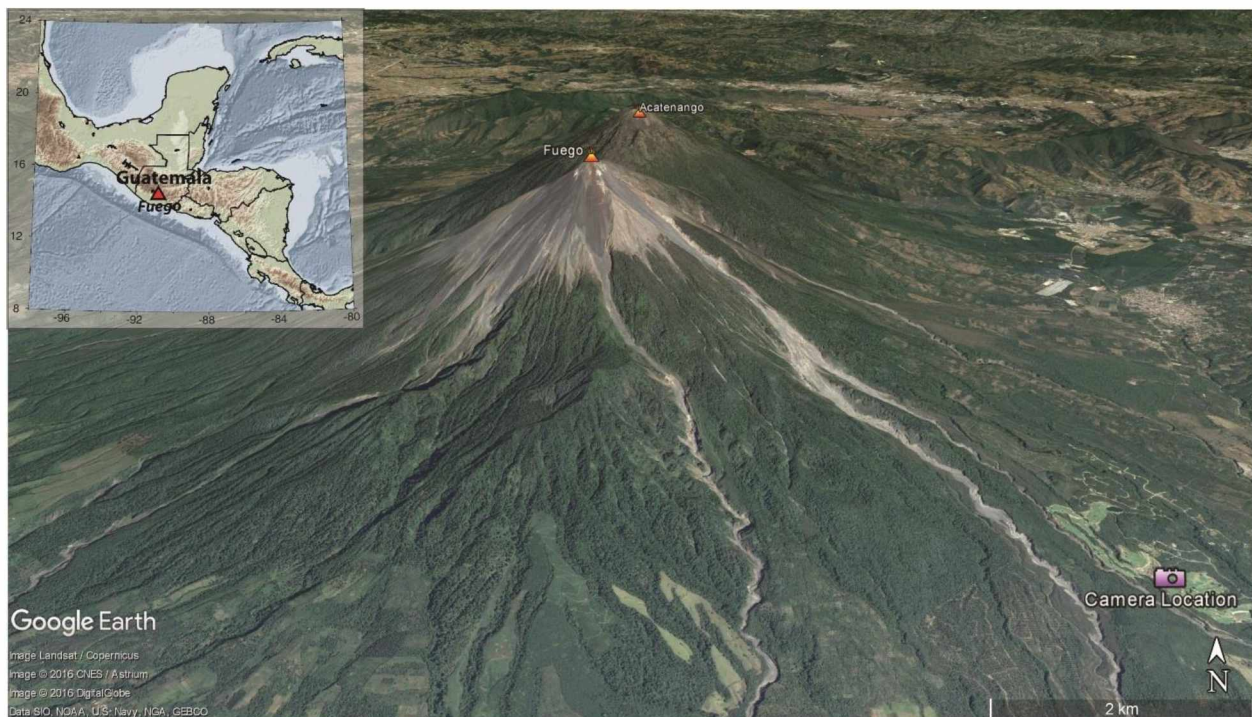


Figure 1.4. View of Fuego volcano, Guatemala, including an inset map showing its location in Central America. A marker identifies the location of the permanently installed webcam. The oblique view is courtesy of Google Earth™.

The oldest recorded dates in this complex are approximately 230,000 years old (Chesner and Halsor, 2006), although the main period of growth at the Fuego edifice didn't begin until 84,000 years ago (Vallance et al., 2001). However, Martin and Rose (1981) suggest the eruptions at Fuego, at their current rate, may have begun as recently as 13,000 years ago.

In the last 500 years, Fuego has experienced more than 60 sub-Plinian eruptions, the most recent of which occurred in 1974, producing a column of ash that reached 17,000 m (55,774 ft) ASL, with tephra deposits thicker than 20 cm reported over 50 km from the main vent (Rose et al., 2007). Since this time, Fuego has experienced over 20 years of quiescence, with only one notable exception occurring in January 1987, which was ended in May 1999 by a VEI 2 eruption (Rodríguez et al., 2004).

Activity since 1999 is described as continuous Strombolian activity, characterized by periods of lava effusion accompanied by discrete explosions, and periods of explosions with no occurrence of lava effusion (Lyons et al., 2010). The explosions observed during periods of effusion contained varying levels of ash and eject ballistic material between 50 m (164 ft) and 300 m (984 ft) above the crater rim, whereas explosions associated with non-effusive periods were typically louder with a higher ash content (Erdem and Waite, 2013).

1.4.2 Popocatépetl Volcano, Mexico

Popocatépetl is a stratovolcano located in Mexico (Figure 1.5), in the central zone of the Trans-Mexican Volcanic Belt (TMVB), resulting from the Cocos and Rivera plates subducting beneath the North American plate (Arámbula-Mendoza et al., 2010). At a height of 5,452 m (17,887 ft) ASL, it is one of the highest active volcanoes in the world and located close to major population centers, including Mexico City, Puebla, and Tlaxcala. These areas are home to over 20 million people with Popocatépetl widely regarded as one of the most dangerous volcanoes in the world (Kuznetsov and Koulakov, 2014).

Activity here has varied greatly throughout its history, ranging from large Plinian eruptions to smaller Strombolian activity. Evidence of five Plinian style eruptions exists from the last 23,000 years, three of which occurred within 5,000 years BP (Siebe et al., 1996). A VEI 4 eruption occurred in 823, the largest eruption during human history in the area, with several moderate VEI 1 - 3 eruptions occurring since, including periods of unrest in 1925 - 1927 (VEI 2), 1994 - 1995 (VEI 2), and 1996 - 2003 (VEI 3) (Kuznetsov and Koulakov, 2014).

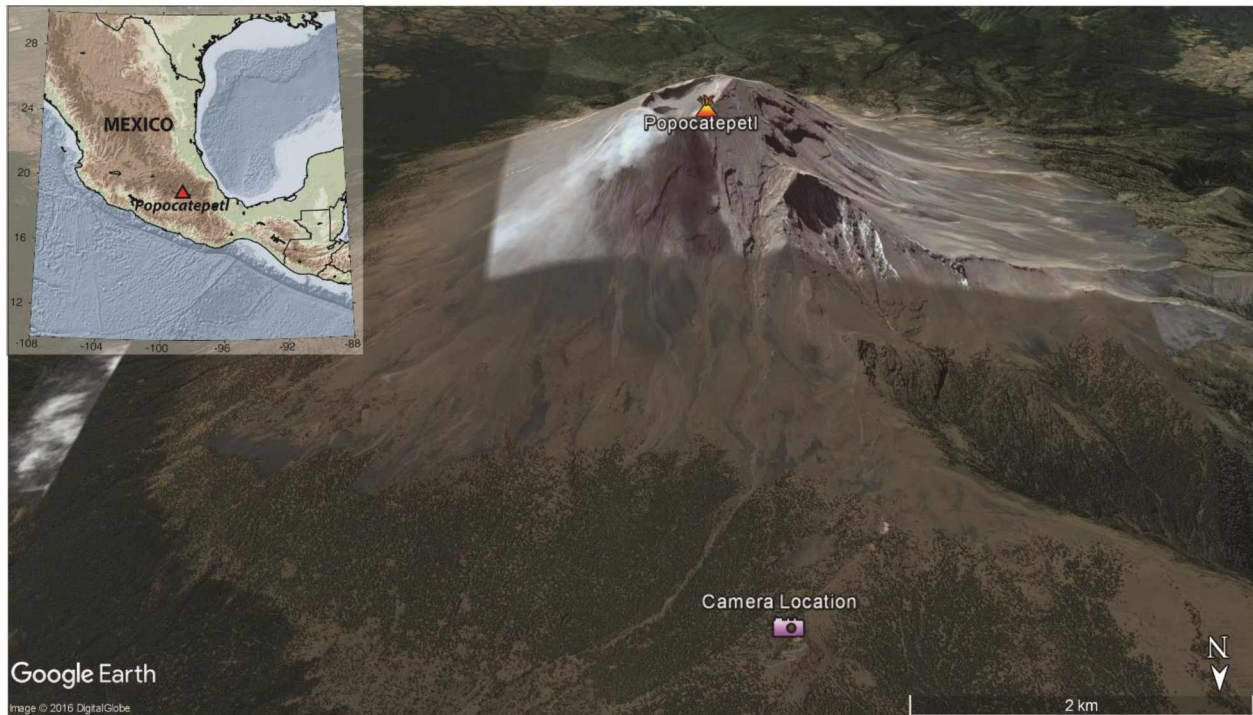


Figure 1.5. View of Popocatepetl volcano, Mexico, including an inset map showing its location in Central America. A marker identifies the location of the permanently installed webcam. The oblique view is courtesy of Google Earth™.

Recent activity consists of repetitive dome-building and dome-destroying episodes. The dome-destroying events are often accompanied by large Vulcanian explosions, with ash plumes reaching up to 8,000 m (26,246 ft) above the crater rim (Zobin and Martínez, 2010). This activity is often described as a VEI 3 eruption or less, resulting in a limited threat to the nearby population (Arámbula-Mendoza et al., 2010).

1.5 Overview

The second section of this thesis will demonstrate laboratory testing of a suite of acquired cameras to determine their spectral response. By understanding exactly how a specific camera functions under a variety of controlled conditions, its behavior during field campaign deployments can be better interpreted and related to actual volcanic activity, to determine levels of activity via proxy.

The third section will focus on the development of the automatic change detection algorithm. A detailed analysis of each step will be demonstrated as the analysis is performed on webcam imagery available from Shiveluch volcano, Russia.

The fourth section will cover fine tuning of and improvements to the algorithm and its application to other target volcanoes including Stromboli volcano in Italy, Fuego volcano in Guatemala, and Popocatepetl volcano in Mexico, as well as higher temporal resolution data from Shiveluch volcano in Russia.

The fifth and final section will cover general conclusions of this investigation, the issues encountered throughout development, and the prospect of future applications for this volcanic activity change detection algorithm.

Section 2: Spectral Response Experiments

The initial stage of this investigation focuses on determining the spectral response of several acquired cameras in a controlled laboratory setting. Each acquired camera imaged an infrared calibration source, also known as a blackbody which was used as a controlled heat source, to document its response. A blackbody is a physical source that absorbs almost all electromagnetic radiation passed into it, with no energy reflected, for all wavelengths and incident angles. A true blackbody has a transmission value of zero, an absorption value of one, and a reflection value of zero, and although perfect blackbodies are theoretical, infrared calibration sources are extremely close to these values. By controlling the heating element of a blackbody, known temperature outputs can be recorded. These values are emitted according to 'Planck's law' (Figure 2.1), which expresses spectral radiance output as a function of observational wavelength and surface temperature for an object with an emissivity of one ($\epsilon = 1$). This allows all the purchased cameras to image the same 'perfect' blackbody at various temperatures to determine their spectral response, which can be used as a proxy for identifying increased temperatures at volcanoes. As an example, humans can see radiation in the visible portion of the electromagnetic spectrum (0.4 to 0.7 microns, μm), meaning objects approximately 600°C will appear dusky red, and anything colder will not be discernable from background temperatures. However, cameras that operate at different wavelengths, such as TIR cameras which operate in the 9 to 14 μm range, can measure lower temperatures of cooler objects and the higher temperatures of objects emitting stronger thermal signals. This allows the imaging of objects with a thermal emittance far below the required threshold for humans to be able to detect a change in thermal output, and is often utilized for safety and crime prevention methods (e.g. Stahl and Schoppmann, 2000). The cameras also image the blackbody at several distances for the entire temperature range, to determine the effect of pixel size on the spectral response. At closer distances, the blackbody target will be larger than an individual pixel, but at greater distances, the target will be sub-pixel in size and the brightness values will be a result of pixel integrated brightness between the brighter target and the darker background. The cameras used in this investigation operate in the near-infrared (NIR) portion of the electromagnetic spectrum and have wavelengths of 0.7 to 1.4 μm .

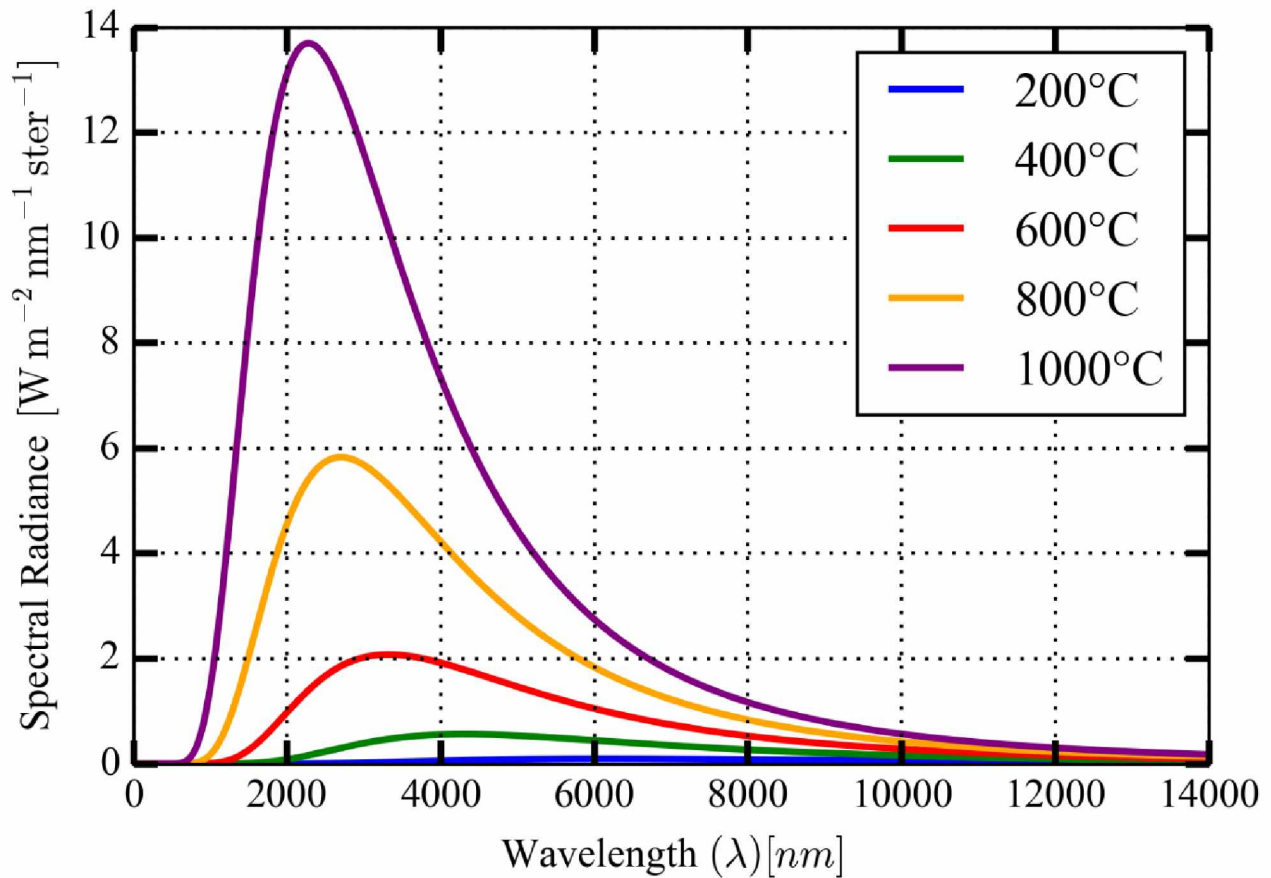


Figure 2.1. Planck's law representing the relationship between spectral radiance and wavelength at blackbody temperatures of 200, 400, 600, 800, and 1,000°C.

2.1 Experimental Set Up

The laboratory used for the controlled spectral response experiments was chosen primarily for its length, but also due to a lack of windows. The room is over eight meters in length, allowing for the cameras to be tested at several distance intervals, and is located in the basement of the West Ridge Research Building (WRRB) at the University of Alaska Fairbanks (UAF), meaning no outside illumination sources are present. A blackbody source was placed at one end of the room, and markers placed on the floor at one, two, four, six, and eight meters (Figure 2.2). Each camera was mounted on a stand to ensure the lens and blackbody cavity lined up perfectly, ensuring the highest level of accuracy and measurement precision. All external sources of illumination in the lab (such as the blackbody temperature display, computer light-emitting diodes (LED's), and

power strip LED's) were covered as well as covering around the door frame with a thick dust sheet to ensure as close to total blackout conditions as possible, leaving the blackbody target as the only source of light. Both the blackbody and chosen camera were connected to a local computer to ensure full automation of this process through a personally developed Python script, removing the likelihood of any human induced errors. Atmospheric room temperature and relative humidity were recorded using a handheld thermometer and hygrometer device at the beginning of each cycle to identify any erroneous data resulting from varying conditions in the room.

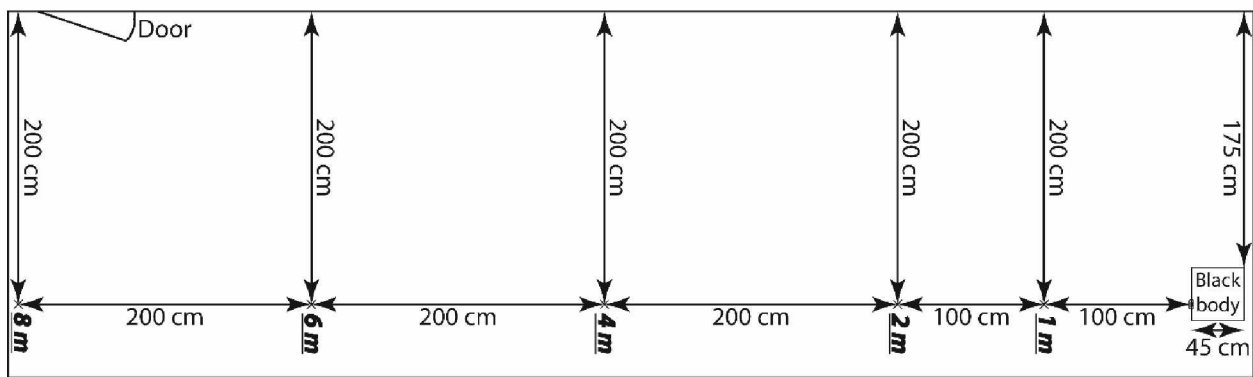


Figure 2.2. Laboratory schematic showing the layout and observation distances for the spectral response experiments.

2.2 Equipment Information

The focus of this investigation is to identify changes in volcanic activity using inexpensive methods to improve monitoring in areas with limited funding, where expensive equipment deployment is unfeasible. Several cameras were purchased for this investigation (Figures 2.3 and 2.4) and ranged in price from \$30 to \$1,000. They all operate in the visible to NIR portions of the electromagnetic spectrum and can capture lower levels of light than can be seen by humans, meaning these cameras are commonly used for household security systems or computer webcams. A summary of all the cameras used in this investigation is shown in Table 2.1. All equipment used for this investigation is publicly available and can be purchased through commercial retailers.

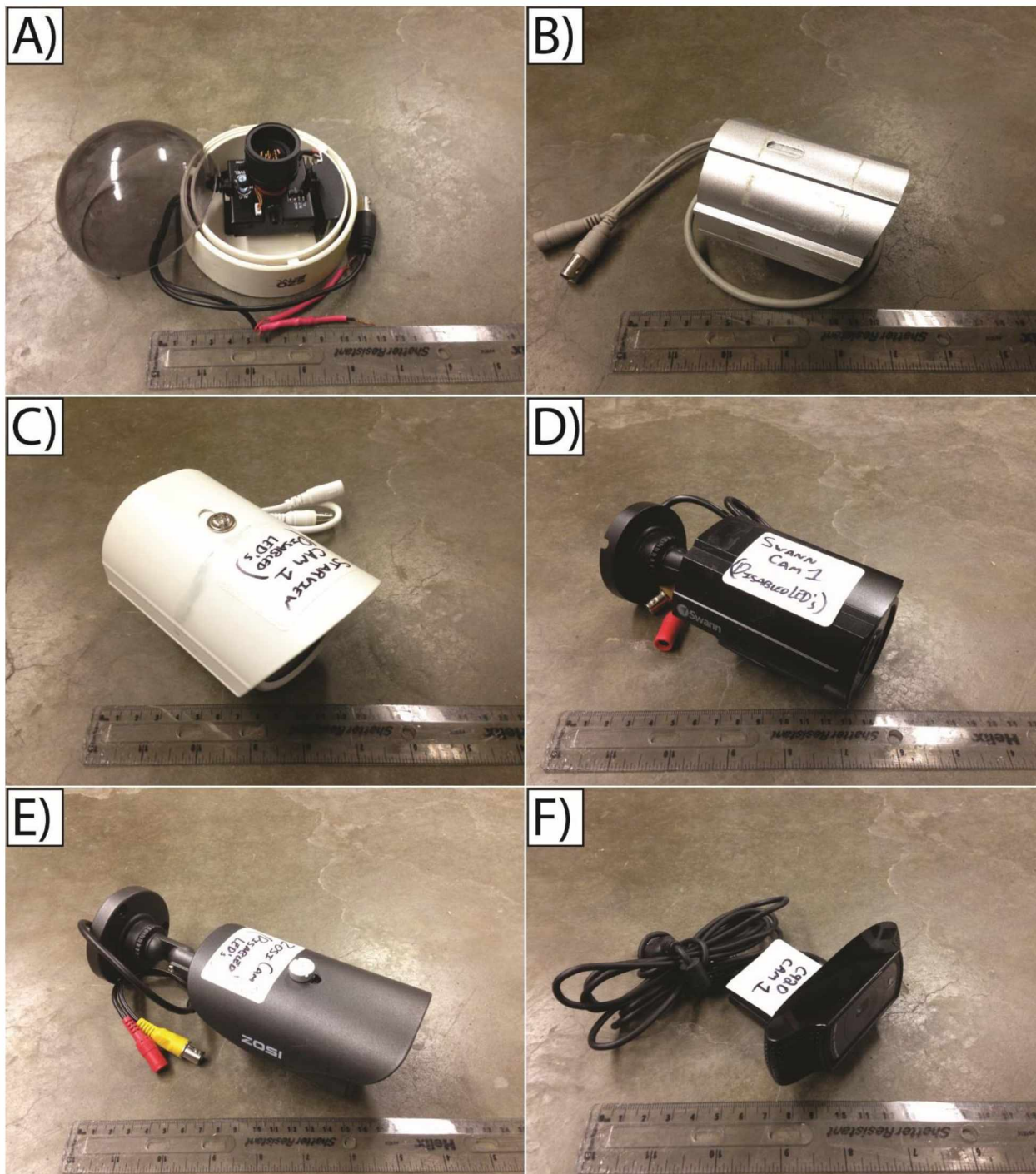


Figure 2.3. Images of six cameras acquired for the spectral response experiments. A) Dome cam, B) Silver cam, C) Starview cam, D) Swann cam, E) Zosi cam, and F) Logitech cam.

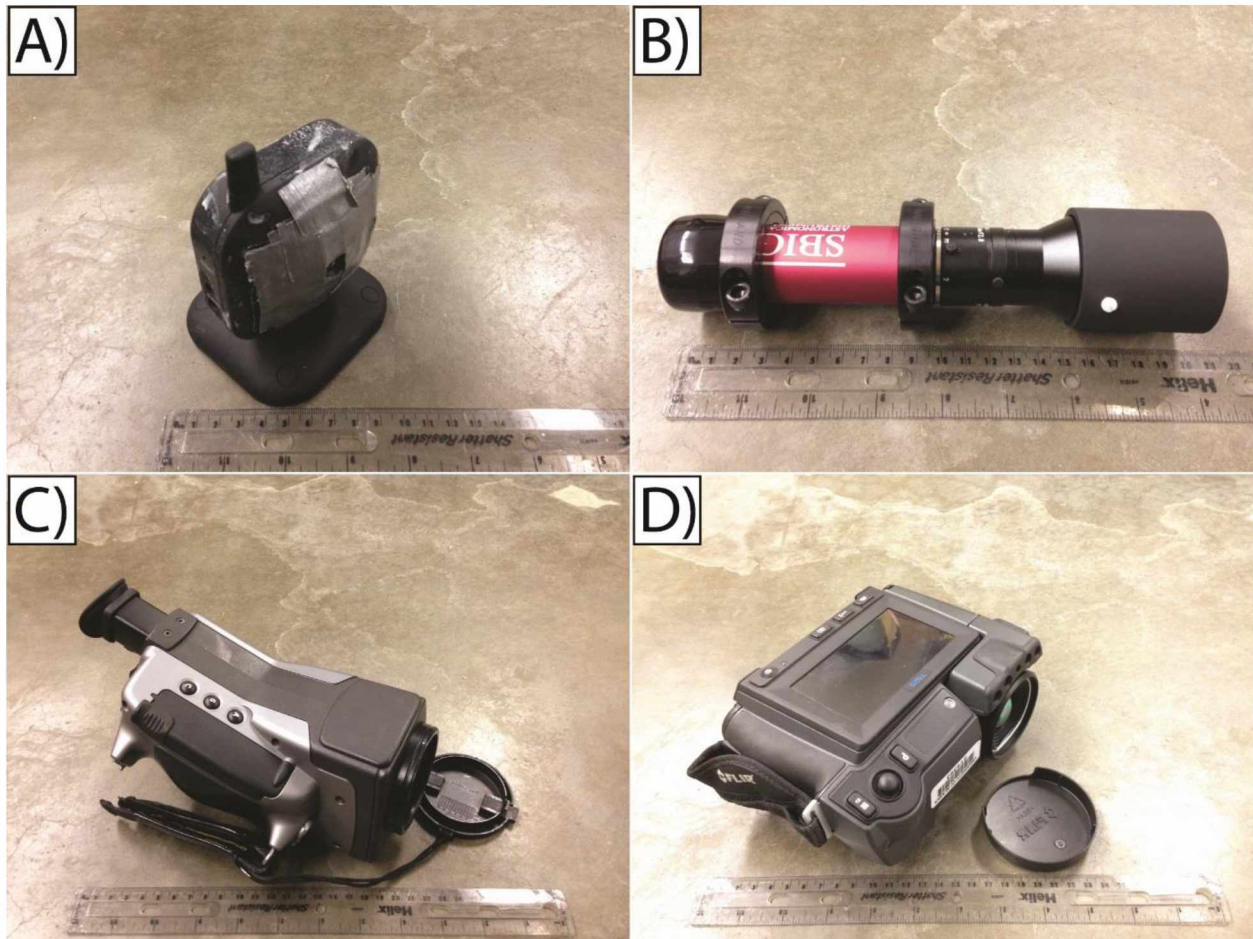


Figure 2.4. Images of four additional cameras used during this investigation. A) Ethernet cam, B) Star cam, C) FLIR S40, and D) FLIR T620.

2.2.1 Omega BB-4A Infrared Calibration Source

The Omega BB-4A infrared calibration source, referred to hereafter as the blackbody, is one of the most expensive items used in this investigation but is only necessary for the spectral response experiments, as it is used to simulate volcanic activity in a controlled laboratory environment. It is a ‘High Temperature Blackbody Infrared Calibration Source’ available from www.omega.com, model number BB-4A, and currently retails for \$3,775. Target temperatures range from 100°C to 982°C (212°F to 1,800°F), the temperature accuracy is $\pm 0.25\%$ or $\pm 1^\circ\text{C}$ ($\pm 1.8^\circ\text{F}$), whichever is greatest, the stability accuracy is $\pm 0.15^\circ\text{C}$ (0.27°F) or less, the cavity emissivity is 0.99, and the cavity opening is 22 mm (0.88”) in diameter. It also includes a built-in

RS-232 port as standard, allowing direct communication with the blackbody via a serial port cable and terminal program.

2.2.2 EverFocus ED350HQ Camera

The EverFocus ED350HQ camera (Figure 2.3A), referred to hereafter as the Dome cam, was formerly used as a security camera and can be purchased for approximately \$30. This particular camera was acquired from the UAF surplus warehouse for free, meaning little information is known about the inner workings of the camera. It is a Red, Green, Blue (RGB) color analog camera with a resolution of 520 television lines (TVL), but the specific wavelength this camera operates at is unknown. It has a 3.8 to 9.5 mm variable lens and a 1/3" charged-coupled device (CCD) sensor. The camera can be connected directly to a computer using a standard Radio Corporation of America (RCA) connector, an analog-to-digital convertor (ADC), and a software package to receive the signal. A commercially available package, 'ION Video 2 PC MKII' was purchased for \$34.99, which included the ADC and 'Cyberlink Power Director' software, allowing a live feed from the camera to be visualized. Images can be captured in several standard formats, of which the bitmap (.bmp) format was chosen, because it is an uncompressed file format and therefore no information will be lost during the acquisition process. Video can also be captured at a 30 FPS sampling rate and stored in a compressed .mpg format if required.

2.2.3 CM711 1/3" CCD CCTV Camera

The CM711 1/3" CCD closed-circuit television (CCTV) camera (Figure 2.3B), referred to hereafter as the Silver cam, is marketed as a 'weatherproof surveillance camera', and predominantly used in home security setups. It was purchased for \$49.99 and has a 1/3" CCD sensor, 650 TVL resolution, a 3.6 mm lens, 12 infrared LED's, and a night mode function. The specific wavelength this camera operates at is unknown. TVL is a resolution measurement for analog cameras, and 650 TVL is approximately equal to 720 by 480 pixel digital resolution, known more commonly as 480p. The LED's on this camera were disabled by disconnecting them from

the motherboard, together with the light sensor, which forced the camera to capture in a grayscale night mode. Images were captured using an ADC and accompanying software.

2.2.4 Starview 4SV558CW48 Camera

The Starview 4SV558CW48 camera (Figure 2.3C), referred to hereafter as the Starview cam, is marketed as a ‘weatherproof bullet surveillance CCTV security camera’, as it is predominantly used in home security setups. It was purchased for \$38.89 and has a 1/3” complementary metal-oxide semiconductor (CMOS) sensor, 1000 TVL resolution, a 3.6 mm lens, 48 infrared LED’s, and a night mode function. The specific wavelength this camera operates at is unknown. As with the previous camera, the LED’s were disabled, the camera forced to capture in grayscale night mode, and images were captured using an ADC and accompanying software.

2.2.5 Swann SWPRO-535 Camera

The Swann PRO-535 cam (Figure 2.3D), referred to hereafter as the Swann cam, is marketed as a ‘multi-purpose day/night security camera’, and predominantly used in home security setups. It was purchased for \$24.99 and has a CMOS sensor, 650 TVL resolution, a 3.6 mm lens, 36 infrared LED’s, and a night mode function. This camera operates in the NIR at 0.85 μ m and has a resolution approximately equal to 720 by 480 pixel digital resolution. Unfortunately, the camera’s gain settings are automatic and cannot be manually controlled, meaning the camera adjusts these settings slightly to produce the ‘optimal image’, which can impact the results slightly. As with the previous camera, the LED’s were disabled, the camera forced to capture in grayscale night mode and images were captured using an ADC and accompanying software.

2.2.6 Zosi ZSZR04BA Camera

The Zosi ZSZR04BA camera (Figure 2.3E), referred to hereafter as the Zosi cam, is marketed as an ‘indoor/outdoor day/night weatherproof vision security camera’, and predominantly used in home security setups. It was purchased for \$24.75 and has a 1/3” CMOS sensor, 1000 TVL resolution, a 3.6 mm lens, 42 infrared LED’s, and a night mode function. The specific wavelength this camera operates at is unknown, and unfortunately, the camera gain settings and back light compensation cannot be manually controlled. As with previous cameras, the LED’s were disabled, but the camera would not maintain a night mode status, so as the blackbody temperature increased and emitted more light, the camera would switch from its grayscale night mode to its color day mode. This was a frustrating problem, but indicative of a realistic issue that may arise with some specific camera purchases and for certain webcams where the chosen camera was purchased by another organization and not optimally chosen for analyzing change detection. Images were captured using an ADC convertor and accompanying software.

2.2.7 Logitech HD Pro C920 Camera

The Logitech HD Pro C920 camera (Figure 2.3F), referred to hereafter as the Logitech cam, is a high definition digital webcam normally used for computer based video calling and recording. It was purchased for \$64.91 and has a 1/4" CMOS sensor with 1920 by 1080 pixel resolution (1080p) for video recording at 30 FPS and up to 5168 by 2907 pixel (15 megapixels, MP) resolution for capturing images. This extremely high resolution means the Logitech cam has an image capture resolution approximately 49 times higher than the other analog cameras tested. The primary function of this camera means that almost all camera settings, such as exposure, focus, and gain are automatically adjusted by the camera to display the best possible image. Fortunately, connecting to the camera via the freely available ‘Logitech Webcam Software’ (LWS), allows these settings to be manually controlled, meaning exact image comparisons are possible by maintaining constant settings for all image acquisitions. For this investigation, all settings were kept constant except for exposure levels, which were sampled at all 15 available settings. Images captured using the LWS were only available in a compressed .jpg format.

2.2.8 Apexis J012 IP Camera

The Apexis J012 IP Camera (Figure 2.4A), referred to hereafter as the Ethernet cam, is an older style RGB color webcam that was purchased for \$42.50. It has a 1/4" CMOS sensor, a 3.6 mm lens, and is powered by a universal serial bus (USB) connection and connects to a computer's Ethernet port, which requires very specific configurations to function on a local Ethernet network, but does allow dedicated image capturing through its local web browser interface with a resolution of 640 by 480 pixels (480p). Unfortunately, the rate of capture was limited to approximately one image every ten seconds due to limitations with the accompanying software, which was insufficient for the needs of this investigation. Therefore, a third-party screen capture software package was installed to capture images at higher rates using the camera's live feed function. However, this does cause issues relating to pixel brightness distortion from color scaling when using the screen capture method compared to the raw data directly from the camera. Relative pixel brightness change was still possible, but absolute values would be difficult to determine.

2.2.9 SBIG ST-i Monochrome Astronomical Camera

The Santa Barbara Instrument Group (SBIG) ST-i Monochrome Astronomical Camera cam (Figure 2.4B), referred to hereafter as the Star cam, is a high resolution (648 by 486 pixels) monochrome camera with a 1/4" CCD sensor purchased from SBIG Astronomical Instruments for a total of \$944 (\$595 for the camera, and \$349 for the guiding kit zoom lens). Its primary function is to observe astronomical activity, but due to the exceptional sensitivity to low light levels required for this task, it was a useful addition to the equipment being used. The additionally purchased guiding kit allowed the connection of a 100 mm lens to the camera, which dramatically increased the viewing distance and narrowed the field of view. It can operate at several wavelengths depending upon the specified exposure and focus settings, and has a KAI-0340 imaging CCD. This camera can capture images at a 16-bit depth and can therefore identify 65,536 unique values, compared to 256 for conventional 8-bit cameras. It connects to and is powered by a USB port and uses dedicated software supplied upon purchase. Images are captured in raw format (.SBIG) and can be converted to a number of conventional image formats.

2.2.10 FLIR S40

The FLIR S40 (Figure 2.4C) is a TIR camera purchased from FLIR Systems, but is an older thermal camera. It was purchased in 2002 for approximately \$35,000 and can capture images at a moderate resolution (320 by 240 pixels). It has an uncooled microbolometer with a thermal sensitivity of $< 0.08^{\circ}\text{C}$ and operates in the TIR portion of the electromagnetic spectrum (7.5 - 13 μm). The camera can image scenes at up to 60 FPS and has a varying field of view (FOV) depending upon the attached lens. This investigation uses the standard lens giving a FOV of $24^{\circ} \times 18^{\circ}$, and a 2x zoom lens (\sim \$5,000 in 2002) reducing the field of view to $12^{\circ} \times 9^{\circ}$. The camera has three temperature ranges of -40°C to 120°C (-40°F to 248°F), 0°C to 500°C (32°F to 932°F), and 350°C to $1,500^{\circ}\text{C}$ (662°F to $2,732^{\circ}\text{F}$), with a temperature measurement accuracy of $\pm 2^{\circ}\text{C}$ ($\pm 3.6^{\circ}\text{F}$) or $\pm 2\%$, whichever is greatest. It stores captured images in .FFF format and video in .SEQ format, which can be opened using dedicated FLIR Systems software.

2.2.11 FLIR T620

The FLIR T620 (Figure 2.4D) is a higher resolution (640 by 480 pixels) TIR camera purchased from FLIR Systems for approximately \$20,000 in 2012. It has an uncooled microbolometer with a thermal sensitivity of $< 0.04^{\circ}\text{C}$ and operates in the TIR portion of the electromagnetic spectrum (7.5 - 13 μm). The camera can image scenes at up to 30 FPS and has a varying field of view depending upon the attached lens. For this investigation, a 25° lens was used resulting in a $25^{\circ} \times 19^{\circ}$ FOV. The camera has two temperature ranges of -40°C to 150°C (-40°F to 302°F) and 100°C to 650°C (212°F to $1,202^{\circ}\text{F}$) with a temperature measurement accuracy of $\pm 2^{\circ}\text{C}$ ($\pm 3.6^{\circ}\text{F}$) or $\pm 2\%$, whichever is greatest. It stores captured images in standard .jpg format and includes raw temperature data with the file, which can be read using specialized FLIR Systems software, and captures video in .mp4 format. The FLIR T620 can also simultaneously capture visible images using a built in 5 MP camera.

Table 2.1. Summary table of all camera information and specifications used in this investigation

	<u>Dome cam</u>	<u>Silver cam</u>	<u>Starview cam</u>	<u>Swann cam</u>	<u>Zosi cam</u>	<u>Logitech cam</u>	<u>Ethernet cam</u>	<u>Star cam</u>	<u>FLIR S40</u>	<u>FLIR T620</u>
Make	EverFocus		Starview	Swann	Zosi	Logitech	Apexis	SBIG	FLIR	FLIR
Model	ED350HQ	CM711	4SV558 CW48	SWPRO- 535	ZSZR0 4BA	C920	J012	ST-i mono- chrome	S40	T620
Cost	\$30.00	49.99	\$38.89	\$24.99	\$24.75	\$64.91	\$42.50	944.00	\$35,000 in 2002	\$20,000 in 2012
Resolution	520 TVL	650 TVL	1,000 TVL	650 TVL	1,000 TVL	5,168 by 2,907 pixels	640 by 480 pixels	648 by 486 pixels	320 by 240 pixels	640 by 480 pixels
Wavelength	Visible	NIR	NIR	NIR	NIR	Visible	Visible	NIR	TIR	TIR
Data type	Analog	Analog	Analog	Analog	Analog	Digital	Digital	Digital	Digital	Digital
Image bit depth	8-bit	8-bit	8-bit	8-bit	8-bit	8-bit	8-bit	16-bit	8-bit	8-bit
Sensor	1/3" CCD	1/3" CMOS	1/3" CMOS	1/3" CMOS	1/3" CMOS	1/4" CMOS	1/4" CMOS	1/3" CCD	Uncooled micro- bolometer	Uncooled micro- bolometer
Lens	3.8 - 9.5 mm	3.6 mm	3.6 mm	3.6 mm	3.6 mm	N/A	3.6 mm	100 mm	24° and 12°	25°
LED's	N/A	12	48	36	42	N/A	10	N/A	N/A	N/A
Power connection	DC battery	DC and AC adaptor	AC adaptor	AC adaptor	AC adaptor	USB	USB	USB	DC and AC adaptor	DC and AC adaptor
Image format	.bmp	.bmp	.bmp	.bmp	.bmp	.jpg	.jpg	.SBIG	.jpg	.FFF

2.3 Preliminary Laboratory Testing

2.3.1 Blackbody Temperature Stability Tests

To accurately image the blackbody heat source without an analyst needing to be present and constantly observing temperature readings, a series of tests were conducted to determine the length of time taken for the blackbody to stabilize to a specified temperature. Temperatures ranging from 100°C to 975°C were analyzed in 25°C increments by utilizing the blackbody's serial communication port and 'pinging' the instrument for accurate temperature readings every 10 seconds. Each increment was set to stabilize for 30 minutes and the final cool down phase back to 100°C was allotted three hours, both of which were sufficient lengths of time according to manufacturer specifications and personal experience. The temperature was considered stable once the temperature reading from the blackbody matched the specified temperature and did not change again, as the manufacturer specifications state a stability accuracy is $\pm 0.15^{\circ}\text{C}$ (0.27°F) or less. This heating and cooling cycle was conducted three times with the maximum value from all tests being selected. An additional time buffer of 10% was added to allow for any additional stability delays.

As expected, stability timings vary according to the specified temperature, with lower temperatures stabilizing faster compared to higher temperatures (Table 2.2). The shortest stability time, including the 10% buffer, was ~18 minutes, compared to the longest time of ~28 minutes, and the cool down phase took ~3 hours, resulting in a total heating and cooling cycle time of ~16 hours for each camera at each specified distance.

Table 2.2. Recorded blackbody temperature stability timings. The length of time for the blackbody to stabilize at a specified temperature was recorded for three separate cycles. The maximum timing was combined with a 10% buffer to determine the total stability time.

<u>Set Temp</u> (°)	<u>Time #1</u>	<u>Time #2</u>	<u>Time #3</u>	<u>Max</u>	<u>Buffer</u> (10%)	<u>Total</u> <u>Time</u>
125	00:17:32	00:16:32	00:17:01	00:17:32	00:01:45	00:19:17
150	00:16:03	00:15:59	00:15:31	00:16:03	00:01:36	00:17:39
175	00:17:53	00:18:13	00:17:03	00:18:13	00:01:49	00:19:42
200	00:17:53	00:17:10	00:17:20	00:17:53	00:01:47	00:19:40
225	00:16:23	00:17:31	00:17:13	00:17:31	00:01:45	00:18:08
250	00:17:31	00:17:51	00:18:13	00:18:13	00:01:49	00:19:20
275	00:20:04	00:19:52	00:18:33	00:20:04	00:02:00	00:22:04
300	00:17:03	00:18:33	00:19:44	00:19:44	00:01:58	00:19:01
325	00:18:53	00:18:53	00:18:44	00:18:53	00:01:53	00:20:46
350	00:19:24	00:18:23	00:18:03	00:19:24	00:01:56	00:21:20
375	00:18:32	00:18:43	00:17:21	00:18:43	00:01:52	00:20:24
400	00:17:43	00:18:30	00:17:31	00:18:30	00:01:51	00:19:34
425	00:17:52	00:18:12	00:18:44	00:18:44	00:01:52	00:19:44
450	00:19:53	00:19:54	00:20:32	00:20:32	00:02:03	00:21:56
475	00:18:43	00:19:54	00:20:43	00:20:43	00:02:04	00:20:47
500	00:19:33	00:19:54	00:21:04	00:21:04	00:02:06	00:21:39
525	00:20:45	00:20:45	00:19:25	00:20:45	00:02:04	00:22:49
550	00:20:35	00:19:04	00:19:23	00:20:35	00:02:04	00:22:39
575	00:20:24	00:20:04	00:21:13	00:21:13	00:02:07	00:22:31
600	00:20:44	00:23:37	00:24:19	00:24:19	00:02:26	00:23:10
625	00:19:55	00:20:37	00:20:04	00:20:37	00:02:04	00:21:59
650	00:21:14	00:19:53	00:20:04	00:21:14	00:02:07	00:23:21
675	00:19:14	00:20:42	00:20:56	00:20:56	00:02:06	00:21:20
700	00:23:05	00:22:06	00:19:54	00:23:05	00:02:19	00:25:24
725	00:20:55	00:21:04	00:23:26	00:23:26	00:02:21	00:23:16
750	00:20:53	00:20:34	00:21:06	00:21:06	00:02:07	00:23:00
775	00:25:28	00:19:52	00:19:22	00:25:28	00:02:33	00:28:01
800	00:22:35	00:18:54	00:21:45	00:22:35	00:02:15	00:24:50
825	00:21:34	00:23:16	00:24:57	00:24:57	00:02:30	00:24:04
850	00:21:36	00:23:27	00:24:06	00:24:06	00:02:25	00:24:01
875	00:20:44	00:22:35	00:24:16	00:24:16	00:02:26	00:23:10
900	00:20:54	00:20:25	00:22:15	00:22:15	00:02:13	00:23:07
925	00:20:43	00:22:25	00:20:14	00:22:25	00:02:14	00:22:57
950	00:24:17	00:23:56	00:21:05	00:24:17	00:02:26	00:26:43
975	00:23:26	00:21:45	00:22:26	00:23:26	00:02:21	00:25:47
100	02:47:14	02:46:27	02:46:15	02:47:14	00:16:43	03:03:57
					Total	15:57:10

2.3.2 Image Capture Sequence

As most of the tested cameras were analog, the likelihood of random noise resulting from signal interference was high, and could skew the results. Therefore, 30 images would be captured for each temperature increment and the results averaged. This number was chosen because it is generally accepted that using at least 30 data points is sufficient to assume a normal data distribution (Davis and Sampson, 1986), meaning any image noise would be statistically insignificant.

This test was also used to determine if there was a significant difference between capturing all 30 images at once, and capturing 10 images at once for three cycles. The difference between the mean pixel value of all images captured 10 at a time was found to be statistically similar to the mean pixel value of images captured 30 at a time. The average difference in mean pixel value between the two methods was 0.54%, with a maximum difference of 1.33% and a minimum difference of 0.13%. These differences are well within the accepted margin of error and are likely produced by variations in random image noise, as opposed to differences in the capturing method. Therefore, all spectral response tests were conducted by capturing 30 images at once, resulting in 1,080 images per cycle for each distance, and a total of 5,400 images per camera.

2.3.3 Camera Testing

As the overall aim of this investigation is to detect increases in volcanic activity as soon as possible, the maximum pixel brightness value within an image will be the most useful, as it will be the first parameter to increase, once a change in activity occurs. The mean pixel value across the full image and standard deviation parameters are also somewhat useful at identifying increases in brightness due to a change in activity, as they will also react to pixel brightness increases. All three results will be shown for each camera at all distances.

To determine a camera's spectral range, the temperature at which the camera initially detects activity, and the temperature at which the camera saturates need to be determined. Initial detection refers to the blackbody temperature where the recorded pixel brightness has become elevated

above background levels, and camera saturation refers to a recorded pixel value of 255, which is the maximum the camera can record, as all these cameras contain 8-bit sensors (2^8 unique possible values ranging from 0 to 255).

As the distance between each camera and the blackbody target increases, pixel integrated brightness will factor into the recorded values, as the blackbody source will represent only a portion of a camera image pixel, and therefore the results will be less sensitive to small changes in temperature. At close distances, the blackbody cavity (22 mm in diameter) is still large enough that the hot target still fills at least one pixel. However, as the distance between the camera and blackbody increases, this hot cavity becomes smaller than an individual pixel (sub-pixel), and darker background values are included. This means the recorded pixel value will be an average of the bright target and dark background values, which will skew the results and detrimentally affect the cameras detection abilities at greater distances.

All acquired images will be cropped to a specific region of interest (ROI). During setup, the camera is specifically lined up with the blackbody cavity in the center of the image. A 100 by 100 pixel ROI crop is performed around the center of the image so only results from the cavity itself are included. For example, as analog cameras age, individual pixel sensors degrade and false 'dead' pixel values are sometimes recorded. By cropping the images to a small ROI, the possibility of this occurring is greatly reduced. The small image crop also reduces computational stress during image processing due to the greatly reduced number of pixels being analyzed.

All color images will be converted to grayscale images so only one channel of data per image exists. Color pixels have three unique digital numbers from the red, green, and blue channel, forming an RGB pixel. The accepted method for a color to grayscale conversion is using a weighted sum for each channel calculated from gamma compressed primary intensities of the image with values of 0.2989 for the Red channel, 0.5870 for the green channel, and 0.1140 for the blue channel used (Grundland and Dodgson, 2005).

2.3.3.1 Dome Cam

The Dome cam test results (Figure 2.5) indicate that this camera operates at visible wavelengths and can observe temperatures similar to humans, with maximum pixel brightness results summarized in Table 2.3. Although the saturation levels for the Dome cam at the first four distances were equal, Figure 2.5A shows that the camera would likely saturate at several different temperatures for each distance, if either more data points were included in the testing or a temperature increment less than 25°C was used. At lower temperatures, constant pixel brightness values of approximately 30 are recorded, resulting from random noise generated by the analog signal of the camera.

Table 2.3. Dome cam spectral response experiment results. Initial detection and saturation temperatures for maximum pixel brightness are shown at each distance using 25°C increments.

	1 m	2 m	4 m	6 m	8 m
Initial detection (°C)	525	550	575	575	600
Saturation (°C)	725	725	725	725	750

2.3.3.2 Silver Cam

The Silver cam test results (Figure 2.6) indicate that this camera can observe significantly lower temperatures than the Dome cam, signifying it operates at longer NIR wavelengths, with maximum pixel brightness results summarized in Table 2.4. Therefore, the Silver cam can detect the simulated volcanic activity sooner, but it saturates at lower temperatures than the Dome cam and cannot be used to detect changes at higher temperatures. A data anomaly occurred just prior to the maximum pixel value (Figure 2.6A) becoming saturated at the six-meter distance, which was likely caused by the camera's automatic gain or gamma control settings erroneously altering the pixel brightness values as saturation levels were approached. By interpolating between the other distances, an expected saturation temperature for the six-meter distance should be approximately 575°C. At lower temperatures, constant pixel brightness values of approximately 38 are recorded, resulting from random noise generated by the analog signal of the camera. In

Figure 2.6C, the standard deviation of the one-meter distance data starts to decrease at approximately 850°C resulting from significant brightness levels from the blackbody at this high temperature, and severe pixel bleed across the entire image. This elevates almost all the pixel brightness values and actually causes a decrease in the standard deviation because the majority of pixels are now completely saturated. This trend continues as the temperature increases to the maximum, and would end with a very low standard value if higher temperatures were imaged, as the entire image would be completely saturated.

Table 2.4. Silver cam spectral response experiment results. Initial detection and saturation temperatures for maximum pixel brightness are shown at each distance using 25°C increments.

	1 m	2 m	4 m	6 m	8 m
Initial detection (°C)	400	400	425	450	450
Saturation (°C)	500	500	525	625	625

2.3.4 Atmospheric Conditions

At the beginning of each image capture cycle the atmospheric room temperature and relative humidity of the laboratory were recorded using a handheld hygrometer (Table 2.5). As these experiments were conducted in a climate controlled room during the dry arctic winter in Fairbanks, Alaska, both the temperature and humidity readings remained relatively constant. Temperatures barely changed, ranging from 20.1°C to 20.4°C (68.18°F to 68.72°F), and humidity ranged between 15% and 19%. Given the extremely small variations during the image acquisition period, any error due to changing atmospheric conditions can be ruled out.

Table 2.5. Atmospheric temperature (°C) and relative humidity (%) as recorded throughout the preliminary spectral response experiments.

	1 m	2 m	4 m	6 m	8 m
Silver cam	20.4°C	20.4°C	20.3°C	20.4°C	20.3°C
	19%	17%	17%	15%	15%
Dome cam	20.3°C	20.1°C	20.2°C	20.2°C	20.3°C
	15%	15%	15%	15%	15%

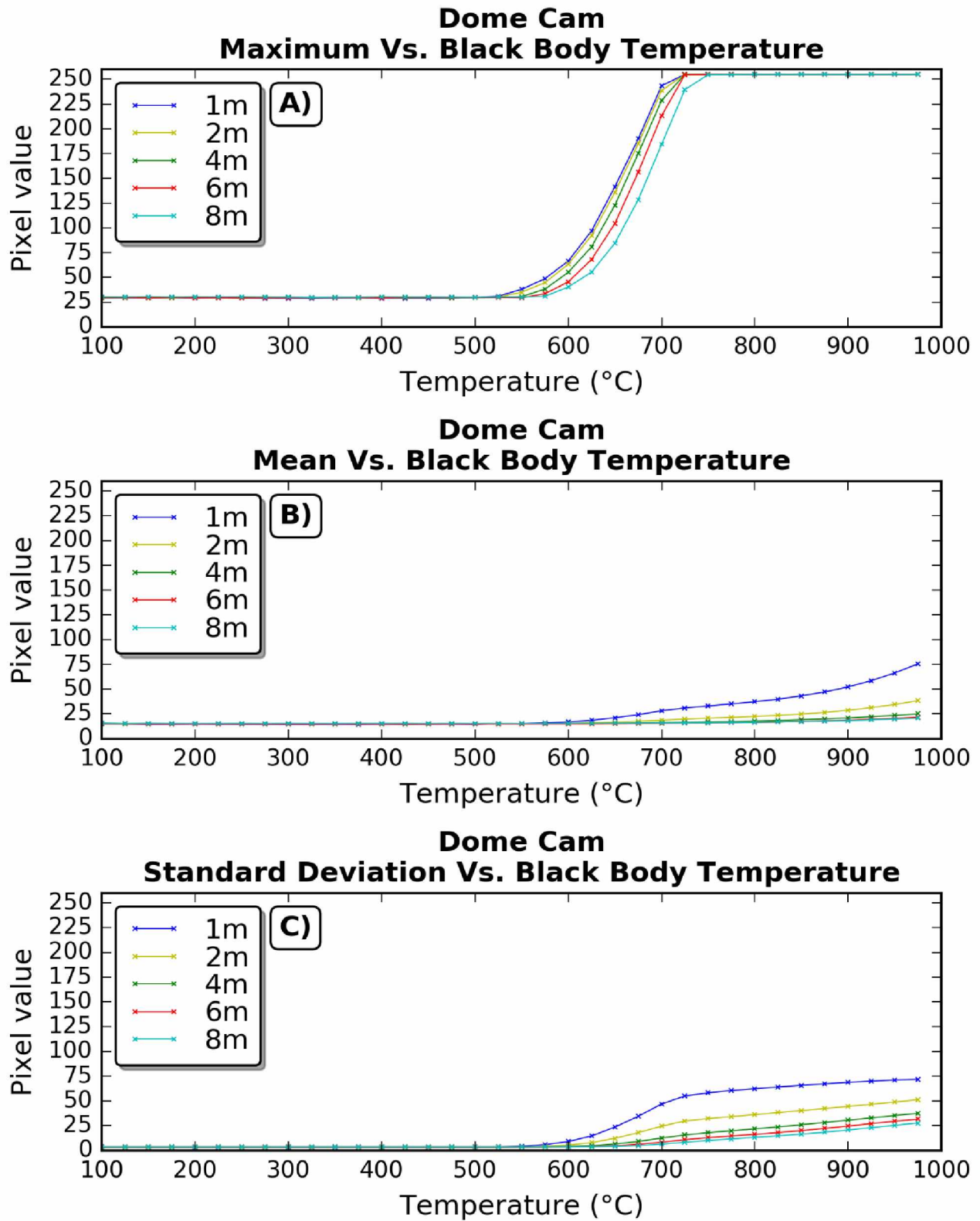


Figure 2.5. Preliminary results from the Dome cam spectral response experiment. All five distances are plotted to demonstrate the effect of pixel integrated brightness. Three different image parameters are shown; A) maximum, B) mean, and C) standard deviation.

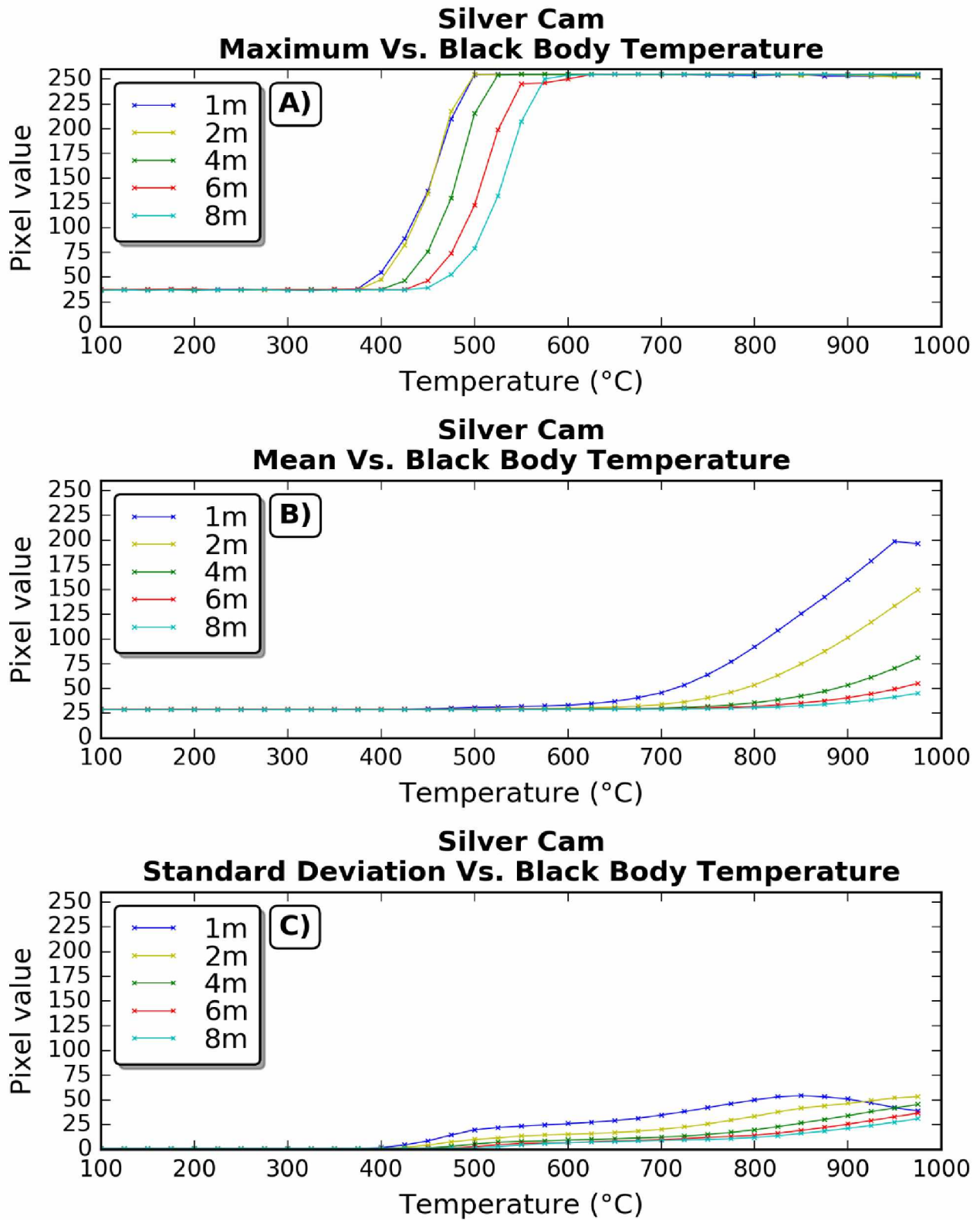


Figure 2.6. Preliminary results from the Silver cam spectral response experiment. All five distances are plotted to demonstrate the effect of pixel integrated brightness. Three different image parameters are shown; A) maximum, B) mean, and C) standard deviation.

2.3.5 Conclusions

These preliminary tests have shown that inexpensive cameras can be beneficial to detect changes in pixel brightness due to increased temperatures. Major differences between the two cameras were identified, resulting from their differing operating wavelengths. The Dome cam operates at visible wavelengths, similar to the wavelengths humans can see, whereas the Silver cam operates at longer wavelengths in the NIR. Therefore, the Silver cam can detect increases in brightness at much lower temperatures, but the Dome cam will observe changes and saturate at higher temperatures. Figure 2.7 shows a side by side comparison of these cameras at various temperatures.

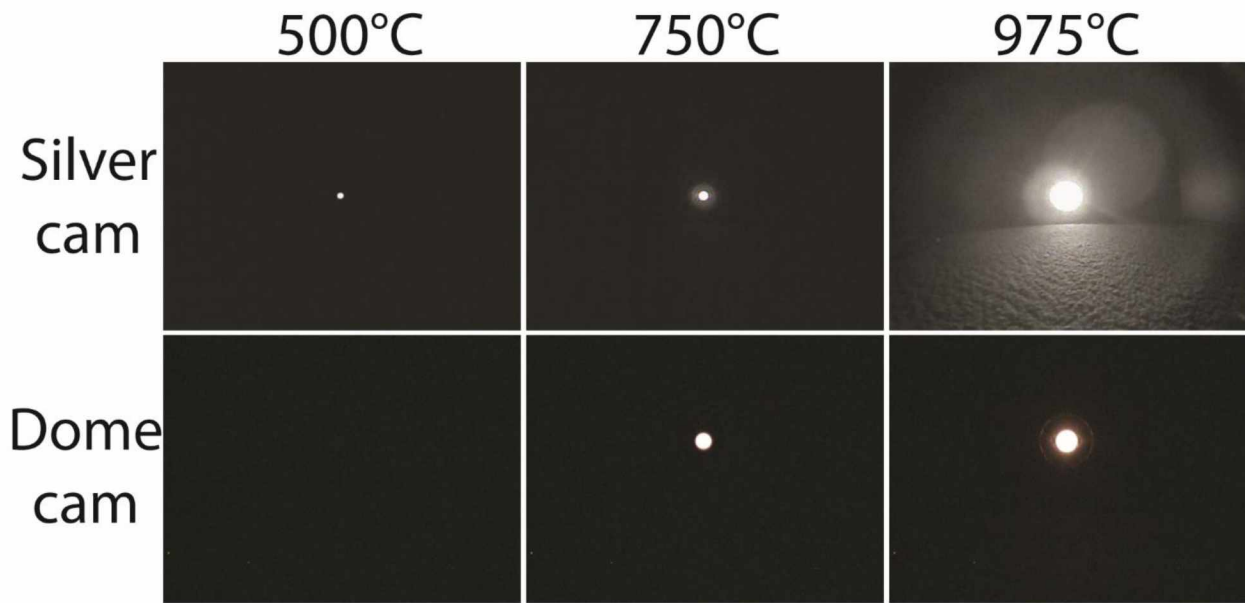


Figure 2.7. Side by side image comparison from the Dome and Silver cameras at three different temperatures during spectral response testing. The Silver cam saturated at 500°C whereas the Dome cam saturated at 750°C. Earlier detection by the Silver cam is clearly visible, but severe pixel bleed occurs at higher temperatures once it has saturated.

2.4 Primary Laboratory Testing

After analyzing the results from the preliminary spectral response experiments, additional extensive spectral response experiments were conducted on both cameras from Section 2.3 as well as the newly purchased Starview, Swann, Zosi, and Logitech cameras. Specific improvements were applied to the data collection method resulting from the identification of sources of error during the preliminary testing. For example, 1) images were captured at intervals of 5°C instead of 25°C to increase the number of data points and identify initial detection and saturation temperatures at a higher level of accuracy, 2) improvements were made to the automatic collection code to improve efficiency and reduce overall capturing time, 3) almost all sources of illumination were powered down instead of being covered to further reduce the chance of light pollution, and 4) cameras were examined for ‘dead pixels’ to ensure they would not affect the results.

Several ‘dead pixels’ were found in the Dome cam imagery, likely resulting from the age and prior treatment of the camera which damaged the camera's image sensor. Fortunately, almost all erroneous pixels were outside the ROI specified earlier, which was designed specifically for this type of situation, except one. The pixel was relatively near to the center of the image and would possibly be included in the ROI. Therefore, the Dome cam's lens was angled slightly down, so the blackbody cavity was slightly higher in the image and far enough away from the erroneous pixel that it wouldn't affect the results.

As the temperature increment for these spectral response experiments was reduced from 25°C to 5°C, new temperature stability tests were conducted to determine how long the blackbody took to stabilize at each temperature. The data collection methods are identical to those described in Section 2.3.1, except the maximum tested temperature increased to 980°C and the tests were conducted seven times for improved accuracy. Once again, stability timings varied according to the specified temperatures, with lower values stabilizing faster compared to higher ones. The shortest time, including the 10% buffer, was ~4.5 minutes, the longest time was ~12.75 minutes, the cool down time was approximately three hours, combined for a total cycle capture time of approximately 22 hours and 30 minutes for each camera at each specified distance.

2.4.1 Dome Cam

The Dome cam test results (Figure 2.8A) show improved accuracy compared to the preliminary tests and the maximum pixel brightness results are summarized in Table 2.6. The initial detection temperature of the Dome cam was 600°C during the preliminary testing, but this has reduced to 575°C during these tests. This is likely a result of the previously described ‘dead pixel’ being included in the preliminary analysis and artificially increasing the background pixel values, which reduced the initial detection sensitivity. The saturation levels for the first four distances are different from the preliminary tests due to the smaller temperature increments measured. However, the eight-meter data shows erroneous results close to saturation temperatures, as experienced with the Silver cam during preliminary testing, likely the result of gain or gamma settings being automatically adjusted by the camera. By interpolating from the other distances, an approximate expected saturation temperature of 740°C is determined, which is supported by the saturation temperature of 750°C during preliminary testing with larger temperature increments. Issues with this camera are likely caused by its age and deterioration of the camera’s image sensor.

2.4.2 Silver Cam

The Silver cam tests results (Figure 2.8B) show improved accuracy compared to the preliminary results with maximum pixel brightness results summarized in Table 2.6. As expected, the Silver cam captures activity at cooler blackbody temperatures, ranging between 130°C and 155°C lower than the Dome cam for all distances. This is very useful for early detection of volcanic activity, but the camera saturates at temperatures 100°C and 230°C cooler, which is a limitation of NIR cameras. The two-meter data is abnormal, as the spectral response curve begins at an expected blackbody temperature, but increases at a faster rate than the one-meter data, and saturates sooner. The differences between the one and two-meter datasets are expected to be very similar (as shown in the preliminary tests) because at those distances, the blackbody cavity is still larger than an individual pixel but the lower temperature saturation of the two-meter data was still unexpected. Possible explanations could be automatic changes to the cameras internal settings, or light

reflecting off a surface in the laboratory and amplifying the pixel brightness values as the temperature increased.

Abnormal behavior of the saturation levels at greater distances also occurs, which is likely caused by internal camera calibrations attempting to prevent the captured image from becoming saturated. A small glitch in the eight-meter data is also apparent at 520°C, which again, is likely the result of altered internal settings as the ‘before’ and ‘after’ gradients of the curve match almost perfectly, indicating something was changed and the remaining data points were offset. Issues with this camera are likely caused by its age and deterioration of the camera’s image sensor.

2.4.3 Starview Cam

The Starview cam test results (Figure 2.8C) show similar behavior to the Silver cam with comparable initial detection and saturation blackbody temperatures, summarized in Table 2.6. This was expected as both these cameras are designed for outdoor nighttime home security, and therefore operate at similar wavelengths. However, automatic setting adjustment issues still persisted with the six-meter data, resulting in an unexpected saturation temperature. By interpolating from the results of other distances, the six-meter saturation temperature should be approximately 560°C.

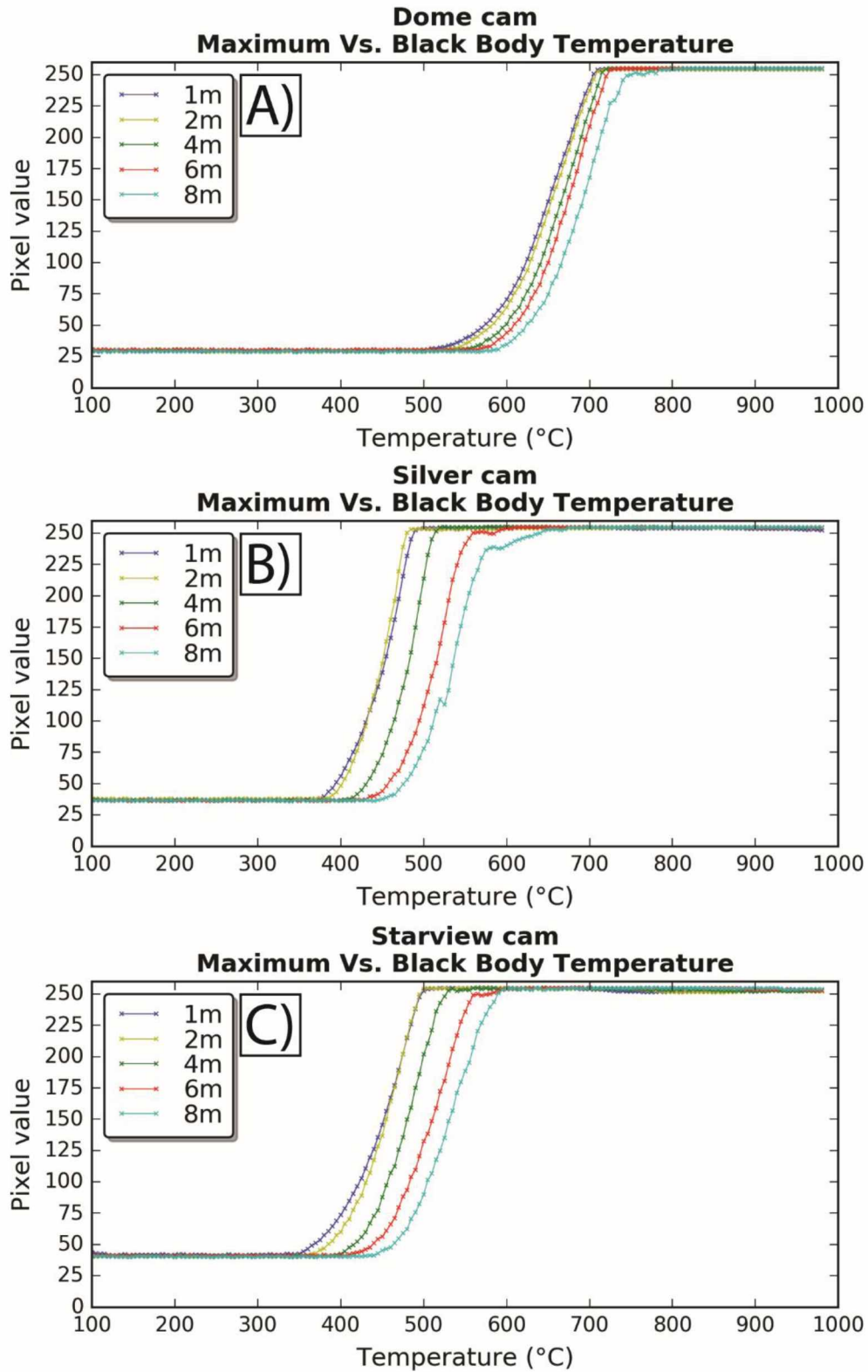


Figure 2.8. Spectral response experiment results at one, two, four, six, and eight-meter distances showing the maximum pixel brightness for A) Dome cam, B) Silver cam, and C) Starview cam.

2.4.4 Swann Cam

The Swann cam test results (Figure 2.9A) show the camera behaves as a good ‘all-rounder’ with initial blackbody detection temperatures slightly above the Silver cam, and saturation levels slightly lower than the Dome cam, as shown in Table 2.6. Therefore, the camera will not be as sensitive to small changes in temperature as a result of its ability to classify signals across a larger range of temperatures. Although the spectral response curves for this camera appear close to expected, the shape and gradient is interesting. The one and two-meter distance results behave almost as expected, but the curves for the remaining distances behave very unexpectedly. As the temperature increase from one interval to the next is linear, a linear increase in pixel brightness would also be expected for the majority of the response curve. However, the four, six, and eight-meter data are quite erratic and are not at all close to behaving linearly. Furthermore, data from all distances appears to alter at a maximum pixel brightness value of approximately 50. All the data show steeper and steeper gradients as the blackbody temperature increases and indicate there is a change in an internal calibration setting around this particular pixel brightness. It should also be noted the background pixel brightness values for this camera, approximately 17, are relatively low compared to all other cameras, which increases the potential to capture activity at lower temperatures that aren’t overcome by the noise.

2.4.5 Zosi Cam

The Zosi cam test results (Figure 2.9B) show the camera has similar initial detection temperature capabilities to the Silver and Starview cams, but it doesn’t saturate as early with maximum recordable saturation temperatures located between the Starview and Swann cam, summarized in Table 2.6. As with several other cameras, automatic internal calibration settings adjustments appear to manipulate the spectral response curves slightly, with glitches occurring in the six and eight-meter distance data. Again, the one and two-meter distance response curves are almost identical, except the two-meter data saturates at a slightly lower blackbody temperature than the one-meter data, likely the result of an automatic adjustment of the internal calibration settings. An abnormality in the two-meter data can be seen when examining the background

brightness values. The pixel brightness is elevated by a digital number of approximately five, compared to all other distances, which is likely the result of improper covering of a computer LED in the laboratory during the capturing process. If a 'dead pixel' were responsible for this elevated data value, it would be apparent at all distances.

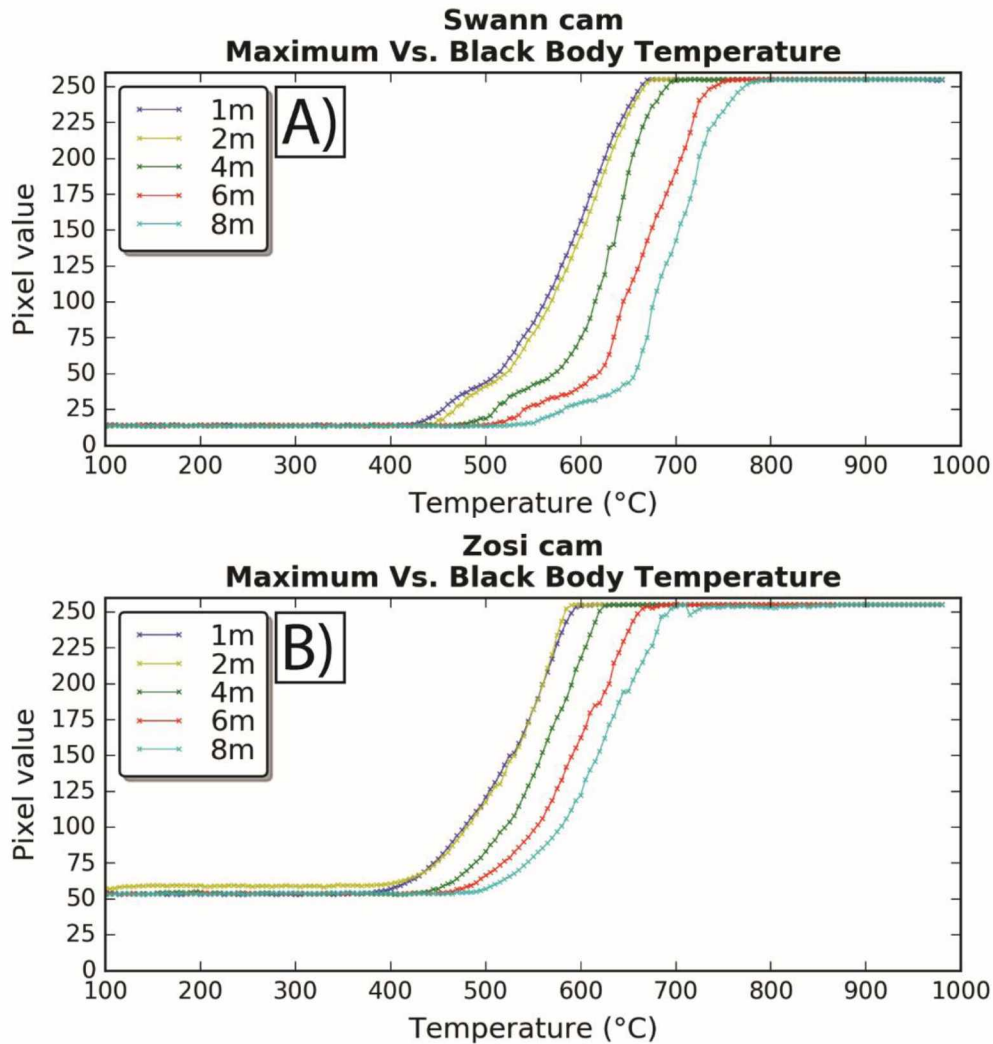


Figure 2.9. Spectral response experiment results at one, two, four, six, and eight-meter distances showing the maximum pixel brightness for A) Swann cam and B) Zosi cam.

2.4.6 Logitech Cam

The Logitech cam test results (Figure 2.10) show the camera responds very differently at various controlled exposure settings, as summarized in Table 2.6. These particular exposure settings represent the longest, shortest, and mid-range settings available. This camera is digital and has a significantly higher resolution (~49 times) than any of the analog cameras tested. As such, the differences between the recorded pixel brightness values at each distance are significantly less, because individual pixels in the Logitech cam images are still smaller than the blackbody cavity diameter at greater distances. The test results show this camera operates at visible wavelengths, with initial detection occurring at blackbody temperatures similar to the Dome cam. At several distances and exposure settings, this camera remains unsaturated at higher temperatures, which could be useful for hotter eruption temperatures. It should be noted that the only adjusted setting was exposure, and further adjustments to gain, brightness, contrast, color intensity, and white balance are possible to further modify the capture sequence to suit individual circumstances.

Table 2.6. Summary of primary spectral response experiment results from all cameras. Initial detection and saturation temperatures for maximum pixel brightness are shown at each distance.

		1 m (°C)	2 m (°C)	4 m (°C)	6 m (°C)	8 m (°C)
Dome cam	Initial detection	510	530	555	565	575
	Saturation	710	715	720	725	785
Silver cam	Initial detection	370	375	410	435	445
	Saturation	490	485	520	610	685
Starview cam	Initial detection	355	370	395	410	445
	Saturation	505	505	530	595	600
Swann cam	Initial detection	425	450	475	500	525
	Saturation	670	675	700	760	790
Zosi cam	Initial detection	385	390	440	455	470
	Saturation	605	595	630	690	710
Logitech cam Exp01	Initial detection	740	755	765	775	780
	Saturation	N/A	N/A	N/A	N/A	N/A
Logitech cam Exp08	Initial detection	625	630	660	685	705
	Saturation	930	935	945	970	N/A
Logitech cam Exp15	Initial detection	525	535	565	580	600
	Saturation	810	815	825	850	890

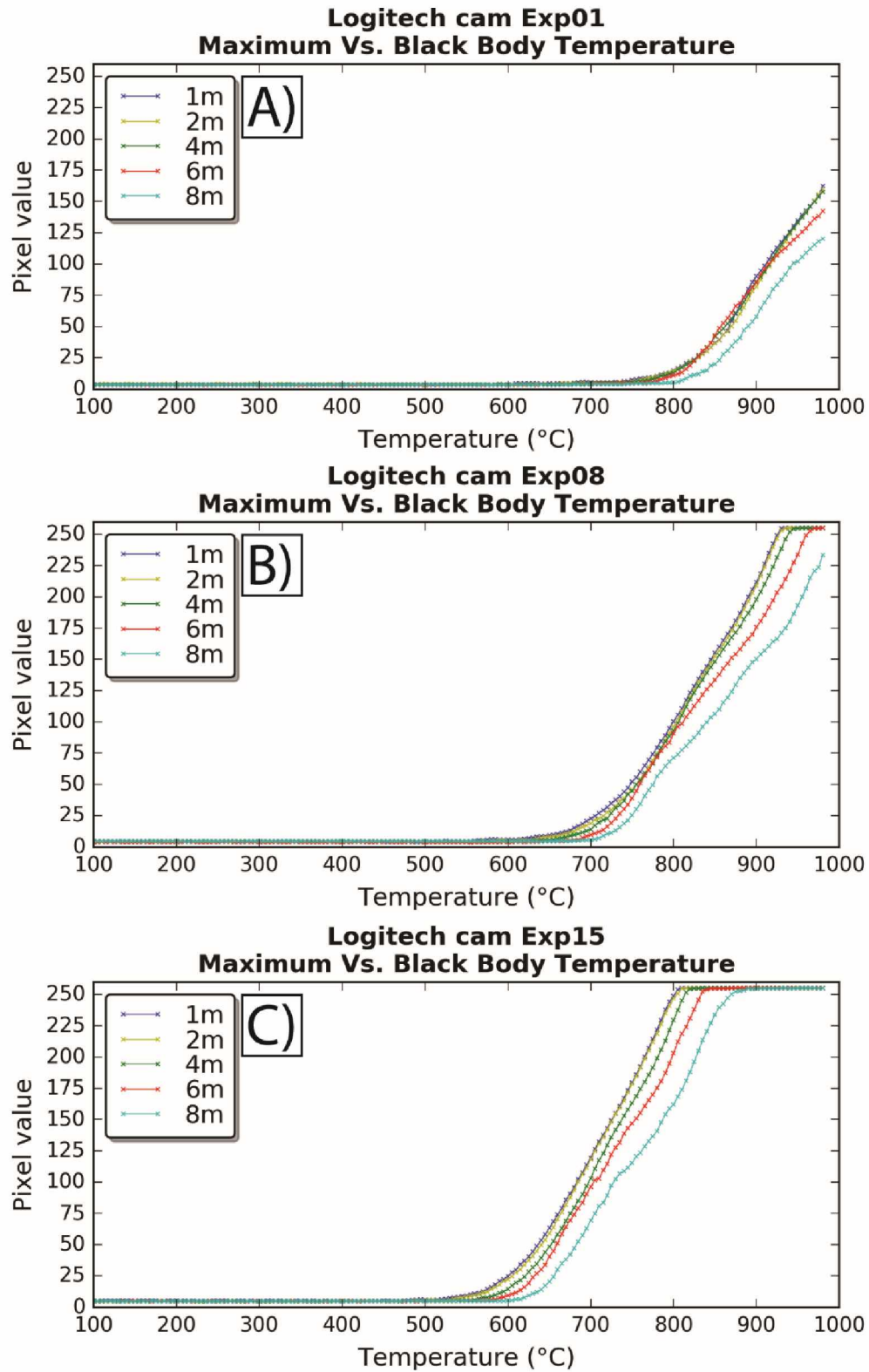


Figure 2.10. Spectral response experiment results at one, two, four, six, and eight-meter distances showing the maximum pixel brightness for the Logitech cam at different exposure settings. A) Exposure 01, B) Exposure 08, and C) Exposure 15.

2.4.7 Conclusions

Overall, all cameras tested during the spectral response experiments performed as expected (Figure 2.11). All NIR cameras (Silver, Starview, Swann, and Zosi), with a primary function of nighttime home security, could detect increases in pixel brightness at lower temperatures than the visible cameras, but also saturated at far lower temperatures. On the contrary, the visible cameras (Dome and Logitech) were only able to detect increases in pixel brightness at higher temperatures, but took significantly longer to saturate, and even remained unsaturated in the case of some Logitech cam exposure settings.

For a field deployment campaign, using a combination of several cameras would be ideal to take advantage of the positive aspects of each device. Depending upon availability of resources, such as power, it is recommended to use two of these cameras coincidentally to image volcanic activity. The Starview cam could be used for initial activity detection, as it has the coolest detection temperatures, and the Logitech cam could be used to image activity at hotter temperatures when other cameras are already saturated. Unfortunately, due to financial restraints, it was not possible to test these cameras during a field campaign deployment.

The distances between the cameras and blackbody source tested in these experiments were relatively small compared to an actual field campaign deployment, but the response of each camera is still applicable to a real-life situation. Whether the distance is large or small, the Starview cam will detect activity before the other cameras and the Logitech cam will likely stay unsaturated at higher temperatures. The distance from camera to source only impacts the area on the volcanic edifice represented by one pixel in the camera. At 1 m (3 ft), the pixel would be sub-mm in size, while at 1,000 m (3,280 ft), it would represent 10's of m. Although beyond the scope of this investigation, further spectral response experiments are encouraged for any future investigations to determine the camera responses at significantly greater distances. For example, testing these cameras in a large warehouse at several hundred-meter distances would be ideal to add more data to these response curves.

Although only relatively short distances were tested here compared to real life scenarios, where cameras would likely be located several kilometers from an active volcano, the initial findings were promising and justified further investigation in a real-life environment.

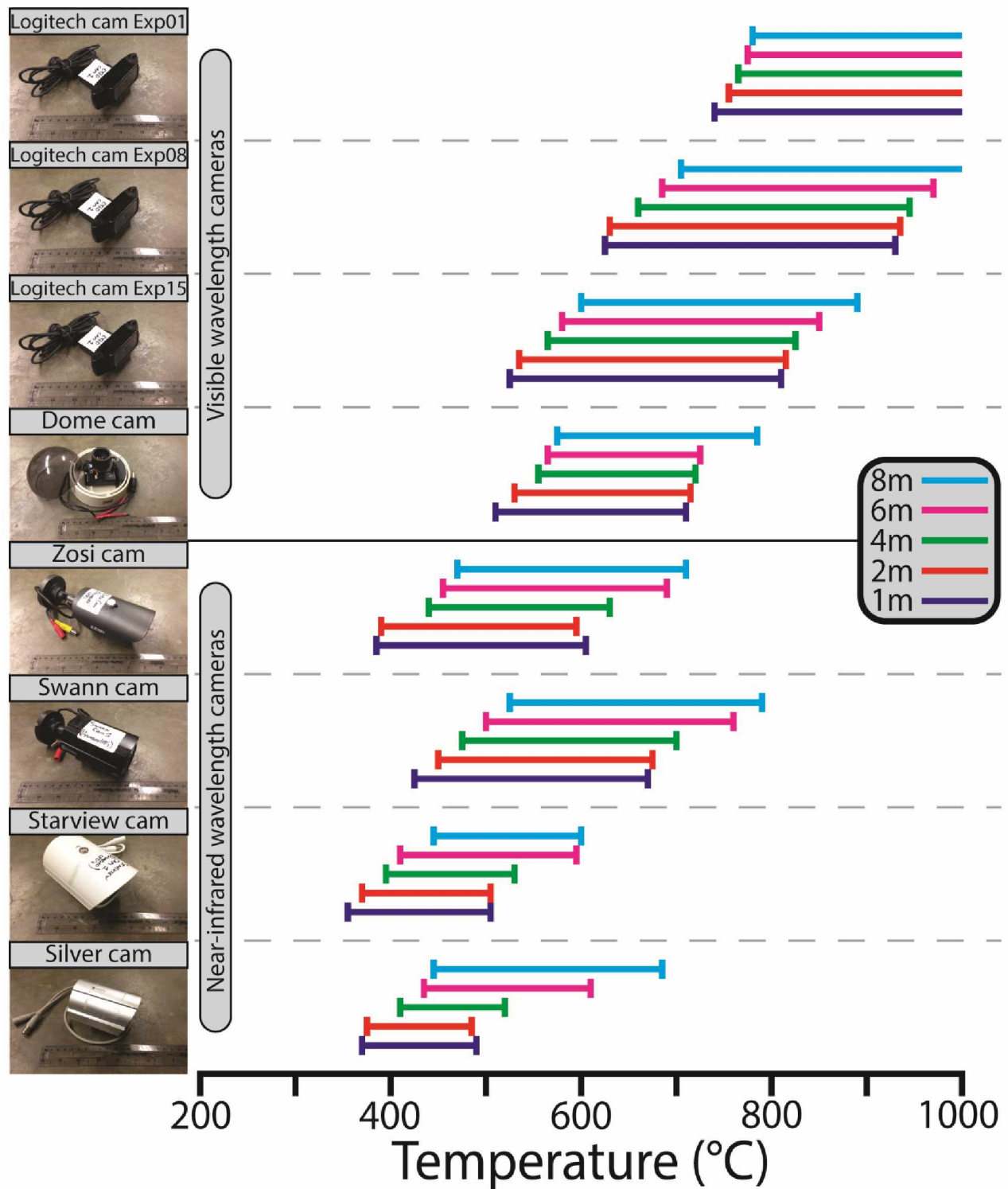


Figure 2.11. Summary of all results from the spectral response experiments showing each camera's initial detection and saturation temperatures at all five distances.

2.5 Field Campaign Deployment (Stromboli Volcano)

After completing the preliminary laboratory spectral response experiments, a dedicated field campaign to Stromboli volcano, Italy was conducted, to determine if pixel brightness could be directly correlated to eruption temperatures. For this campaign, all simulated methods performed in the laboratory setting were applied to actual volcanic eruptions at Stromboli. A total of 250 GB worth of data was acquired from 10 days of image acquisition on Stromboli island at various locations. This field campaign was completed with the assistance of Dr. Peter Webley, Dr. Jonathan Dehn, Christina Dehn, and Jill Shipman.

2.5.1 Location

Three different locations were used to capture images of the active crater terrace, which consists of three distinctly unique active craters (Figure 1.1). Location 1 is at the summit of the volcano on the eastern side with a primary view of eruptions from the northeast (NE) crater, as well as some central crater eruptions, approximately 400 m away. Location 2 is at the base of the volcano, on the terrace of a local restaurant, approximately 1,900 m away, with a primary view of the NE crater eruptions, although eruptions from other craters are captured if large enough. Location 3 is at the summit of the volcano on the southwestern side with a primary view of eruptions from the southwest (SW) crater, and some central crater eruptions, approximately 300 m away.

2.5.2 Equipment Used

Both the Dome and Silver cameras (Figure 2.3A and 2.3B) from the prior laboratory testing, and the Ethernet and Star cams (Figure 2.4A and 2.4B) were used for this field campaign to Stromboli. Additionally, both FLIR TIR cameras (Figure 2.4C and 2.4D) were used to enable direct comparisons between temperatures and pixel brightness in a volcanically active

environment. The Dome and Silver cameras were both wired to be powered by a 12-volt battery, which could be carried to the summit of the volcano. The Ethernet and Star cams are powered by a USB connection and required a dedicated laptop computer. Both FLIR cameras are able to capture images standalone, but for high frame rate recording additional dedicated laptops were required.

2.5.3 Results

2.5.3.1 Location 1

To attempt to capture coincidental and collocated images, several standard and TIR cameras were deployed at the summit of Stromboli, where they were connected to 12 volt batteries, if required, and Windows 7 operating system laptops. By connecting to these cameras through a computer, and therefore dedicated software, fast frame rate capturing was possible. Eruptions from the NE crater were captured at 30 FPS by all cameras for the duration of their battery life. This initial stage of the investigation attempted to determine approximate eruption temperatures from pixel brightness, by collocating a standard and TIR camera, and capturing identical events to perform a pixel by pixel comparison between the two images. The Ethernet and Star cam were not used here because they cannot capture images at a high enough frame rate and rely on third party software screen captures, which can detrimentally affect the pixel values.

Figure 2.12 shows an example sequence of images from the Silver cam (A, C, and E) and FLIR S40 (B, D, and F), captured simultaneously. Due to the nature of the spatter from this eruption (a Type 1 eruption as described by Patrick et al., 2007), it was possible to identify individual projectiles in both image sequences and compare brightness values directly to temperature values. Unfortunately, as the Silver cam saturates at relatively low temperatures, most of the pixels are completely white (the saturation color for 8-bit, 0 - 255 data) and offer little temperature information. This was also true for the Dome cam even though in the spectral response experiments, it saturated at a higher temperature.

One conclusion that can be made from these coincidental images is that under these capturing conditions (such as location, distance, and atmospheric conditions), if the Silver cam images an eruption and saturates, the eruption temperature was greater than approximately 150°C, as some of the coolest pixels that have saturated are colored blue/purple. However, from the spectral response experiments prior to this field campaign, the Silver cam was determined to saturate at approximately 500°C, at only a few meters' distance. Therefore, it is likely several issues are apparent in this environmental setting. For example, the FLIR S40 may not be capturing true eruption temperatures due to ash or gas obscuring the signal, issues with capturing images against a background of empty sky instead of a solid mass, or averaged pixel temperatures resulting from pixel integration.

The same conclusion was drawn from all other coincidental and collocated imagery at this location, with severe issues of camera saturation arising. A decision was made in the field to attempt the same method at a greater distance to see if the results could be improved and a direct comparison performed, if possible.

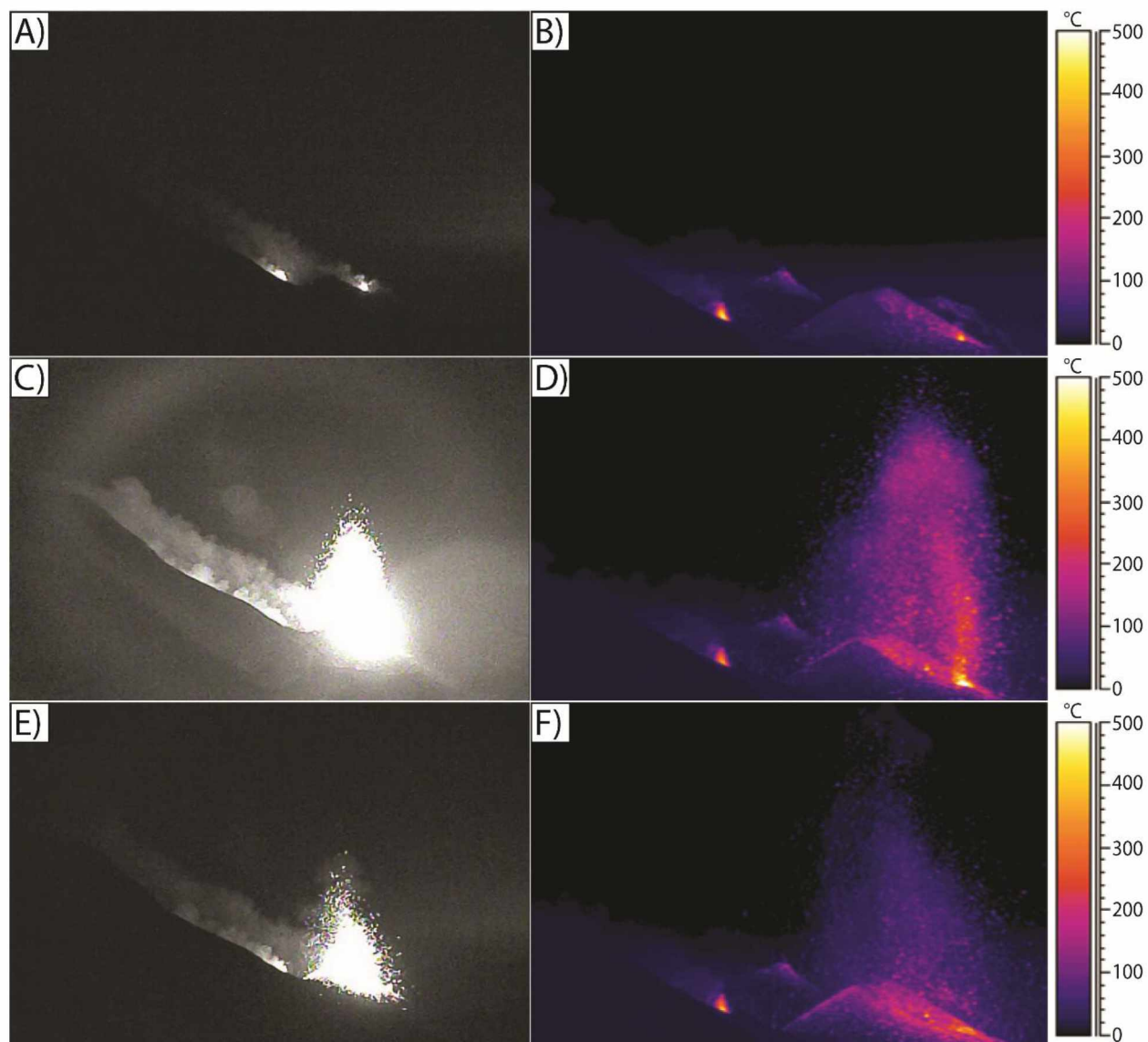


Figure 2.12. Coincidental image sequence of an event at Stromboli volcano on June 29, 2013 from Location 1. A), C), and E) are Silver cam images and B), D), and F) are FLIR S40 images. A) and B) were captured at 19:23:26 UTC, C) and D) at 19:23:34 UTC, and E) and F) at 19:23:39 UTC.

2.5.3.3 Location 2

Figure 2.13 shows an example of collocated and coincidental images from Location 2 of an eruption from the NE crater. All four NIR cameras used for this field campaign, as well as the two TIR cameras, captured the same event, allowing for direct comparisons. As previously identified from the results at Location 1, issues arose regarding the saturation of the NIR cameras when attempting to compare individual pixel brightness values to temperatures recorded by the FLIR cameras. However, identifying the onset of individual eruptions resulting from spikes in pixel brightness values and changes in histogram distribution is feasible in an actual field deployment setting.

As this location was approximately 1,900 m from the source of activity, data captured here is highly relevant to developing a real-time monitoring tool, as most webcams deployed at volcanoes are situated in nearby villages, with a regular power source, or at locations several kilometers from the active crater, where the risk of damage to the installation and telemetering station is significantly reduced. This also allows a wide view angle of the volcanoes flanks in case elevated signals do not originate from the summit crater. A high-resolution camera at the summit of an active volcano would be ideal and provide incredibly detailed imagery, but the risks involved surrounding installation and maintenance are too dangerous. A camera at this location also has a high likelihood of being damaged or destroyed during an eruption, which would cause increased safety issues and financial burden when it needed replacing.

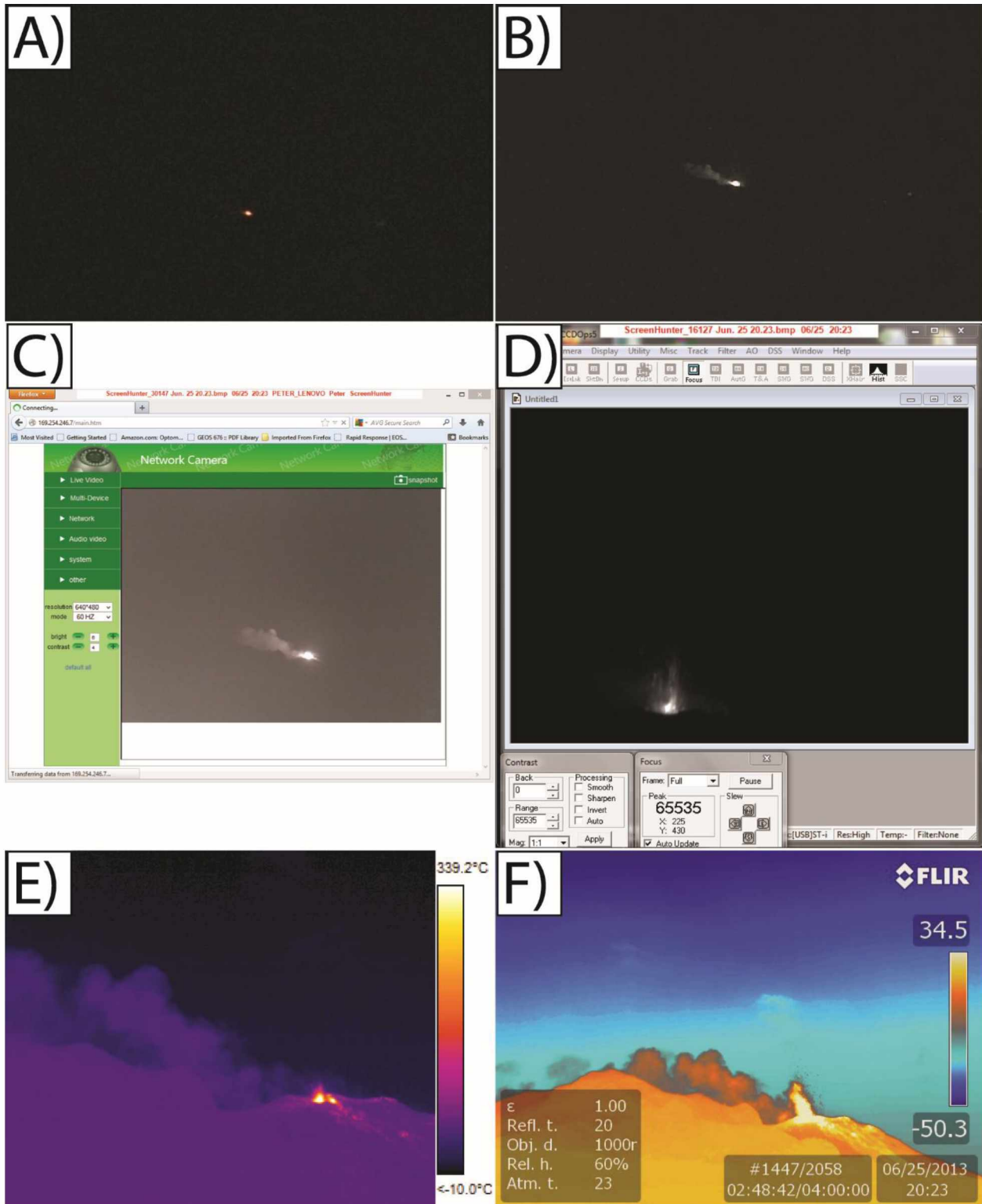


Figure 2.13. Coincidental images of an event at Stromboli volcano on June 29, 2013 at 20:23:28 UTC from Location 2. A) Dome cam, B) Silver cam, C) Ethernet cam, D) Star cam, E) FLIR S40, and F) FLIR T620.

2.5.3.2 Location 3

Due to the serious problems encountered with determining eruption temperature from pixel brightness at Locations 1 and 2, a different approach was utilized at the second summit location (Location 3). Instead of focusing on temperatures after an eruption had occurred and the data had been processed, the investigation pivoted towards detecting volcanic eruptions in near real-time. It was still possible to utilize the cameras available on the field campaign, but the focus would be on relative pixel brightness values, instead of absolute values. For example, if a camera detects a sudden increase in pixel brightness, this would be indicative of an eruption, regardless of the specific values identified.

At Location 3, which primarily viewed activity from the SW crater, several of the NIR cameras were set to record at one second interval speeds to capture several images throughout the duration of an event, which typically lasts 10 - 20 seconds. Figure 2.14 shows a sequence of images captured by the Ethernet cam before, at the onset, during, and after an eruption, with accompanying histogram frequency distributions. By comparing the shape and number of pixels at each unique value in the histogram, it is possible to identify the beginning of an eruption, the size and intensity, and when the eruption is over. For example, Figure 2.14B shows a relatively normal distribution histogram, which suggests only noise and background pixel values are apparent. Figure 2.14D doesn't show a significant change in the overall shape of the histogram, but it does show a small spike in the number of pixels with a high brightness value, which is indicative of the onset of an eruption. Figure 2.14F shows a significantly altered histogram distribution shape, and many pixels with high brightness values, indicating the middle of a large eruption, and Figure 2.14H shows a relatively normal distribution histogram, indicating the eruption is over and pixel brightness values have returned to background levels. Some bright spatter is apparent in the final image, but it is not significant enough to be identified on the histogram distribution plot.

Using the process of identifying bright anomalous pixels compared to average background values, this investigation attempted (Section 3) to create an effective volcanic activity monitoring tool by utilizing webcam imagery and implementing a change detection algorithm to automatically identify events.

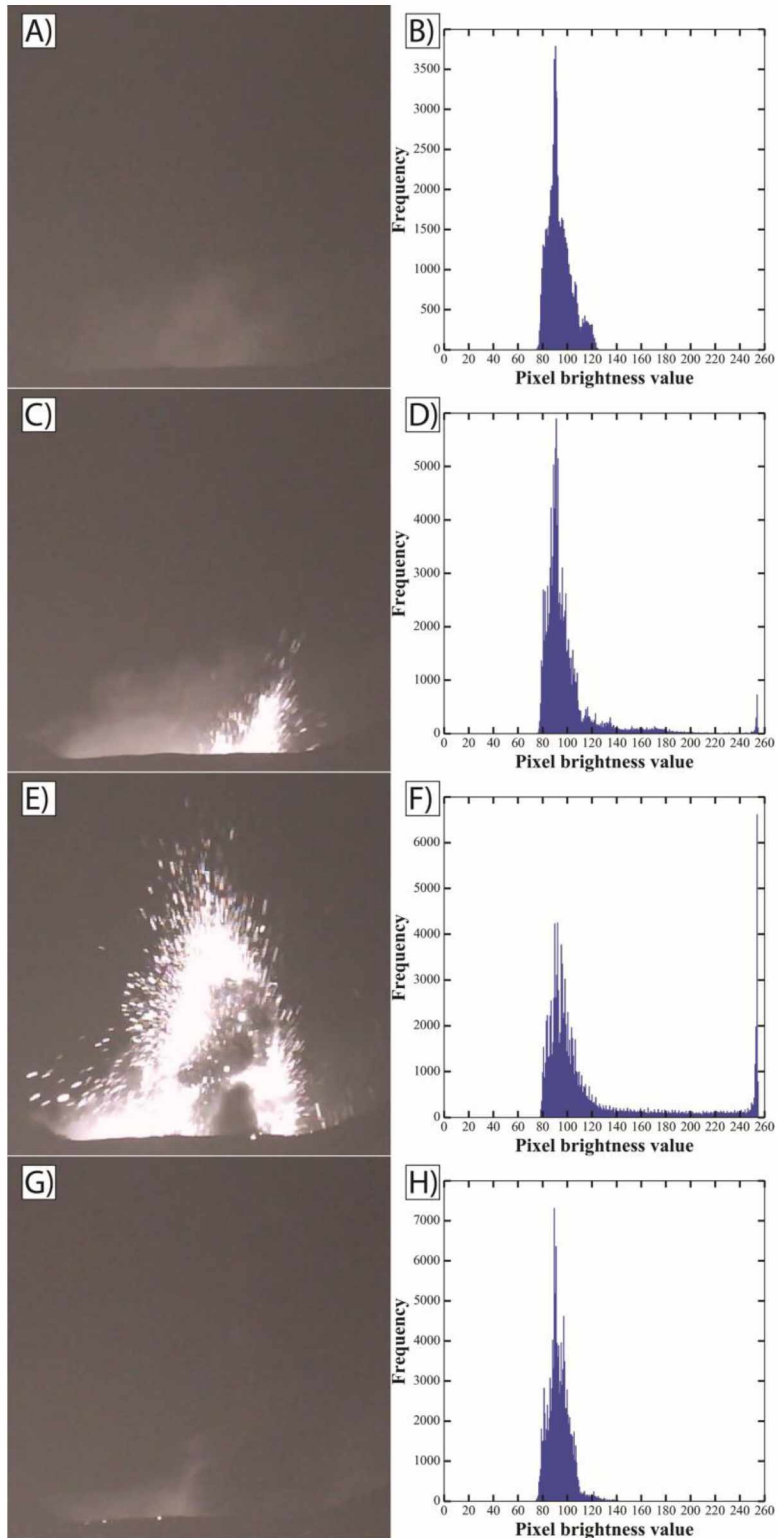


Figure 2.14. An eruption sequence captured by the Ethernet cam at Stromboli volcano from the summit at Location 3 on June 23, 2013. A), C), E), and G) are actual captured images and B), D), F), and H) are the corresponding distribution analysis histograms. A) was captured at 18:37:35 UTC, C) at 18:37:37 UTC, E) at 18:37:40, and G) at 18:37:52 UTC.

2.5.4 Limitations

The primary issue faced during this field campaign was camera saturation. All the RGB and NIR cameras used were almost constantly saturated during an event, even from Location 2 which was almost two kilometers away from the source of activity. Unfortunately, this severely limited direct pixel brightness to eruption temperature comparison capabilities. To mitigate this problem, the camera would either need to be located further away from the active crater, to increase the physical area represented by each pixel, or use adjusted settings relating to exposure times. Both techniques would potentially reduce the effects of camera saturation, but the ability to capture the onset of small scale activity would be detrimentally impacted. Due to this issue, the focus of this investigation was pivoted towards developing change detection algorithms for a real-time monitoring tool instead of direct brightness to temperature comparisons.

Another issue encountered during the campaign was adequate power supply for the equipment. Most of the cameras were powered by dedicated 12 volt batteries and were able to capture images for several days, due to the low power consumption of these cameras. These batteries are designed for motorbikes and jet skis and had to be purchased at the field campaign location as they cannot be transported on commercial flights. However, some of the cameras required a USB connection for power, and all the cameras need a connection to a computer to capture at high frame rates, which requires the use of laptops in the field. Therefore, the time spent recording eruptions at summit locations was limited by the performance of the laptop batteries. Generally, they ran for two hours recording at high frame rates, before needing to be recharged, so careful logistical planning is needed to accommodate this limitation.

Issues also arose from no direct line of site to the source of an eruption at Locations 2 and 3. Sometimes small events or activity from another crater could be heard by the field campaign team or was captured by a TIR camera, but would not be observable in the NIR cameras due to their location and FOV. This causes problems of being able to detect activity as soon as it occurs as the source vent may increase in temperature and brightness significantly before an actual eruption, but is not observable given the location of the cameras relative to the vent. Therefore, valuable time could be lost to provide warnings and alerts of a change in activity, if it is not possible to directly observe the active vent.

Problems of impaired viewing were also encountered as a result of unfavorable atmospheric conditions, where summit clouds would pass between the cameras and the active craters. Although this issue cannot be mitigated, it is one of the largest limitations of this approach. Cloudy data will always have a significant negative effect on most remote sensing methods, from both ground and space observations. There has been some development of real-time space borne algorithms attempting to classify the amount of cloud cover within the data, as far back as the mid 1990's (Higgins and Harris, 1997), while more recent algorithms used multispectral data and atmosphere prediction models to assist in classifying the local atmospheric conditions (Pavolonis et al., 2006). Knowledge of the local atmospheric conditions, and where possible the cloud cover, would help in building a real-time change detection system. However, this would require the inclusion of a weather station in the camera deployment costs, as well as the development of a program capable of connecting this data to the pixel brightness signals from the webcam. Then, a reduction in brightness can be assessed to be truly volcanic or a result of a change in local atmospheric conditions.

Section 3: Automatic Detection Algorithm Development

3.1 Introduction

Over the last decade, the variety, and availability of different methods used to monitor volcanoes has increased dramatically, allowing for a more rapid and accurate detection of eruptions and greatly mitigating the risk associated with this natural hazard. Though this is a significant development for real-time monitoring, almost all these techniques, such as permanent seismic station networks or space borne satellites, require large financial contributions to purchase, calibrate, transport, deploy, and maintain. However, ground based remote sensing, using commercially available equipment, offers a less expensive alternative and can provide essential near real-time data, albeit in a localized area. Several studies have used ground based thermal cameras to monitor volcanic activity, which operate at 7.5 - 13 μm TIR wavelengths, but the cost of purchasing and maintaining a TIR camera can still cost thousands of dollars (Calvari et al., 2004; Harris et al., 2005; Patrick et al., 2007; Patrick et al., 2014; Patrick et al., 2016; Stevenson and Varley, 2008).

A less expensive alternative uses webcam video surveillance systems, now becoming typical monitoring tools at most volcano observatories around the world, which can provide an accurate yet inexpensive way to monitor changes in volcanic activity (Furukawa, 2010; Paskievitch et al., 2006; Patrick et al., 2010; Patrick et al., 2015; Sentman et al., 2006). Webcams that operate at visible wavelengths (0.4 - 0.7 μm) have several limitations that may restrict their usefulness for monitoring volcanic activity. For example, a limited view of the vent due to thick volcanic gases, ash plumes or atmospheric clouds, as well as difficulties in identifying hot volcanic material, which may not be incandescent enough for detection or stand out significantly from natural background illumination. However, cameras that operate at longer NIR wavelengths (0.7 - 1.4 μm) are capable of imaging incandescent targets at lower temperatures, although the effects of volcanic and atmospheric clouds still cause an adverse effect on data collection, albeit to a lesser degree than visible cameras (Sentman et al., 2006).

This section focuses on the development of an algorithm designed to automatically identify changes in volcanic activity resulting from increases in thermal incandescence, and is designed

with financial constraints in mind. Many of the world's most active volcanoes are in remote or financially limited areas where the deployment of expensive monitoring equipment is not feasible (Aspinall et al., 2011). The proposed algorithm will use the freely available Python programming language, to remove the cost of expensive software license fees to potential end users, and all images will be downloaded from freely available websites provided by local volcano observatories. This ensures the algorithm would work smoothly if incorporated directly into the observatories data streams. A preliminary proof of concept investigation was conducted at Stromboli volcano, Italy, using a permanently installed high-cost TIR camera, with the main algorithm development focused on Shiveluch volcano, Russia, with data from a low cost NIR camera.

3.2 Preliminary Investigation: Stromboli Volcano

3.2.1 Image Acquisition

Images are downloaded to a local UAF server from a permanently installed TIR camera located on the flank of Stromboli volcano, approximately 1,440 m (4,724 ft) from the active crater terrace (Figure 1.1), which is streamed live online, courtesy of the University of Florence, Italy (http://193.206.127.20/Stromboli_Monitoring/DatabasePNG/snap_ROC.jpg). The camera has a direct line of sight of any activity occurring from the NE crater, while events from the central and SW craters can also be captured, if they are large enough. The camera operates at TIR wavelengths, but the freely available online images are .jpg files and only contain standard 8-bit image pixel brightness data, ranging from 0 to 255, instead of the raw TIR temperature values. The image resolution is relatively poor compared to most modern cameras, at 352 by 288 pixels, while a new image is downloaded every four seconds, which is the shortest possible time interval given the connection speeds between Alaska (site of the local server for algorithm development) and Stromboli, Italy (site of the camera). This results in approximately 20,600 images every day at ~20 KB each, totaling around 120 GB of data a month, which can cause issues resulting from limited file storage capacity. As eruptions at Stromboli usually last approximately 10 to 20 seconds this rate of sampling is almost guaranteed to capture every individual event, and could potentially be

lowered if there are issues with storage, while still maintaining an extremely high potential of capturing every event.

3.2.2 Manual Observations

To perform a direct comparison between automatically detected and human detected events (equivalent to observatory personnel analyzing every image), a manual analysis was performed on an eleven-day period of events from March 30 to April 9, 2014. During this time, images were captured at approximately eight second intervals, totaling ~226,000 images, which were manually inspected for any signs of an eruption. Note that the temporal frequency of observations was increased to every four seconds after this time period, after the required storage and probabilities of missing an event were assessed. As the images are captured in the TIR, the chosen color scale represents cooler colors as dark blue, and hotter eruption events as white, yellow, orange, or red, depending upon the captured temperature. An event is identified when any form of activity from any of the three vents elevates the pixel brightness above background levels, which can range from strong ballistic eruptions to passive degassing. In the manual analysis performed by the trained observer, the number of unique events were counted instead of the total number of images with an elevated pixel brightness to increase the overall accuracy, as one individual eruption could be captured in multiple images under the right conditions.

3.2.3 Automatic Algorithm Analysis

The automatic detection algorithm initially applied a spatial filter and cropped the target image to a predetermined region of interest (ROI), which removed all potential sources of erroneous data including timestamps and color scale bars, and reduced the computational load required for processing. The maximum pixel brightness was compared to the mean of the entire image plus ten standard deviations. This particular value was chosen because the background of each target image can be classified as relatively constant compared to the previous image, allowing a fixed detection threshold to be used, similar to constant false alarm rate (CFAR) detection

methods used in radar remote sensing (Robey et al., 1992). Both lower and higher numbers of standard deviations were tested, but the results contained either significantly higher false alarms, or significantly lower successful detections respectively. Therefore, ten standard deviations were used to mitigate either of these issues as much as possible. If the maximum value breached this threshold, activity was likely captured in the image with a bright event contrasted by a dark background and labelled accordingly. Each individual image was analyzed in this manner and assigned a one or a zero. Additionally, a temporal filter was applied to determine the number of unique events, compared to the number of flagged images, by restricting a new event assignment to a minimum of once every five images (~40 seconds). This meant that after a successfully flagged event, subsequent images that contained either an ongoing eruption or spattered material around the vent, which was still significantly hot enough (and therefore bright) to be considered thermally anomalous, were not included in the results. Worden et al. (2014) estimated that the majority of eruptions at Stromboli cool significantly enough below TIR camera detection thresholds within seconds to a few minutes, depending upon the eruption type and intensity. However, for this investigation the 40 second temporal filter was deemed significant enough given the images were captured using a relatively high range temperature scale, where objects below approximately 50°C were incomparable to the background.

Comparative results from both the manual and automatic analysis (Table 3.1) show the automatic detection algorithm successfully identified 75.34% of the manually identified events, with the actual number of detected events shown in Figure 3.1. The lowest percentage agreement between the two methods occurred on April 5, when cloudy data severely impacted the number of detectable events and therefore a single missed event accounts for a significantly larger percentage of the total. The total number of automatically detected events includes events that matched the manual analysis and events that were identified an image or two later than the manual analysis, which were small or had a slow start to the eruption that didn't initially breach the automatic threshold at the same time as the manual observation. The overwhelming reason the automatic image analysis underperformed compared to the manual analysis was due to very small events going undetected by the automatic method because they were below the detection thresholds, accounting for an average difference between the two methods of 17.90%. Other sources of discrepancy include noisy images causing a false detection (2.78%) and two individual events occurring within the temporal time frame from different craters being classified as one event

(2.08%). Overall, the automatic algorithm had an average false detection rate of less than 3%, which was more than sufficient to apply it to a much larger dataset and use the number of events per day as a proxy for eruption activity levels.

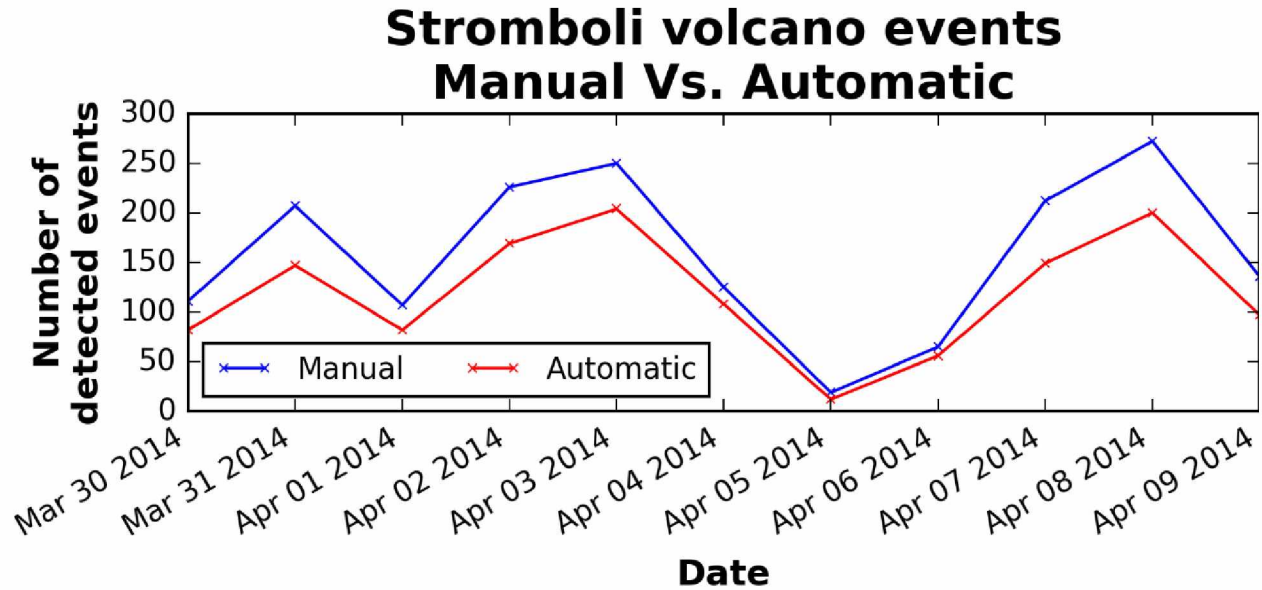


Figure 3.1. Detected events at Stromboli volcano from manual analysis and automatic image processing between March 30 and April 9, 2014.

Table 3.1. Manual and automatic analysis results from an eleven-day period of eruptions at Stromboli volcano.

Date	Number of events detected		Correct (%)
	Manual	Automatic	
2014-03-30	111	82	73.87
2014-03-31	207	147	71.01
2014-04-01	107	82	76.64
2014-04-02	226	169	74.78
2014-04-03	250	204	81.60
2014-04-04	125	108	86.40
2014-04-05	19	12	63.16
2014-04-06	65	56	86.15
2014-04-07	212	149	70.28
2014-04-08	272	200	73.53
2014-04-09	136	97	71.32

3.2.4 Automatic Eruption Activity Detection

The results from the test phase of this preliminary automatic detection algorithm were encouraging with 75% of events successfully identified positive and only a 3% false detection value. The algorithm was then applied to the same dataset for a longer period of time at Stromboli, ranging from March 30 to August 31, 2014, which had an image time interval of four seconds for the overwhelming majority of the data, resulting in approximately 3.25 million images being analyzed. These results are shown in Figure 3.2 and show a dramatic change in levels of activity. An initial activity level of 100 to 200 events per day (EPD) was identified throughout the first half of April, before a decrease in activity levels (10 to 50 EPD) was noted from mid-April to mid-May. A significant increase in levels of activity occurred throughout late May and peaked in late June at 450 to 500 EPD, before dropping off relatively fast and almost ceasing after August 6 (10 to 20 EPD) signaling an end to this increased level of activity. These observations compare closely with actual field observations of increased activity levels beginning in June, and a sustained high number of eruptions until August 6, 2014 when a fracture in the volcano opened, resulting in an effusive eruption (Rizzo et al., 2015) and a cessation of most of the explosive activity.

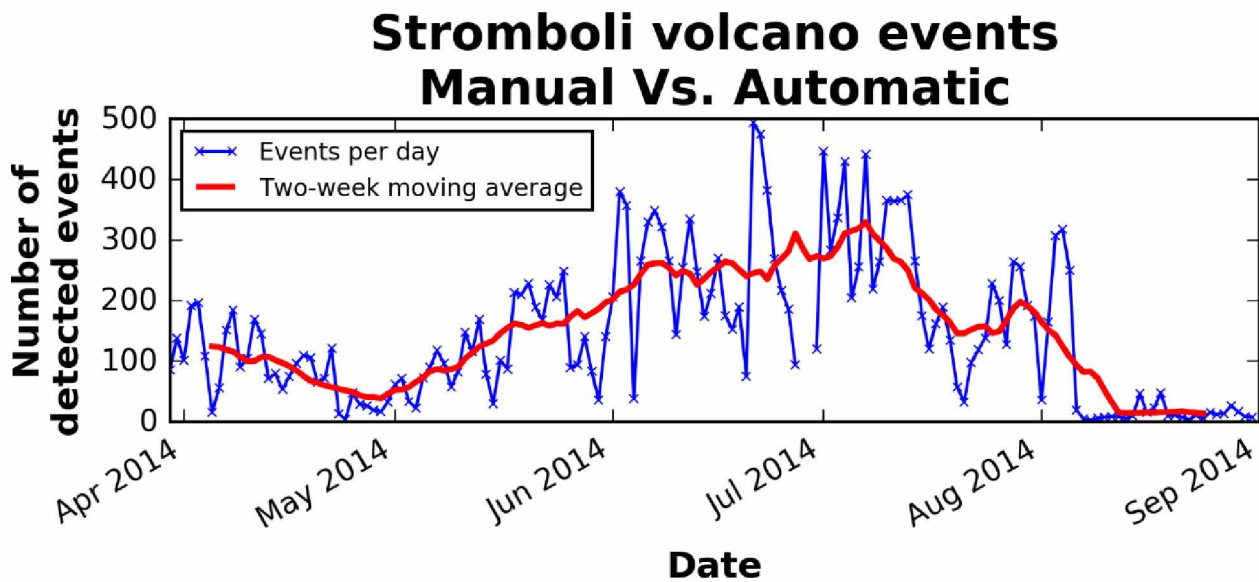


Figure 3.2. Automatically identified events at Stromboli volcano from March 30 to August 31, 2014. Number of unique events per day is plotted in blue with a two-week moving average in red.

3.3 Primary Investigation: Shiveluch volcano

3.3.1 Image Acquisition

Images are downloaded to the same local server at UAF as used for the Stromboli data collection, for data analysis and algorithm development, from a permanently installed NIR webcam located in the village of Klyuchi at the ‘F. Yu. Levinson-Lessing Kamchatkan Volcanological Station’, approximately 45 km (27.96 mi) southwest of the active dome (Figure 1.3). These images are uploaded to the internet courtesy of the Russian Institute of Volcanology and Seismology and the Kamchatka Volcanic Eruption Response Team (KVERT) (http://volcano.febras.net/archive/latest_Shv2.jpg). The camera is an Axis P1364 fixed network camera and has a direct line of sight to the active dome, with a field of view ranging from 92.3° by 67.9° to 33.3° by 24.7° depending upon the focal length used. It has a 1/3” RGB CMOS sensor and operates at visible wavelengths during the day and NIR wavelengths at night. The image resolution is 800 by 600 pixels is updated every 60 seconds. However, for this section of the investigation, an image refresh rate of 15 minutes is used as the higher rate of capture was not available, resulting in ~96 unique images every day at ~40 KB each, totaling ~80 MB a month.

3.3.2 Algorithm Development

The algorithm, hereafter referred to as the Harrild and Webley (HAW) algorithm, was written in Python and tested primarily on a Windows operating system computer, with slight modifications required to run the algorithm on a Linux server. An overview flow diagram is shown in Figure 3.3, which demonstrates the execution process of the entire algorithm. Before running the code, users must read all comments to confirm if all the appropriate Python modules are installed to allow the algorithm to run correctly. For example, most Python packages are pre-loaded with NumPy and SciPy kits, but is it necessary to install them if they are not already available. Here, the graphical user interface (GUI) Spyder was used, which is part of the Anaconda Python package, containing many pre-loaded Python modules used to develop the algorithm.

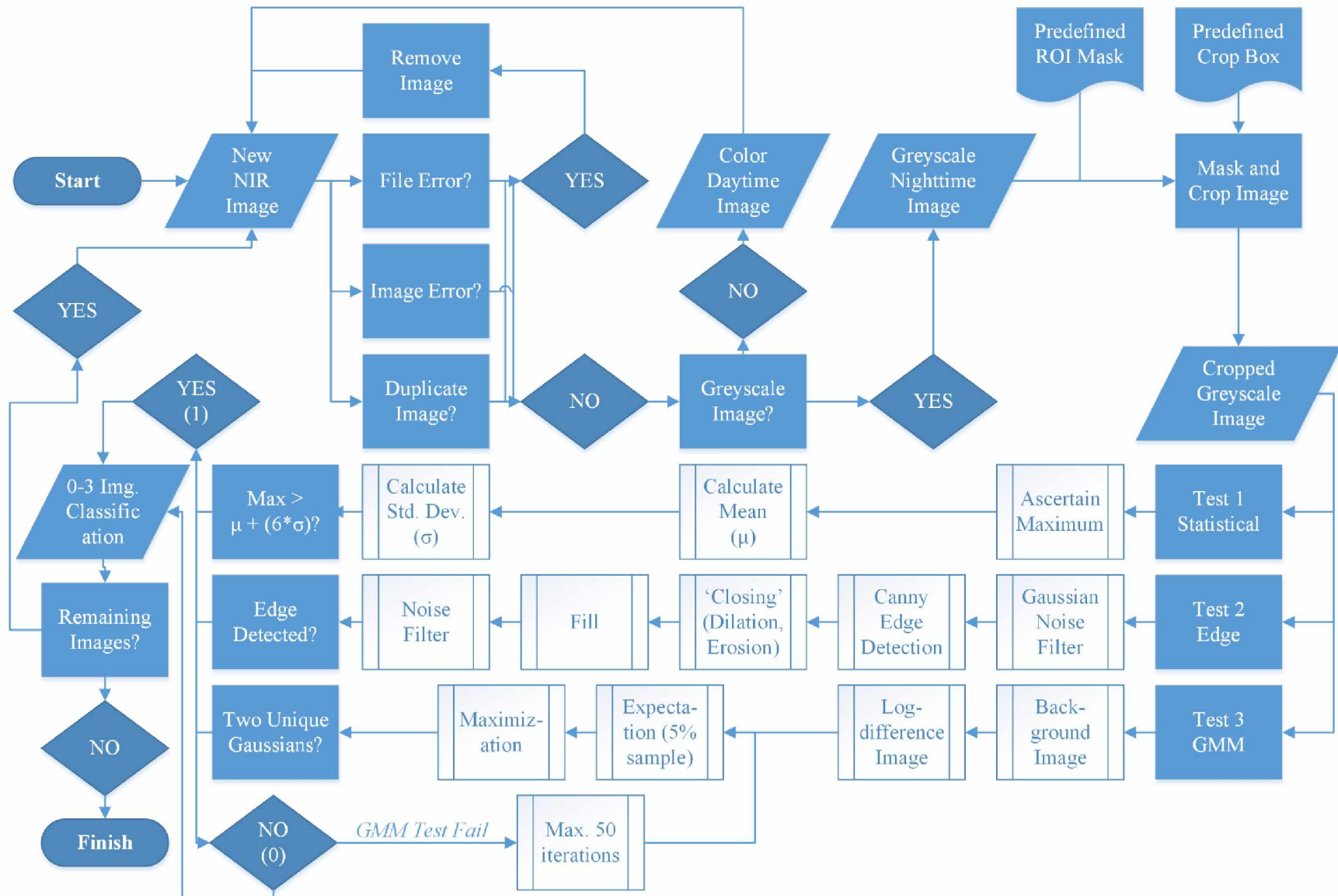


Figure 3.3. Overview flow diagram of HAW algorithm analysis process

3.3.2.1 Remove Erroneous Files

The initial step in the HAW algorithm is to evaluate each new image in the target directory to determine if it is credible. Sometimes, erroneous files are downloaded which have the potential to skew the output results or cause the analysis code to fail. The sources of error usually result from the automatic downloading of images from the target KVERT website, while the camera is trying to write a new image to the directory. This results in completely corrupt files or images that are missing sections of data, which need to be excluded from the analysis. Erroneous files are filtered into two distinct categories, file errors and image errors. File errors refer to completely corrupt image files that will not load into any image viewing application, whereas image errors refer to images that can be read, but display empty or false data. If an image successfully triggers either of these tests, it is deleted and the HAW algorithm proceeds to the next available image.

If an image does not trigger these tests, it is deemed ‘usable’ and is passed to the next stage. Here, images are tested to determine if any duplicate files exist. Sometimes, duplicate files are downloaded from the source website at KVERT when the automatic script is triggered before the host website has updated the image, or if the camera fails and the same image is downloaded over and over again because it is not being updated. The current image file size is compared to the previous image, and if the sizes are identical, the image is deemed a duplicate, deleted, and no further analysis is performed on this image.

3.3.2.2 Image Type

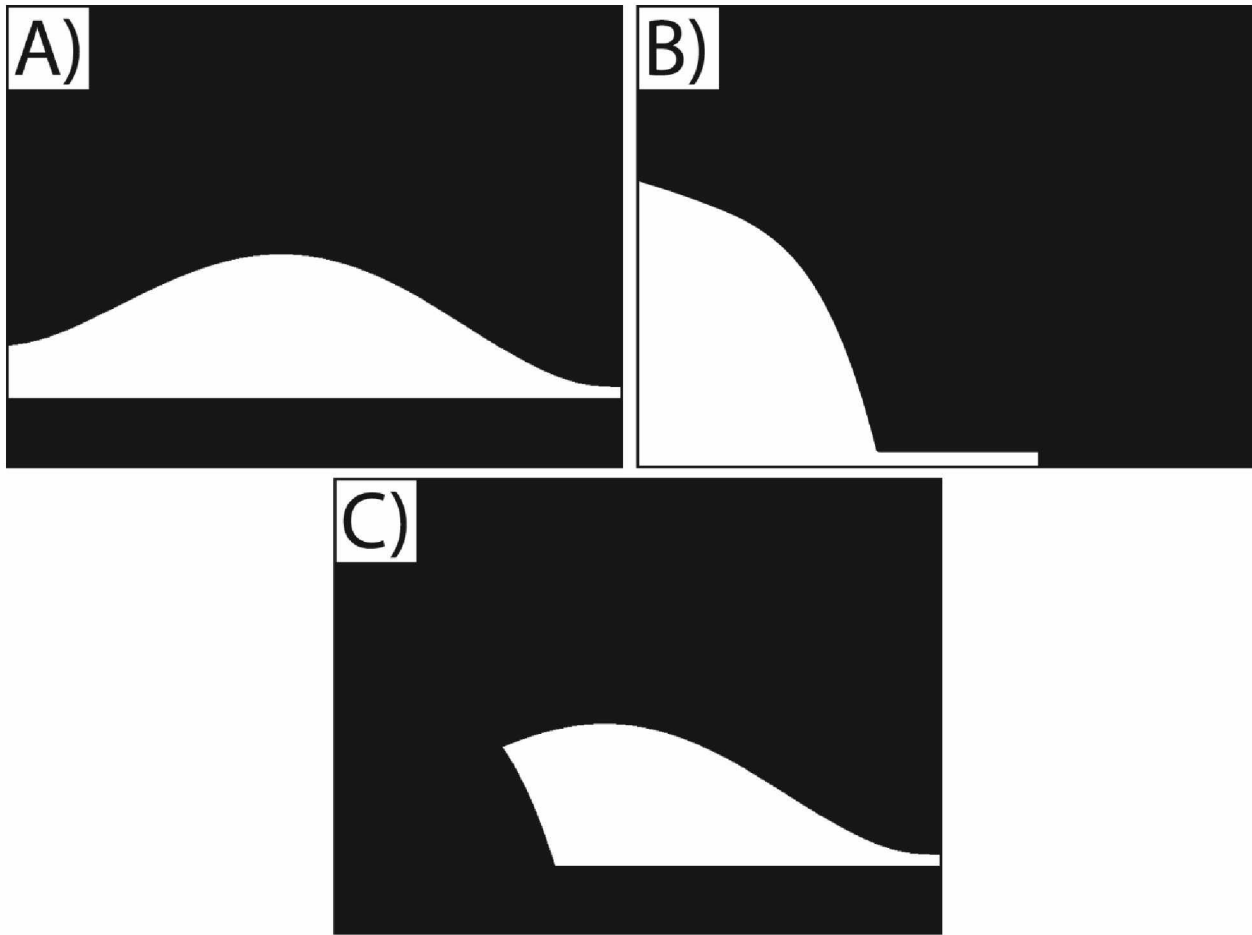
As this investigation focuses on detecting changes in activity resulting from thermal incandescence, nighttime images are used because activity levels have to be extremely high for incandescence to be visible during the day. At Shiveluch, this particular camera has a built-in nighttime detection function. A light sensor on the camera will automatically determine if light levels are too low and will transition from an RGB color image to a grayscale image. However, the image data is not saved as a true, single channel 8-bit grayscale image, but is saved as a 24-bit RGB image with the same values in all three channels. Therefore, to determine if an image is

useful for this investigation, all three channels are compared to one another to verify if every pixel has the same value. If they don't, the image is classified as a color daytime image, not used, and the HAW algorithm proceeds to the next available image. If this image color test is positive, the image is converted to a true 8-bit grayscale image by selecting only one channel of the image and passed to the next section. Here, it does not matter which channel is selected as all three channel values for each pixel are identical.

3.3.2.3 Region of Interest (ROI)

This is the first step of the analysis section of the HAW algorithm with all files making it this far being deemed useable, grayscale, nighttime images. To reduce the chance of false detections and computational power required for processing, the target image is cropped to a predefined region of interest (ROI) using a binary system, which is manually defined and specifically adjusted for each target volcano. For Shiveluch volcano, a ROI is created around the summit of the volcano where the active dome is located (Figure 3.4A). This constitutes a positive event probability map (P_E) and removes almost all other sources of erroneous event detections. For example, any bright pixels from background stars, the moon, or a bright twilight behind the volcano are excluded. Another ROI is created for any sources of likely false detection (Figure 3.4B) to create a false alert probability map (P_{FA}). This includes erroneous sources of bright pixels such as street light illumination in the image or timestamp text. For Shiveluch, this is most of the lower left corner where pixel bleed from street light illumination is an issue. Combining these two maps will output a final probability map (Figure 3.4C) to use for our ROI (Equation 3.1). The target image is then cropped to this ROI leaving only useful pixel brightness values and reducing computational processing requirements. It should be noted that this ROI is predefined and static, which does cause issues if the camera moves slightly throughout the study period. Future algorithm improvements could attempt to define the ROI automatically based upon previous images.

$$P_F = P_E \times (1 - P_{FA}) \quad (\text{Equation 3.1})$$



*Figure 3.4. Predefined region of interest to remove erroneous sources of data using a binary system. White areas equal one, black areas equal zero. A) Mask defining area where probable events would likely occur. Approximately conforms to shape of the volcano, B) mask defining area where false detections would likely occur. Identifies area of pixel bleed from street light illumination and image time stamp, and C) mask defining the final region of interest (Probable events * (1 - False Detections)).*

3.3.2.4 Test 1: Statistical Analysis

The HAW algorithm performs three separate tests on each target image to determine if it is anomalously bright resulting from increased levels of thermal incandescence, the first of which is a statistical comparison test. This involves calculating and testing specific statistical parameters from the image pixel brightness data to determine if thermally incandescent activity is detected. The maximum pixel brightness value is extracted from the cropped ROI image and compared to the mean value of the image plus six standard deviations. This approach was chosen after comparing the results of this test for standard deviation values ranging from two to ten, to a manually (trained observer) analyzed time period. Using six and seven standard deviations yielded similar results (seven standard deviations detected fewer events, but had less false alarms) but using a value of six had a higher positive detection to false alarm ratio and was included for all subsequent tests. If the maximum pixel brightness is greater than this threshold (mean plus six standard deviations) then the image is deemed thermally incandescent and assigned a result value of one for Test 1.

3.3.2.5 Test 2: Edge Detection

The second test performs a Canny edge detection (Canny, 1986), which consists of first applying a Gaussian filter to smooth the image and remove any noise artifacts. A two by two Gaussian filter size is used here as it consistently performed the best during testing. Examples of different Gaussian filter sizes applied to the same image are shown in Figure 3.5. Note that the kernel with $\sigma = 2$ (two by two) captures the edge of the thermal feature and doesn't detect the noisy signals when using a kernel with $\sigma = 1$ (one by one). Four filters are applied to the smoothed image to detect edges in four different directions (horizontal, vertical, and both diagonals) before applying non-maximum suppression to narrow the width of any detected edges to one pixel. Finally, edge pixels are kept or removed using hysteresis thresholding designed to suppress weak and disconnected 'edges' (Figure 3.6C).

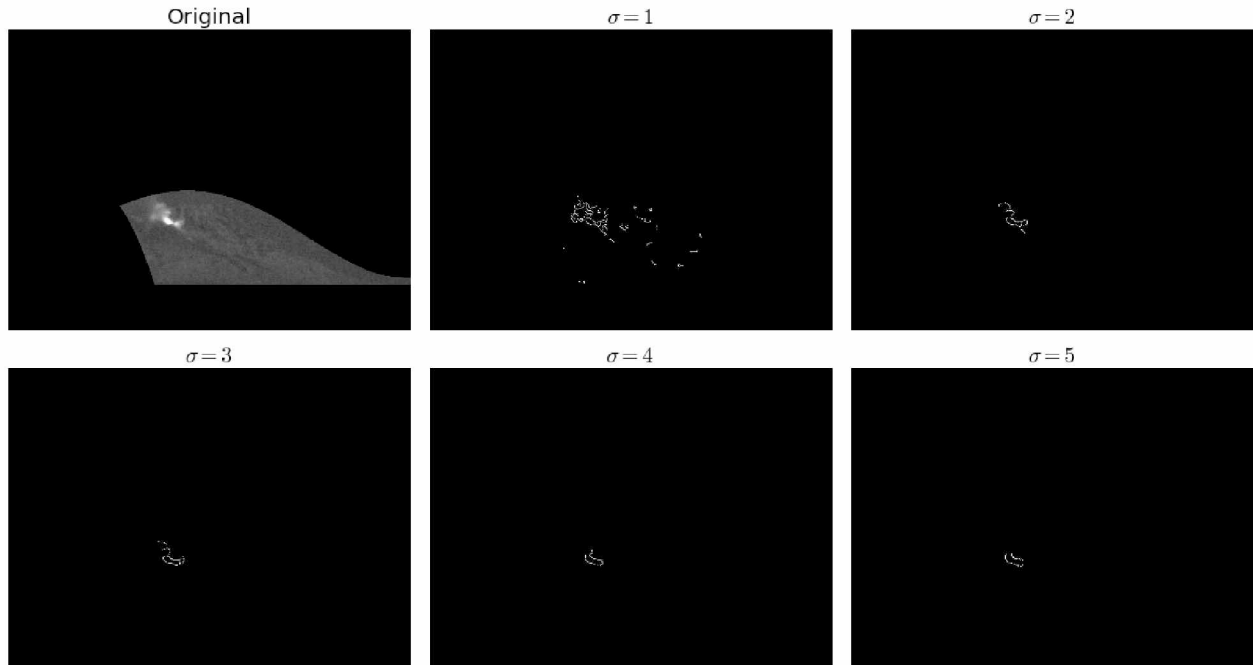


Figure 3.5. Example images demonstrating the effect of varying Gaussian kernel sizes (sigma (σ), from 1 to 5) used to smooth noise in the image before the Canny edge detection process.

The resulting image containing the edge detection is then ‘closed’ (Figure 3.6D). This is the name commonly used to refer to an image dilation, followed by an erosion, using a specified structuring element. Here, a spherical array with a diameter of five pixels is used. This dilation and erosion process will close any small gaps in the edge detection process and create a solid edge. Completed edge detections are then filled to create a large solid area where thermally incandescent activity has been identified (Figure 3.6E). A final step cleans the output image and removes any artifacts that are less than 10 pixels in size, which eliminates any noise resulting from the edge detection process (Figure 3.6F). If the final processed image contains any pixels that have a value greater than one (the edge detection process uses a binary approach, one = edge, zero = no edge) the original image is rated one and determines that the Canny edge detection test identified a significant thermal anomaly, and the algorithm concludes the second test.

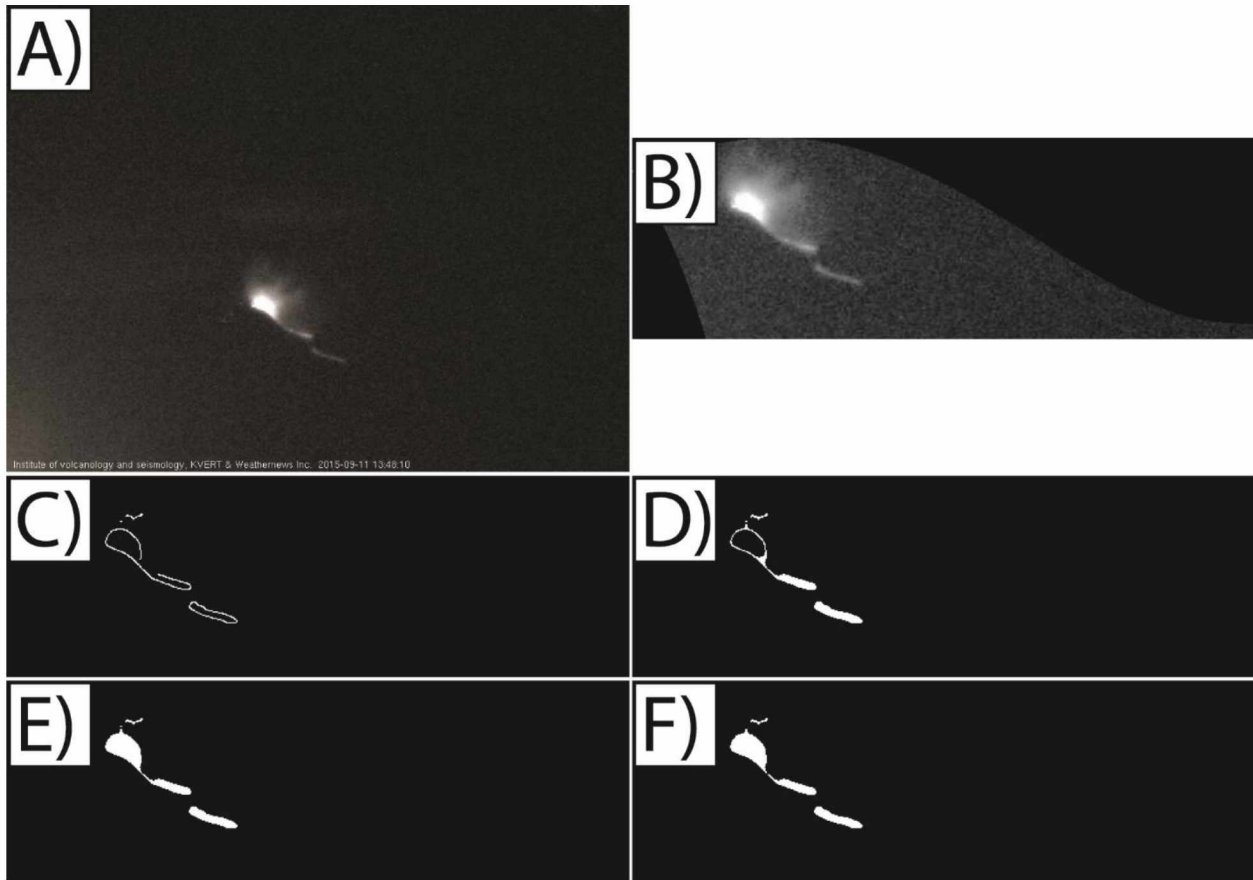


Figure 3.6. Example images of the edge detection process. A) Original target image, B) cropped target image, C) Canny edge detection image, D) 'closed' image (dilation and erosion) to fill small gaps in edge detection, E) filled image to create a solid region of activity, and F) cleaned image to remove any small noise artifacts.

3.3.2.6 Test 3: Gaussian Mixture Model (Expectation-Maximization)

The third and final test in the HAW algorithm was developed to determine if there are two separate Gaussian distributions within the frequency histogram data from each captured webcam image. One distribution would represent the background and the other would represent the thermal signal from an event. To be able to perform this test automatically, a Gaussian mixture model (GMM) approach was used. This is a probabilistic model that analyses a set of data points and assumes the values are generated from a fixed number of unique Gaussian distributions with unknown parameters (Moon, 1996). The model runs through a specified number of iterations and determines the probability of each data point belonging to a specific and unique Gaussian distribution. For example, with an image of a thermally incandescent volcano, the model will initially assign equal weighting to all image pixels, but as each iteration completes, the model is able to determine two definitive data point clusters, namely the pixel brightness values from the anomalously hot target compared to the darker background.

1) The initial step of this test is to create an averaged background image which is then compared to the current target image. For this investigation, a time frame of seven days was chosen. While this time frame provides the algorithm with a background signal from the previous week, and eliminates any possible errors that could occur from longer time periods and any seasonal changes, this value is arbitrary and can be altered depending upon computational resources and image capture rates. The date and time of the target image is determined, and all prior images within the seven-day time frame were allocated for use in the background image. However, problems arose from using all images in the background window, as data points from shortly after sunset (where twilight is still apparent) were being directly compared to much darker images in the middle of the night. Therefore, a two-hour window was applied to the selection of background images which encompassed the target image timestamp. For example, if the target image was captured at 03:00 UTC, each image from the past seven days within a timeframe of 02:00 to 04:00 UTC, which had not been classified as including thermally anomalous activity and excluding the target image itself, would be allocated for use in the background image. With a new image captured every 15 minutes from this webcam, the average background is normally comprised of 56 unique images.

2) The second step of this test is to create a log-difference image for use in the GMM test, similar to methods used by Ajadi et al. (2016), to suppress the background and highlight any anomalous signals. The target image is divided by the averaged background image calculated in step 1, and logged using a base 10 log. By log-transforming the data, differences between any potentially bright pixel values and the darker background are accentuated and more easily identified, erroneous sources of background illumination are significantly reduced, and the resultant data will have a zero mean and would fit a true Gaussian distribution if no activity was present. This step is important so the GMM test is able to evaluate the presence of two Gaussian distributions.

3) The final step in this test is to apply the GMM test to the log-transformed difference data. The GMM approach is designed to probabilistically assign each pixel brightness value to one of two distinct Gaussian distributions. The number of distributions can be altered in the Python code, but for this investigation two distributions are used to represent a darker background and a brighter incandescent target. If an image has no bright thermal activity, only one distribution will be present and the GMM approach will be unable to find the two requested distributions. In this instance, the image would not pass this test and would be given a zero on Test 2. The number of iterations and sample size can also be altered in the code depending upon available computational power. Here 50 iterations and a 5% data sample size are used, where the sample size value represents the number of pixels in the ROI being selected for the GMM test.

Each iteration is comprised of two distinct steps, 'Expectation' and 'Maximization'. The expectation step assigns a weighted probability to each individual pixel based upon the likelihood that particular pixel belongs to one of the two Gaussian distributions. The first iteration initially assigns every other pixel to one of two distinct distributions, creating a completely random distribution with no bias. All subsequent iterations assign a calculated probability to each pixel computed by the maximization step, which determines statistical criteria including distance from the center of both Gaussian distributions and the standard deviation of each pixel. This information is iteratively calculated and used to assign varying probabilistic values to individual pixels during the expectation stage. As the number of iterations increases, each pixel is determined to belong to one of the two Gaussian distributions.

Figure 3.7 demonstrates the GMM testing process for a thermally incandescent example image at Shiveluch volcano. All pixels in the image are given an equal weighting to one of the two Gaussian distributions to create a completely random initial assignment in the first expectation step (Figure 3.7B). In this example, red and blue colors are used to visually represent both distributions and demonstrate the GMM process. The first maximization step calculates the probability each pixel has of belonging to both distinct Gaussian distributions. This results in an almost equal assignment to both the red and blue distributions because of the uniformed allocation of the initial assignment, causing all the pixels to show as purple for the second expectation step. All subsequent maximization steps that occur for each iteration recalculates each individual pixel's probability of belonging to one of the defined Gaussian distributions using the updated values from the previous iteration. This method of adaptive learning is very effective at identifying unique clusters of pixels that belong to different Gaussian distributions, as shown in the remaining images in Figure 3.7, until two distinct clusters of background and thermally incandescent pixel values are identified.

If the statistical probability of all pixels belonging to a unique Gaussian distribution does not change from one iteration to the next, accurate to six decimal places, the test is considered successful and the image receives a value of one for the GMM test, which concludes the second test. However, if no distinct distributions can be identified, normally due to the image only consisting of dark background pixel values and therefore only one Gaussian distribution, the test will run for the full number of iterations before finishing and determining that the target image does not pass this test, assign a value of zero, and begin the subsequent stage of the HAW algorithm.

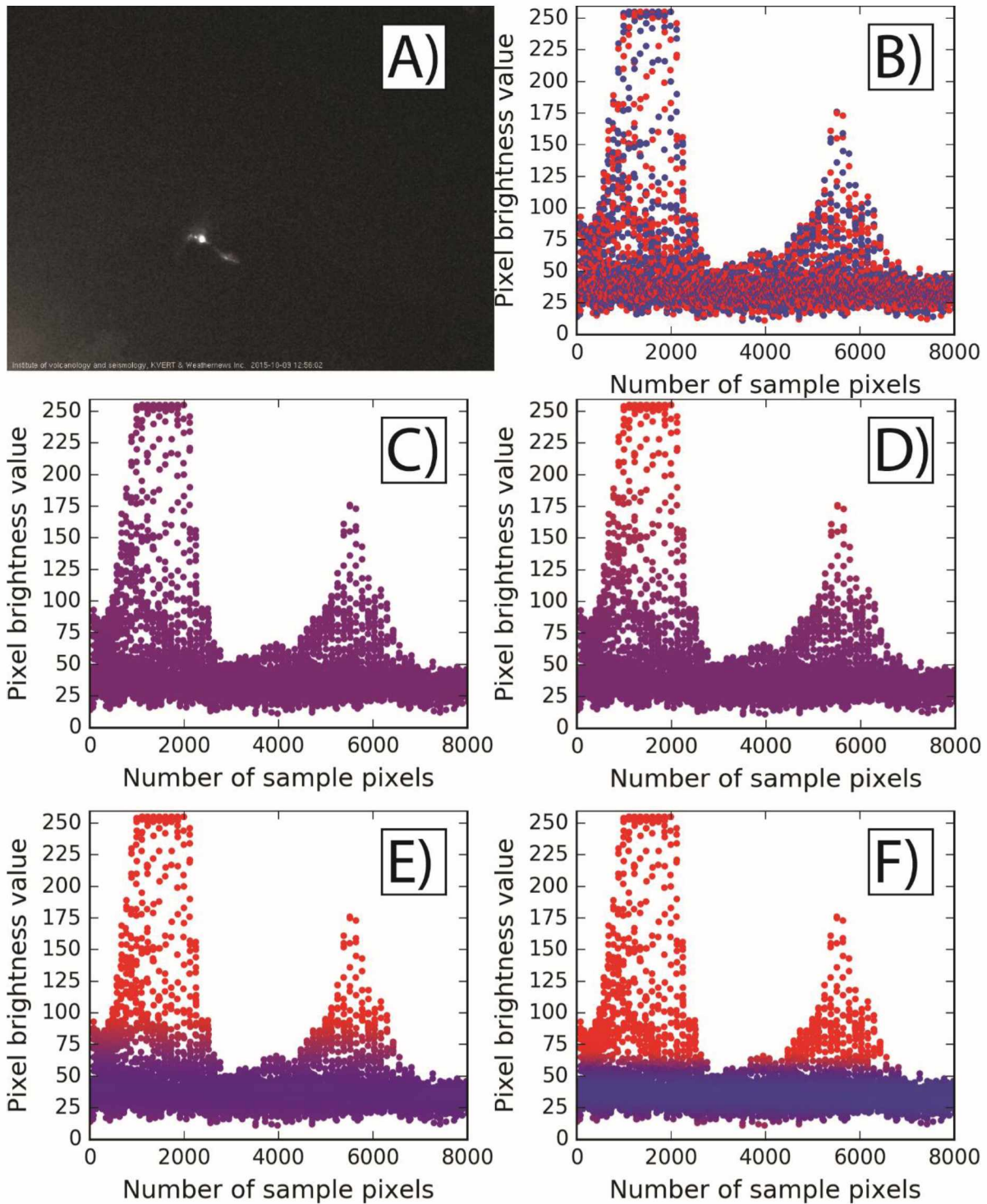


Figure 3.7. Results from the Gaussian Mixture Model test applied to a thermally incandescent image. A) Original image of volcanic activity, B) initial expectation step assigns all pixels to one of two Gaussian distributions, C) first iteration loop with all pixels equally assigned, D) second iteration loop with higher value pixels assigned increased probability of belonging to the 'red' distribution, E) third iteration shows distinct boundary between background and incandescent distribution, and F) fourth iteration with clear image divide between two Gaussian distributions.

3.3.2.7 Conclusion

Overall, each target image will be assigned a thermally incandescent likelihood value between zero and three. A rating of zero means all three tests determined no activity occurred, whereas a rating of three means all tests identified a source of incandescence. Any images receiving a rating greater than zero will be considered successfully identified by the automatic algorithm for this investigation, as the image could have been detected by any of the three available tests.

3.3.3 Algorithm Application

Images from the permanently installed camera at Shiveluch volcano were downloaded throughout 2015 and provide a stable dataset to test the effectiveness of this algorithm. A total of 29,662 images were downloaded and manually analyzed to accurately identify all thermally incandescent activity. Here, any level of activity is manually classified as an event, regardless of size or intensity. From the HAW algorithm analysis, a total of 1,715 images were determined to contain varying levels of activity, 17 images were identified as erroneous, and 25 images were identified as duplicates. It should be noted that most detected activity (~63%) occurred during the first and last three months of the year. However, this bias is a result of longer hours of light during the summer reducing the number of incandescent images detected and is not necessarily indicative of increased activity.

3.3.3.1 Successes

The HAW algorithm correctly identified and deleted 100% of the erroneous files throughout the year, with 17 images removed leaving 29,620 images for analysis (after removing the duplicate files too). All 17 files were removed from the analysis because of ‘image errors’ resulting from the file being downloaded while the source camera overwrote the data with a new

image instead of corrupt ‘file errors’. Of the 29,620 useful images, 100% of the daytime color images (17,921) were identified and disregarded, leaving 11,699 grayscale images for analysis, which are summarized in Table 3.2. There are many ways to critically analyze the overall success of an algorithm (Table 3.3), several of which are outlined in Schaefer (1990), which are used to assess the success of weather forecasting, and Webley et al. (2009b), which are used to determine the success of volcanic ash modeling. To develop a more comprehensive evaluation criteria of the HAW algorithm’s success, several basic parameters in the confusion matrix are defined as follows: an event that is both manually and automatically identified is referred to as a True Positive (TP); a non-event that is both manually and automatically identified is referred to as a True Negative (TN); a non-event that is manually identified but flagged as an event by the algorithm is referred to as a False Positive (FP); an event that is manually identified but missed in the automatic analysis is referred to as a False Negative (FN).

Table 3.2. Confusion matrix results from the comparison of a manual analysis and the automatic algorithm analysis for NIR webcam images at Shiveluch volcano from 2015.

		Manual analysis	
		Event	No event
HAW algorithm analysis	Event	Type: True Positive (TP) Number of images: 1,540	Type: False Positive (FP) Number of images: 175
	No event	Type: False Negative (FN) Number of images: 293	Type: True Negative (TN) Number of images: 9,691

In addition to the TP, TN, FP, and FN, the overall accuracy of the HAW algorithm is defined as the ratio of all events where the manual and automatic analysis agree to the total number of events (Equation 3.2). Here, 11,231 eruptions images were classified the same by the manual and automatic analyses for an overall accuracy value of 0.96 or 96.0%.

$$\frac{TP+TN}{TP+TN+FP+FN} \quad \text{(Equation 3.2)}$$

The frequency of detected events was calculated, which describes the total number of manually identified events divided by the total number of images analyzed (Equation 3.3), which can be used to determine an average time interval between eruptions. Here, 1,833 eruption images were manually identified from a total of 11,699 images for a frequency value of 0.157, meaning 15.7% of all images analyzed will likely contain an event.

$$\frac{TP+FP}{TP+TN+FP+FN} \quad \text{(Equation 3.3)}$$

The probability of detection (POD) is defined as the ratio of events that are correctly identified by the manual and automatic analysis to the total number of events (Equation 3.4). Here, a POD value of 0.840 or 84.0% is calculated. This value essentially determines the likelihood that if an event is automatically detected by the HAW algorithm, it is a true event. If the number of FN images is zero, then the POD value is one. This POD means approximately eight in ten images are accurately identified, which is a relatively high value for the initial test of this algorithm.

$$\frac{TP}{TP+FN} \quad \text{(Equation 3.4)}$$

The false alarm rate (FAR), the inverse of the success rate (SR), is defined as the ratio of the number of false alarms to the total number of predicted events (Equation 3.5). Here, a FAR value of 0.102 or 10.2% is calculated. This value is relatively low and shows the HAW algorithm in its current form will only falsely detect one in every ten images analyzed.

$$\frac{FP}{TP+FP} \quad \text{(Equation 3.5)}$$

The critical success index (CSI) is defined as the ratio of successful event detections to the total number of events, excluding the TN images (Equation 3.6). Here, a CSI value of 0.767 or 76.7% is calculated, which shows that combined with all other previous statistical success tests, the HAW algorithm has a high probability of automatically detecting a thermally anomalous event.

$$\frac{TP}{TP+FP+FN} \quad \text{(Equation 3.6)}$$

Table 3.3. Statistical classification confusion matrix results from the HAW algorithm analysis of NIR webcam images at Shiveluch volcano from 2015.

	Equation	Value (Percentage)
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$	0.960 (96.0%)
Frequency	$\frac{TP + FP}{TP + TN + FP + FN}$	0.157 (15.7%)
Probability of detection (POD)	$\frac{TP}{TP + FN}$	0.840 (84.0%)
False alarm rate (FAR)	$\frac{FP}{TP + FP}$	0.102 (10.2%)
Critical Success Index (CSI)	$\frac{TP}{TP + FP + FN}$	0.767 (76.7%)

By further analyzing the individual confusion matrix results from each test (Table 3.4), it is possible to determine their overall success. The statistical test performed extremely well with a high POD, low FAR and a CSI value similar to the entire algorithm. Ideally, a reduction in the FAR value would significantly improve this test, as a relatively high number of false negatives were identified due to ambient light being included in the ROI as the camera moves slightly throughout the year. The edge detection test had a slightly lower overall accuracy than the statistical test, but had a significantly low FAR with only 16 events falsely identified for the entire year. Again, a reduction in the false negatives, where small events are indistinguishable from background levels, would greatly improve this test. Finally, the GMM had the lowest performance values of all tests. Although the overall accuracy is still relatively high (89%) the majority of the

images successfully identified contained no activity, represented by a low POD value of 35.8%. Also, although the number of false negatives for this test was only 10%, it was still the highest of all three tests. Ideally, improvements to this particular algorithm would reduce this number and increase the overall CSI index of this test.

Table 3.4 Statistical classification confusion matrix results from individual tests in the HAW algorithm analysis of NIR webcam images at Shiveluch volcano from 2015.

	Equation	Statistical test	Edge detection test	GMM test
True Positives	TP	1,419	985	657
True Negatives	TN	9,749	9,850	9,795
False Positives	FP	117	16	71
False Negatives	FN	414	848	1,176
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$	0.956 (95.6%)	0.926 (92.6%)	0.893 (89.3%)
Frequency	$\frac{TP + FP}{TP + TN + FP + FN}$	0.131 (13.1%)	0.08 (8.0%)	0.06 (6.0%)
Probability of detection (POD)	$\frac{TP}{TP + FN}$	0.774 (77.4%)	0.537 (53.7%)	0.358 (35.8%)
False alarm rate (FAR)	$\frac{FP}{TP + FP}$	0.08 (8.0%)	0.02 (2.0%)	0.10 (10.0%)
Critical Success Index (CSI)	$\frac{TP}{TP + FP + FN}$	0.728 (72.8%)	0.537 (53.7%)	0.345 (34.5%)

3.3.3.2 Failures

The 25 images that were identified as duplicates by the HAW algorithm were all falsely labelled as such. Although the number of images erroneously assigned as duplicates is relatively low, it shows the current method of image duplicate identification is not completely robust and needs to be improved upon to limit falsely discarded images, particularly when analyzing higher temporal resolution datasets where the likelihood of duplicate images is significantly increased.

Most false positives resulted from twilight conditions behind the volcano summit appearing bright compared to the volcano itself and triggering the tests. This is particularly difficult to mitigate as the camera physically moves throughout the year due to weather conditions or human interference, so it is extremely difficult to determine a ROI that will not include any background for all twelve months. Another significant source of false positives were very cloudy images causing the statistical method to falsely identify activity because of a small standard deviation relative to the maximum pixel value, due to almost homogenous pixel brightness values.

Most false negatives resulted from major cloud obstruction or very small events that were difficult to automatically identify using these three specific tests. The system would often detect the main portion of an event, but miss a small number of images leading up to and shortly after the strongest and brightest volcanic activity, when activity levels were much lower. Another significant source of missed events for the statistical test resulted from large events with many bright pixel values. The ROI would become almost entirely saturated from these large events causing the statistical difference between the event and the background to become insignificant. This has the same effect as an ROI filled with all dark background pixels. The GMM test's performance on its own was surprisingly poor, when compared to the time and computational power required to perform it. Of all 1,833 manually identified images, the GMM test successfully identified 657 (35.84%). However, of these 657 images, only 49 (15.22%) were successfully identified solely by the GMM test, which raises questions of whether this test should be included moving forward as the length of time required to perform this test is significant when trying to analyze data in near real-time for volcano monitoring. However, if the test is not included due to time and computational limitations 49 potentially significant events would have stayed undetected.

3.3.3.3 Improvements

The main improvement moving forward is to reduce the number of detected false positives. Most of these images were identified shortly after sunset or shortly before sunrise causing a relatively bright signal to be included in the algorithms tests. Ideally, the camera would remain in a fixed position throughout the year and the ROI would mitigate this problem, but realistically, this scenario is unlikely to occur. Therefore, a test to determine the level of twilight in the image

needs implementing to ensure the number of false detections is reduced. This could be in the form of another test that determines sunrise and sunset times based on a volcano's specific location (latitude and longitude) and only evaluates images that are within a specific time frame, centered on the specified sunrise and sunset times, which reduces the chance of erroneous twilight images being identified. Also, by performing all three tests on the log-difference image used in the GMM test, this should remove a significant amount of the erroneous bright background values.

A general area of improvement for the algorithm is its overall performance. As these testing methods are designed for use in near real-time, to detect increases in volcanic activity as soon as possible, speed and total completion time are extremely important factors. Significant improvements to the GMM test are also required to continue to apply this test in the algorithm as, while it can detect events not flagged by the other tests, the contribution to the overall results is relatively low compared to the time and computational power required to perform the test. For example, if this test was excluded from the algorithm, the overall completion time would fall by approximately 90% but only 49 images of the total 1,833 images (2.67%) would not be successfully identified. Therefore, alterations to the required number of GMM test iterations and sample size are required. For this investigation, 50 iterations and a 5% sample size were used, but these values could be significantly increased. However, further testing is required because the GMM test attempts to determine two individual clusters of pixels, which may in fact be achieved faster and more efficiently by using a lower sample size and increasing the sampling space between selected data. However, this would likely cause an increase in the number of images classified as false negatives. As with all automatic detection systems, a compromise between accuracy and efficiency has to be made.

It should also be noted that although the number of significant events identified is relatively high, the number of these events that actually produced a hazardous eruption is unknown. For example, the lava dome at Shiveluch volcano may exhibit an increased thermal flux, and as a result become thermally incandescent, but this type of activity is relatively harmless, until a collapse occurs. The fine line between effective event identification and over-reporting non-threatening activity needs to be carefully defined.

3.4 Conclusions

Presented here was an efficient and effective tool to monitor changes in volcanic activity. Overall the algorithm performed better than expected at identifying thermally anomalous activity. During preliminary testing of data from Stromboli volcano, a ten-day period of activity from March 30 to April 9, 2014 was manually analyzed to identify how many images contained thermal anomalies and compared to the automatic analysis results. The images were downloaded from a website which live streamed the data, allowing an eight second temporal resolution for this dataset, the shortest connection speed times between Alaska and Italy. Over 226,000 images were examined and 1,730 unique eruptions were manually identified, compared to 1,306 events that were automatically detected (Figure 3.1). An overall accuracy rating for this relative rudimentary initial algorithm was approximately 75%, after accounting for false detections and missed events.

Building upon the marked success of this algorithm to successfully identify activity, a six-month period of activity was analyzed from March 30 to August 31, 2014, approximately 3.25 million images. Figure 3.2 shows the dramatic changes in levels of activity which were successfully identified by the automatic algorithm, which included a decrease in ‘normal’ levels of activity before dramatically increasing for the duration of summer 2014. This activity all but ceased on August 6, 2014, which coincided with a local observatory report of the beginning of an effusive eruption. This almost exact correlation to the observed behavior on the ground marked a significant success in the beginning of this automatic detection algorithm.

Following the success at Stromboli volcano, the algorithm was drastically redeveloped to include a greater number of more robust and comprehensive tests for attempting to identify thermally incandescent activity in a large dataset of images. Initially, all images were analyzed to determine their credibility to reduce the chance of erroneous files affecting the overall results, but primarily to prevent the algorithm from failing, so it would be able to run 24/7 at a volcano observatory. If an image was deemed to be a ‘file error’, ‘image error’, or ‘duplicate image’, they were discarded and the algorithm moved to the subsequent image. Next, a ‘type test’ was performed to determine if the image was a nighttime grayscale image or a color daytime image. As thermal incandescence is rarely observable in daytime images with only significant events captured, only nighttime data was used for this test. All images that successfully make it to this

stage of the algorithm are cropped to a predefined ROI to remove erroneous sources of illumination such as street lights or image time stamps.

Three unique tests were performed on the resulting ROI data. First, a statistical test compares the maximum pixel brightness value to the mean plus six standard deviations to determine if thermally anomalous activity is present. Second, a Canny edge detection aims to identify thermally incandescent activity from the pixel brightness gradient across the image. Third and finally, a Gaussian Mixture Model (GMM) test aims to classify all the pixels into two unique Gaussian distributions (one for the dark background, and one for the brighter activity) by iteratively calculating the probability that each pixel belongs to a specific distribution and assigning these new values (Expectation-Maximization). Each test assigns a one or a zero based upon whether the image passes or not. Images rated a total of zero have no significant activity, whereas images rate higher than zero are detected as having significant activity by at least one test.

This newly improved algorithm was applied to data from 2015 at Shiveluch volcano in Kamchatka, Russia. A permanently installed camera captured at a temporal resolution of fifteen minutes, totaling 26,665 images for the year. The automatic event detection algorithm was applied to this dataset and compared to a manual analysis which was used to represent the absolute truth. An overall accuracy rating of 96.0% was achieved, with a critical success index value of 76.7% when all images successfully identified as a non-event are not included.

This increase in accuracy compared to the preliminary investigation results is significant, but it should be noted the original preliminary testing was performed on pixel brightness data from a TIR camera image and not the raw thermal data. Although the raw temperature values were not used, a TIR camera is much better at identifying volcanic activity compared to an analog NIR camera. However, the success rates seen for the Shiveluch volcano webcam demonstrate the capability of the algorithm to detect volcanic activity.

Further work is needed to reduce the false detection rates and to ensure that as many events as possible are detected as true positives. This algorithm could then be added to a local observatories data stream and its automated alerting of volcanic activity could be correlated to any available seismic, GPS, infrasound, or satellite remote sensing data. The work presented here illustrates how webcam data can be taken from the qualitative to the quantitative world and increase its use in real-time monitoring networks.

Section 4: Automatic Detection Algorithm Testing

4.1 Introduction

Monitoring changes in volcanic activity is not only paramount to understanding the processes that drive active volcanoes, but can also drastically improve the level of risk mitigation. This risk results from the human interaction with a potential volcanic eruption, such as population centers located close to an active volcano or aviation routes that could be detrimentally affected by an ash cloud. Combined with appropriate hazard education and effective communication, identification of precursory signals allows advanced warnings to be disseminated to any potentially affected individuals and could possibly save lives or reduce financial impacts of such a natural disaster (Gregg et al., 2015; Prata and Rose, 2015; Twigg, 2002).

Expensive monitoring equipment is extremely effective when it is utilized, but remote locations, the risk of damage or destruction, and financial restraints can significantly impact the level of monitoring at certain volcanoes. Even relatively inexpensive ground based monitoring methods have total costs of several thousand dollars (Janssen et al., 2002; Harris et al., 2005). As such, many volcanoes around the world are rudimentarily monitored, if at all. Aspinall et al. (2011) determined that of the 441 active volcanoes in 16 developing countries, 384 of them have rudimentary or no monitoring at all, including 65 volcanoes identified as ‘high risk’ to large local populations.

A recent eruption of Bogoslof volcano, Alaska occurred on December 21, 2016 with no warning and was identified by nearby pilots as no ground-based monitoring equipment exists on the island (AVO website, 2016). An eruption last occurred here in 1992 (Miller et al., 1998) and the island is located in the Aleutian island chain 838 miles (1,349 km) from the nearest large city of Anchorage, which means permanently installed monitoring equipment is extremely expensive to deploy and maintain here. Fortunately, no injuries or fatalities occurred resulting from the unexpected eruption, but ideally, monitoring equipment on the ground is needed at remote and potentially dangerous volcanoes like this.

Here, the HAW algorithm developed in Section 3 is improved upon and applied to several other volcano image datasets to determine its effectiveness at detecting pixel brightness resulting from thermally incandescent activity at several different volcanoes around the world. These particular volcanoes were chosen for this section of the investigation as they represent a variety of geolocations and eruptive styles, ranging from the sub-polar regions of the North Pacific to the tropical regions of Central America to the islands of the Mediterranean and volcanologically, from Strombolian to sub-Plinian eruptive styles. Each particular chosen volcano has a webcam capturing images of eruptions at relatively frequent intervals that is freely available online and has allowed the collection of a significant dataset of images spanning several years for each volcano.

4.2 Algorithm Improvements

Resulting from the investigation conducted in Section 3, several areas of improvement were identified that could potentially increase the automatic detection algorithms overall success, which are detailed below. The final algorithm developed here will be referred to as the ‘HAW 2.0 algorithm’.

4.2.1 Grayscale Image Classification

A significant source of error for the original HAW algorithm was the inclusion of images that contained relatively bright backgrounds resulting from twilight conditions, which were included in the ROI and falsely identified as thermally anomalous activity. Under ideal conditions, the ROI would be designed to mask out any of the image background, but as volcanic activity often occurs at the summit and the cameras tend to move slightly throughout the year, the ROI cannot completely remove all background signal. Therefore, a completely different approach was implemented to attempt to reduce the number of false detections.

Originally, all grayscale images were used where inclusion for data analysis was determined by comparing all pixel values in each of the red, green, and blue channels of the data

to each other. As each pixel has an identical value in each channel when captured in the nighttime grayscale mode, when $R=G=B$, the image was designated as grayscale. The installed camera at Shiveluch automatically captures images in this mode when an onboard sensor detects decreased light levels, but the specific value is unknown. This approach to classifying the image type (color or grayscale) would include some images, particularly during the summer months, which contained a dark volcano in the foreground, and a relatively bright horizon in the background, which was often falsely detected as thermally anomalous activity.

Therefore, a new method was developed for the HAW 2.0 algorithm that specifies for target images to be included for the analysis only if they passed a 'twilight test'. For this twilight test, the volcanoes unique latitude and longitude were required to be manually adjusted at the beginning of the algorithm code, which allows for the retrieval of twilight timing data pertaining to natural light levels in that specific area. The date and time of the image filename, and hence time of capture are extracted and passed into an application program interface (API) tool that provides the required information via a JavaScript Object Notation (JSON) file. The date and time relating to three different levels of twilight are extracted from the result and can then be further used in the algorithm. These three levels of twilight are: (1) civil twilight, defined as the time when the sun is less than 6° below the horizon, (2) nautical twilight, defined as between 6° and 12° below the horizon, and (3) astronomical twilight, defined as between 12° and 18° below the horizon (Rozenberg, 1963). For this investigation, all images between the beginning and end of astronomical twilight are included in the grayscale classification for testing, as images in this period will have the lowest level of natural illumination apparent, resulting in a lower percentage of false positive detections in the algorithm. However, this may cause some events to be missed if they occur shortly after sunset, but the benefit of reducing the false alarm rate was deemed greater than the detriment of missing a handful of events.

4.2.2 Duplicate Image Classification

This investigation examines image datasets with a significantly higher temporal resolution than the original HAW algorithm (one minute vs. fifteen minutes for the Shiveluch camera, and four second data for the Stromboli camera), therefore the likelihood of duplicate image files is

dramatically increased. The original algorithm compared file size to determine if an image was duplicated, but this is not a completely robust method as demonstrated in Section 3.3.3.2. There is a possibility that two images could be different but have the same file size. Therefore, the duplicate image classification test was improved upon for the HAW 2.0 algorithm to compare the current file to the previous file using their unique image hash values. This new method is very robust because even if the file names are different, or the image file sizes are the same but are actually two unique images, the image hash will only identify truly duplicated images. As before, if an image is deemed a duplicate, it is deleted and the algorithm moves to the subsequent image.

4.2.3 Average Background Image

This improvement doesn't change the overall outcome of the background image (see Section 3.3.2.6 for a detailed explanation of the derivation of the background image), but improves upon the computational time required to create it. Originally, as the HAW algorithm progressed through each image in the target directory, each image within a specified time window of seven days prior to the current image was added to a list. This list was then examined for specific images within a one hour time window both sides of the target image, and used to create the background image. As only the image filename, date, and time were stored, every time a background image was created, the algorithm would loop through the list, read the data for all appropriate images, and create the averaged background image. While this approach accurately determines the correct background images, the reading of data from each image in the list is relatively inefficient and takes several seconds per target image, slowing the overall performance of the algorithm when performed on a long-term time series dataset. Therefore, for this investigation, instead of re-reading all the images on a specified list, the HAW 2.0 algorithm is designed to take advantage of Python's, specifically NumPy, fast array manipulation abilities. By adding each images pixel values to a three-dimensional NumPy array, the background image can be computed at a fraction of the time compared to the original method (~0.26 seconds), thus increasing the overall computational efficiency of the HAW 2.0 algorithm.

4.2.4 Gaussian Mixture Model Test

Another improvement made to the original HAW algorithm is to strive to improve the overall effectiveness and computational efficiency of the GMM test. This GMM test takes significantly more time and computational resources to complete compared to the other two tests when analyzing long term time series data, and had the lowest accuracy rating of all the tests. The GMM test was performed using 50 iterations, a 5% sample size, and attempted to identify two potential Gaussian distributions, accurate to six decimal places. For this section of the investigation, these GMM test parameters have been altered to attempt to improve the overall accuracy. The GMM test now uses 100 iterations, a 10% data sample size, and assess' the probability of two distinct Gaussian distributions to an accuracy of eight decimal places. These alterations will provide the GMM test with significantly more data points to increase the likelihood of a true positive detection and the only downside of these alterations is the length of computational time required. As previously stated, the GMM test is already very computationally intensive and while these changes may increase the detection capabilities of the test, they will significantly increase the time taken to complete a long time series dataset analysis.

4.2.5 Log-Difference Image

Previously, the GMM test was performed on a log-differenced average background image (Section 3.3.2.6), which helped to not only enhance any potentially thermally incandescent signals but also mitigate the effects of erroneous artifacts in the target image. Therefore, for the HAW 2.0 algorithm, all three tests are performed on a log-differenced average background image. Combined with other improvements to this algorithm such as the twilight test, using the log-differenced image for all tests will reduce the FAR, as the main cause of false alarms (bright erroneous background images) would be differenced out. This alteration would also decrease the number of FN images where small events signals are usually missed, but are amplified by the log-difference method.

4.3 Algorithm Application

To test the applicability of this HAW 2.0 detection algorithm, four volcanoes were selected that have regularly updated webcam data freely available. Images from Shiveluch volcano during 2015 were analyzed again, but with a temporal resolution of one minute. The three other volcanoes selected for analysis were Fuego volcano in Guatemala, Popocatepetl volcano in Mexico, and Stromboli volcano in Italy. These volcanoes were not only chosen because of the availability of high temporal webcam data, but also because of their varying geographical locations (higher latitude vs. tropical locations) and the type of volcanic activity experienced (Strombolian vs. Vulcanian vs. lava dome growth)

4.3.1 Shiveluch Volcano, Russia

4.3.1.1 Image Acquisition

The images from Shiveluch volcano are acquired in an almost identical way to Section 3.3.1 by downloading the images to a local server in Fairbanks from a website (http://volcano.febras.net/archive/latest_Shv2.jpg) maintained by the Russian Institute of Volcanology and Seismology and the KVERT. The camera operates at NIR wavelengths which makes it particularly susceptible to identifying increases in object brightness resulting from increases in temperature. It is located 45 km (27.96 mi) from Shiveluch in the local town of Klyuchi (Figure 1.3), captures images in color during the day, and in grayscale at night when light levels are sufficiently diminished. The image resolution is 800 by 600 pixels and is updated every 60 seconds, resulting in approximately 1,440 images every day. As each image is approximately 40 KB in size, a total of 1 GB of images is collected every month.

4.3.1.2 Results

Webcam data from a three-week period in February 2015 (Figure 4.1A) were analyzed using the HAW 2.0 algorithm to determine its effectiveness and identify periods of heightened volcanic activity. A total of 30,406 images were analyzed during this period with 1,731 images automatically identified as thermally anomalous. It should be noted that due to the high latitude location of Shiveluch volcano, daylight hours during the summer months are significantly longer. Using the astronomical twilight test to ignore images with a bright atmospheric background resulted in a period of approximately three weeks either side of the summer solstice (June 21st) to be ignored due to increased levels of light. Therefore, for this particular volcano, the twilight test was altered to use nautical twilight instead.

The HAW 2.0 algorithm identified a unique period of increased activity beginning on February 17, 2015 that lasted for approximately five days, as identified in green on Figure 4.1A. During this time, several large eruptions occurred, ejecting ash plumes to altitudes ranging from 3,000 to 5,000 m (9,843 to 16,404 ft) ASL, which posed a significant threat to aircraft in the area. A Volcano Observatory Notice for Aviation (VONA) was released by KVERT on February 17 at 21:30 UTC confirming the continual growth of a lava dome, with a viscous lava flow effusion on the northern flank, and an ash plume eruption to 4,000 m (13,123 ft) ASL (KVERT website, 2016a). Figure 4.1B shows the initial eruption at 17:40:48 UTC as captured by the webcam and identified by the automatic detection in the HAW 2.0 algorithm, with an extremely large thermally incandescent explosion followed by an ash plume in subsequent images.

Several thermally anomalous hotspots were identified from National Oceanic and Atmospheric (NOAA) Advanced Very High Resolution Radiometer (AVHRR) sensor data in the several days preceding this event, which is not uncommon for this volcano as it is currently extremely active, and correspond to the smaller data spikes seen on February 15 and 16 in Figure 4.1A. However, a hotspot detected on February 16, 2015 at 04:46 UTC was the last thermally anomalous identification of activity until the large eruption occurred just over 36 hours later, although a potential steam plume was identified on February 17, 2015 at 03:08 UTC. The last satellite image available before the eruption occurred (15:09 UTC, 2.5 hours prior) showed no sign of any thermally anomalous activity and a hotspot wasn't identified, using satellite remote sensing,

until 18:18 UTC, approximately 40 minutes after the event occurred and was identified by the HAW 2.0 algorithm. It should be noted that no significant increase in detected temperature from the satellite remote sensing methods was observed making a large event hard to distinguish, and the data acquisition is limited to the time of satellite overpass, approximately every few hours in this particular region. Furthermore, according to the released VONA, the ash cloud was identified at 20:07 UTC in NOAA satellite 18 AVHRR sensor data. Compared to the initial detection of the eruption through webcam data, by the time it was detected by the satellite sensor, the ash cloud had reached a height of 4,000 m ASL and travelled 234 km (145 miles). By implementing the HAW 2.0 algorithm into real-time operations at this particular volcano, this specific eruptive event would have been detected approximately 40 minutes prior to satellite thermal remote sensing methods and 2.5 hours prior to satellite ash cloud remote sensing methods.

A similar event occurred on February 1, 2015, which was initially detected by the HAW 2.0 algorithm at 14:09:31 UTC. Several thermal anomalies (hotspots) were identified the previous day but the majority of February 1, 2015 was extremely cloudy. The final hotspot detection prior to the eruption occurrence was identified on February 1, 2015 at 06:00 UTC, approximately eight hours prior to the eruption, and the first hotspot was identification after the eruption occurred at 14:47 UTC, approximately 40 minutes after the event. Additionally, the associated ash cloud remained officially undetected until a NOAA 18 satellite AVHRR image became available at 19:50 UTC, a detection time difference of almost six hours (KVERT website, 2016b). By this time, the ash cloud had reached a height of 5,000 m (16,404 ft) ASL and drifted 780 km (485 miles). Although no damage to aircraft or local infrastructure was reported from these eruptions, the risk posed by an undetected ash cloud is significant.

These two examples demonstrate how effective the HAW 2.0 algorithm could be when deployed in a near real-time volcano observatory monitoring environment, as other monitoring methods sometimes fail and are inhibited by their limitations. For this example, the main limitation of satellite remote sensing data is atmospheric cloud obscuring the volcano. Although clouds are also detrimental to the HAW 2.0 algorithm, generally high-level clouds affect satellite remote sensing data whereas only low-level clouds affect the algorithm. Therefore, these two methods of remote sensing could be used in conjunction to support each other, when individual limitations are encountered.

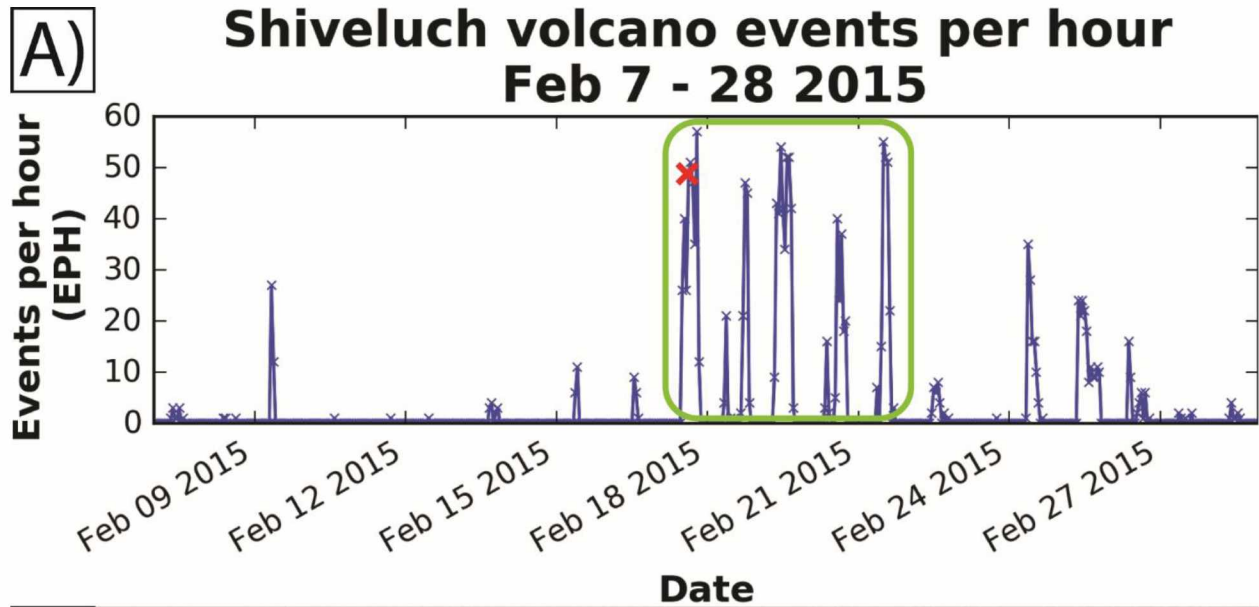


Figure 4.1. Automatic volcanic activity detection algorithm results from Shiveluch volcano. A) A significant period of increased activity is identifiable beginning on February 17 2015, identified in green and B) a webcam image of the initial eruption on February 17 2015 at 17:40:48 UTC, identified by the red cross.

Unfortunately, the HAW 2.0 algorithm is not always completely effective as FP and FN events are sometimes identified. A selection of these events are shown in Figure 4.2. Figure 4.2A shows an example image from February 1, 2015 at 07:25:31 UTC, where the statistical test falsely identified activity, likely resulting from the maximum pixel brightness breaching the specified threshold due to similar pixel brightness values across the ROI creating a low standard deviation. An improvement for future algorithm development could be to implement a minimum standard deviation requirement for this particular test, which would limit the effect of low standard deviations resulting from cloudy data. Figure 4.2B shows an example image from February 24, 2015 at 07:29:42 UTC where the edge detection test falsely identified activity. This is likely the result of moonlit snow on the summit of the volcano appearing brighter compared to the darker surrounding rocks. Figure 4.2C shows an example image from February 5, 2015 at 20:13:00 UTC, where the GMM test falsely identified activity, likely resulting from a significant difference between the dark volcanic summit and the whiter atmospheric cloud in the foreground. These two features led to the GMM test finding two distinct Gaussian distributions in the image, as if there was volcanic activity present. It should be noted the bright twilight background is not a source of error as it was masked from the ROI.

Figure 4.2D shows an example image from February 1, 2015 at 14:18:31 UTC, where both the edge and GMM tests were successful, but the statistical test failed, likely because of a large standard deviation in the data within the ROI because of a large number of saturated pixels, as well as the darker background pixel data. Figure 4.2E shows an example image from February 17, 2015 at 12:37:48 UTC, where both the edge and statistical tests were successful, but the GMM test failed, likely because the level of thermal incandescence was not significant enough to be identified as a second Gaussian distribution and hence the test could not find the required two distributions within 100 iterations. It should be noted that the edge detection test never failed to detect an event that was identified by one of the other tests.

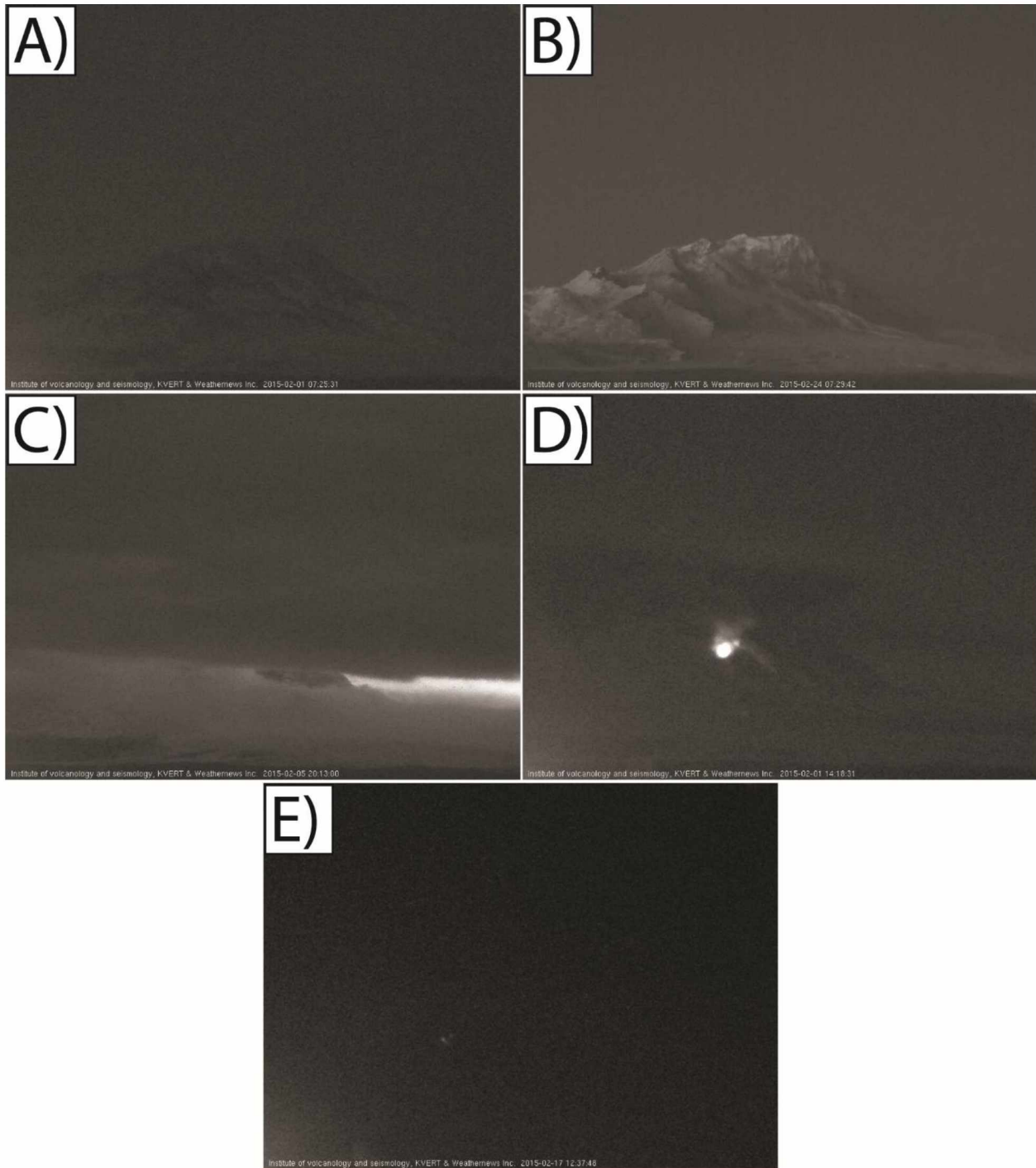


Figure 4.2. Example images where the algorithm falsely detected activity or failed to detect activity at Shiveluch volcano. A) A false detection from the statistical test, B) a false detection from the edge test, C) a false detection from the GMM test, D) a failure to detect activity by the statistical test, and E) a failure to detect activity by the GMM test. In images D) and E), both additional tests successfully identified incandescent activity.

4.3.2 Fuego Volcano, Guatemala

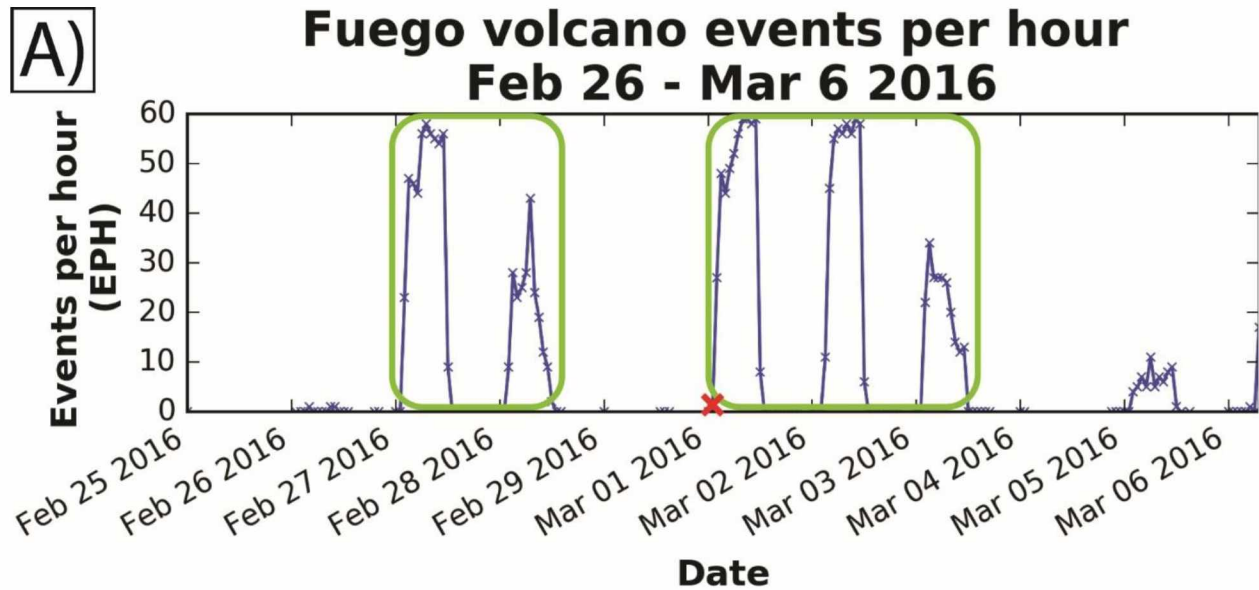
4.3.2.1 Image Acquisition

The images from Fuego volcano are downloaded from a website (http://ovfuego-norte.geo.mtu.edu/raw_jpgs/latestfgo2.jpg) maintained by the Instituto Nacional de Sismología, Vulcanología, Meteorología E Hidrología (INSIVUMEH) and Michigan Technological University (MTU). The camera is located approximately 7 km (4.35 mi) southeast of the summit at La Reunion Golf Resort. Little information is known about the camera itself except that it automatically captures in a grayscale night mode when low levels of light are detected. The image resolution is 640 by 480 pixels and is updated every 60 seconds, resulting in approximately 1,440 images every day. Each image is approximately 25 KB in size, totaling almost 0.5 GB of images collected each month.

4.3.2.2 Results

Webcam data from February to May 2016 were analyzed using the HAW 2.0 algorithm. A total of 87,309 images were analyzed during this period with 15,890 automatically identified as thermally anomalous. To demonstrate the effectiveness of this algorithm, a shorter ten-day period between February 26 and March 6, 2016 was selected to determine the tests success.

During this ten-day period the HAW 2.0 algorithm identified two significant groups of increased activity from February 27 to February 28, and March 1 to March 4, which are identified in green on the left and right respectively (Figure 4.3A). It is likely that these two groups of events are part of one larger cluster and period of activity, but the significant data gaps throughout the webcam time series for this volcano impeded this decision. During this time, significant thermal anomalies were identified at the summit of Fuego volcano leading up to an explosive event on March 1, which was identified by the HAW 2.0 algorithm at 01:10:57 UTC (Figure 4.3B).



<http://ovfuego-norte.geo.mtu.edu/> 2016-03-01 01:10:57



Figure 4.3. Automatic volcanic activity detection algorithm results from Fuego volcano. A) Two significant periods of increased activity are distinguishable beginning on February 27 and March 1, 2015, identified in green and B) a webcam image of the initial explosive eruption on March 1, 2015 at 01:10:57 UTC, identified by the red cross.

This event detection closely supports reports and seismic data provided by INSIVUMEH of a large eruption and the beginning of a paroxysm event in the early hours of March 1, 2016 (Figure 4.4), identified by the large seismic tremor, in black and circled in orange, at approximately 01:09 UTC. Each line represents fifteen-minutes of data, with various colors used to visually segregate the information. The increased seismic signals seen for the following several hours, moving down the figure, is indicative of sustained eruptive activity and supported by the HAW algorithm, with thermally incandescent rockfalls and lava flows detected throughout the night.

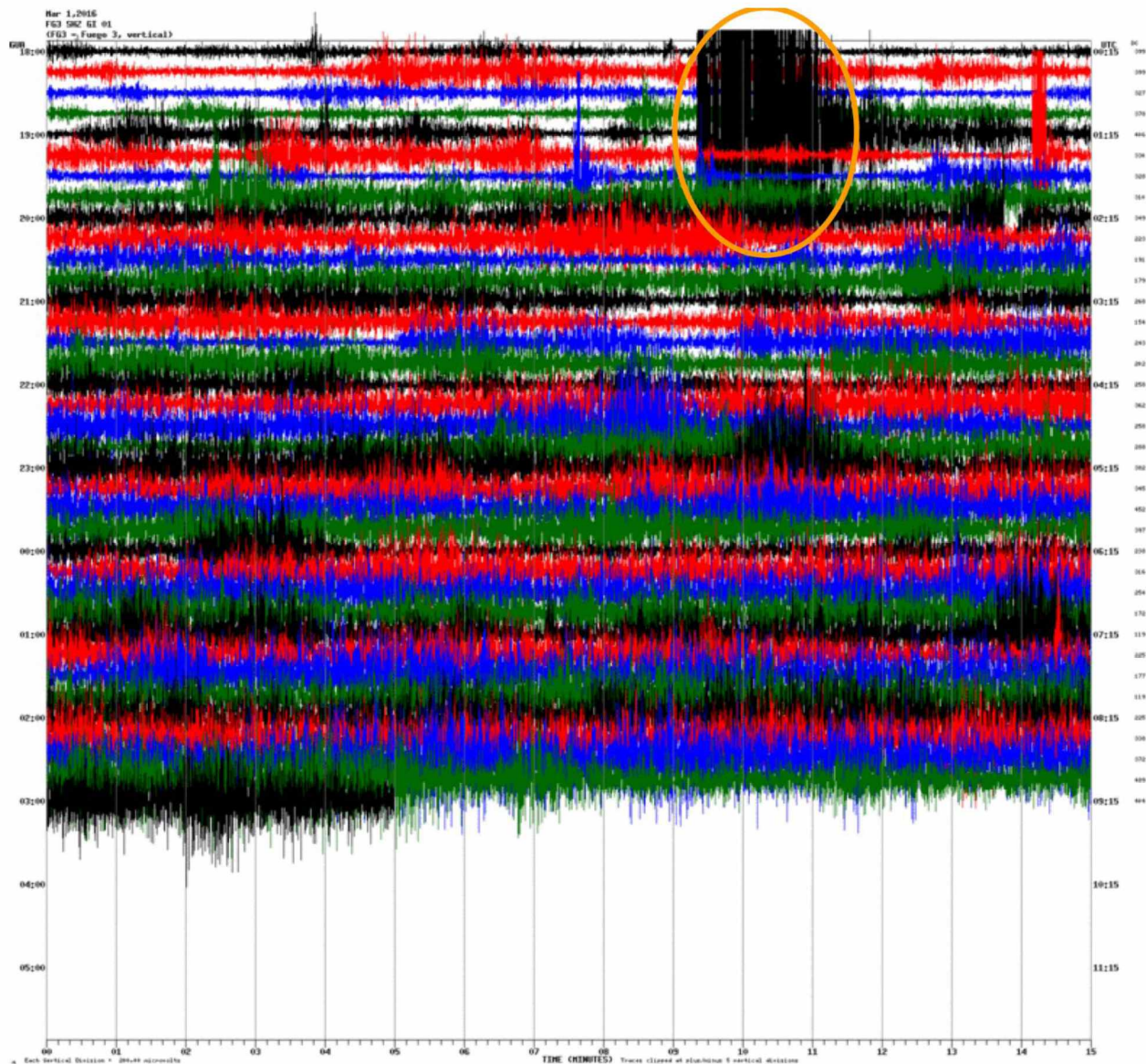


Figure 4.4. A seismogram from Fuego volcano courtesy of INSIVUMEH on March 1, 2015 showing increased activity and initial stages of a new paroxysm event. The initial large eruption is circled in orange, occurring at approximately 01:09 UTC.

As before, the HAW 2.0 algorithm is not always completely effective and suffers some shortcomings in specific areas. Figure 4.5 shows a selection of FP and FN images for Fuego volcano. Figures 4.5A, 4.5B, and 4.5C (March 4, 2016 at 01:05:53, February 5, 2016 at 11:44:26 UTC, and February 28, 2016 at 03:41:09 respectively) demonstrate example data where the statistical, edge, and GMM tests falsely detected activity respectively. These three images appear relatively similar in their red, green, blue channel pixel brightness values, and the primary cause of these FP detections appears to be the bright source of illumination in the lower left corner. This causes significant pixel bleeding across most of the image, where the signal from the illumination source elevates the brightness data of pixels that are in close proximity. In turn, this causes the camera to remain in its daytime color mode and capture images with significantly high image noise. Ideally, this light wouldn't be visible in the image or the camera would be positioned at a slightly different location, but unfortunately, ideal scenarios are not always feasible.

Figure 4.5D shows an example image from March 2, 2016 at 09:35:01, where both the edge and GMM tests were successful, but the statistical approach failed. This likely occurred because of the high number of saturated pixels compared to a dark background, causing a high standard deviation. Therefore, the maximum pixel was insufficient to breach the predefined threshold used in the HAW algorithm (mean plus six standard deviations). Figure 4.5E shows an example image from February 23, 2016 at 08:17:19 UTC, where both the GMM and statistical tests were successful, but the edge detection test failed to identify any activity. This is likely because of the relatively small number of thermally anomalous pixels which were filtered out during the noise reduction stage of the edge test in the HAW 2.0 algorithm. Potentially, a smaller Gaussian kernel or a larger 'closing' morphology array may mitigate this problem. Figure 4.5F shows an example image from March 3, 2016 at 07:09:01 UTC, where both the edge and statistical tests were successful, but the GMM test failed, likely due to the relatively small event not being sufficiently sampled in the test, where only one in ten pixels were used for the Gaussian distribution analysis. Note, this sampling rate was defined for computer efficiency to allow the integration of the HAW 2.0 algorithm at locations without high performance computing capabilities. However, a higher sample rate percentage can easily be implemented and would likely mitigate this problem, but would significantly increase computational resources required.

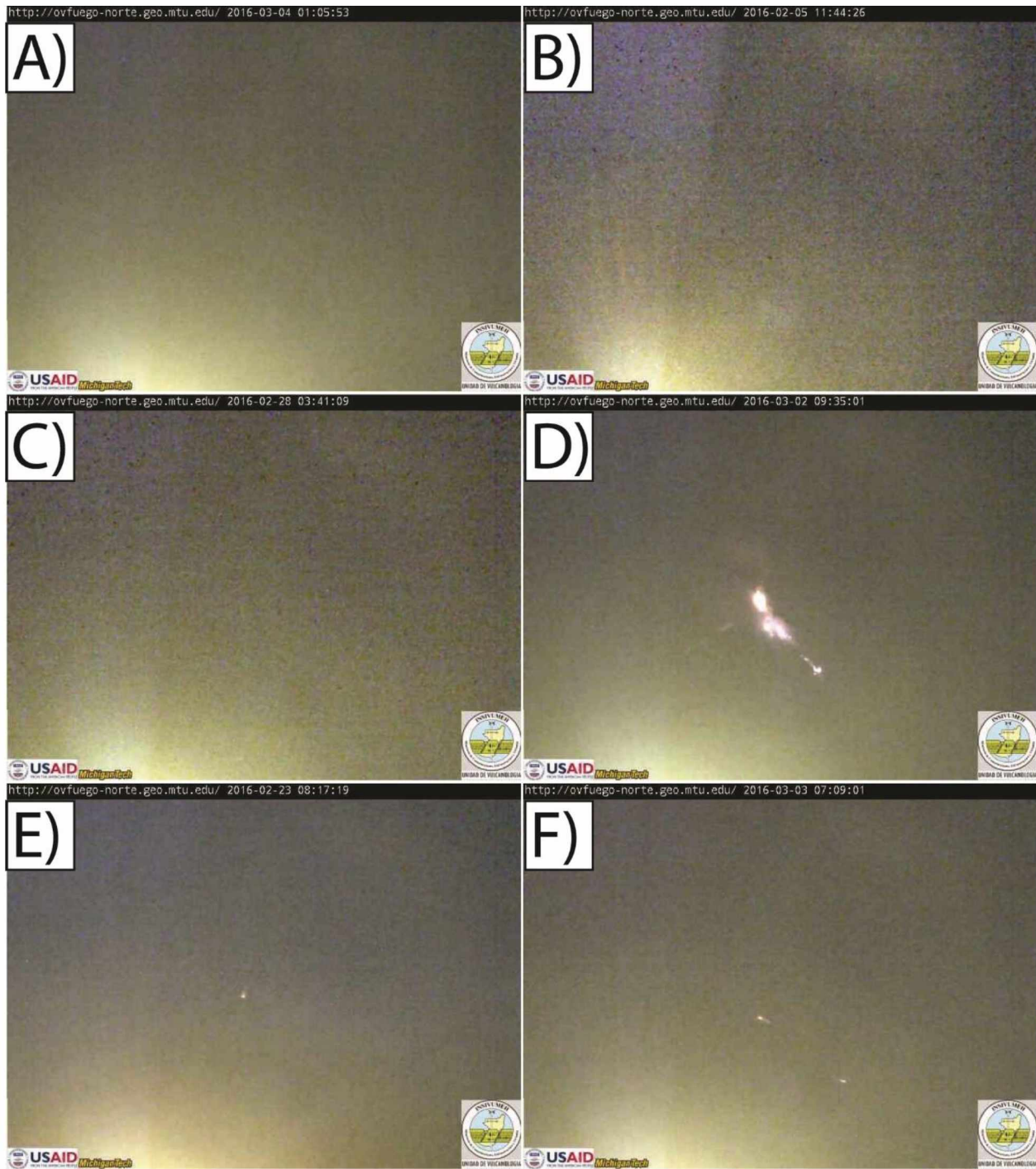


Figure 4.5. Example images where the algorithm falsely detected activity or failed to detect activity at Fuego volcano. A) A false detection from the statistical test, B) a false detection from the edge test, C) a false detection from the GMM test, D) a failure to detect activity by the statistical test, E) a failure to detect activity by the edge test, and F) a failure to detect activity by the GMM test. In images D), E), and F), both additional tests successfully identified incandescent activity.

4.3.3 Popocatepetl Volcano, Mexico

4.3.3.1 Image Acquisition

Images are downloaded from a website maintained by the Centro Nacional de Prevención de Desastres (CENAPRED) (<http://www.cenapred.unam.mx/images/popoTlamacas2.jpg>). The camera is an Axis 214 camera located 5 km (3.11 mi) south of the summit, but doesn't have a direct view of the active vent. It has a 1/4" CCD and automatically captures in a grayscale night mode when low light levels are detected. The image resolution is 704 by 480 pixels and the image is updated every 60 seconds, resulting in approximately 1,440 images every day. Each image is approximately 50 KB in size, totaling approximately 1.5 GB of collected images each month.

4.3.3.2 Results

Webcam data from January to April 2016 were analyzed using the HAW 2.0 algorithm. A total of 159,715 images were analyzed during this period with 14,325 automatically identified as thermally anomalous. To demonstrate the effectiveness of this algorithm, a shorter seven-day period between April 14 and April 20, 2016 was selected to determine the tests success.

During this seven-day period, the HAW algorithm identified a significant group of increased activity from April 15 to April 18, 2016, identified in green (Figure 4.6A). During this time, significant thermal incandescence was apparent at the volcano before a large paroxysm event occurred on April 18, 2016 at approximately 08:00 UTC (03:00 local time). Strombolian activity ejected material up to 1.6 km (1 mile) from the summit, and an ash plume which rose to approximately 8,000 m (26,247 ft) ASL (SIGVP, 2016). This eruption was the strongest event to occur since 1994. Figure 4.6B shows the initial paroxysm eruption occurring at 08:02:30 UTC (03:02:30 local time) as captured by the webcam and automatically identified in the HAW 2.0 algorithm, with a large thermally incandescent explosion accompanied by an ash plume. CENAPRED reported this large paroxysm event gave no warning, but during the 90 minutes to this event, a steady increase in the summit crater brightness is observed, as shown in Figure 4.7.

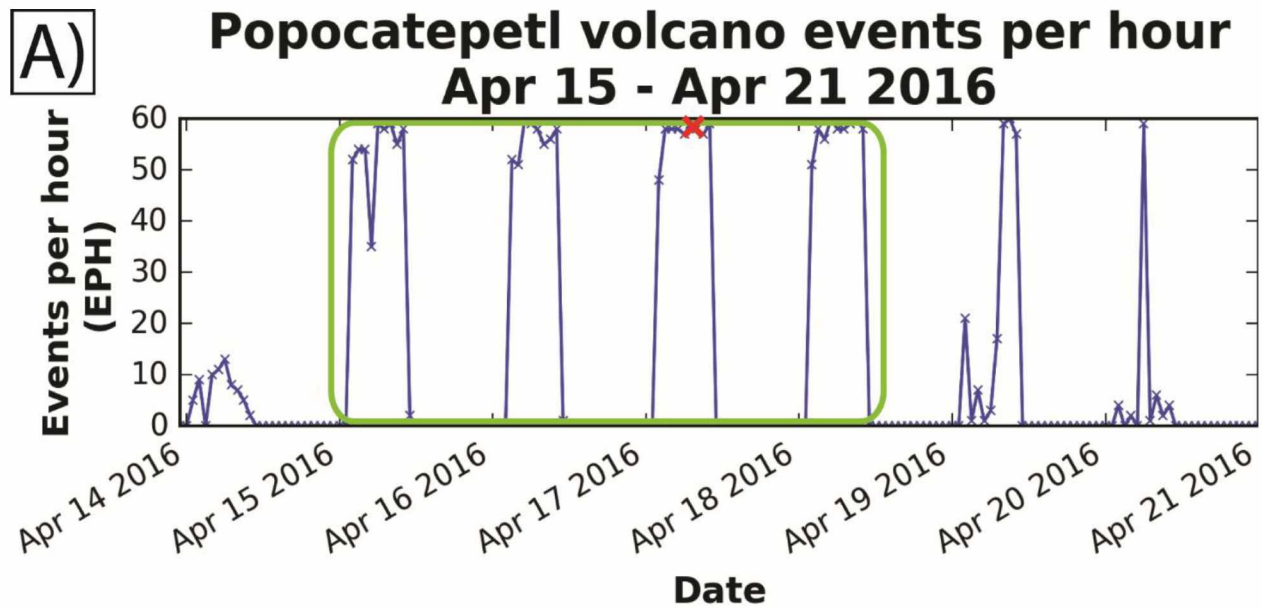


Figure 4.6. Automatic volcanic activity detection algorithm results from Popocatepetl volcano. A) A significant period of increased activity is identifiable beginning on April 15, 2016, identified in green and B) a webcam image of the initial eruption on April 18, 2015 at 08:02:30 UTC (03:02:30 local time), identified by the red cross.

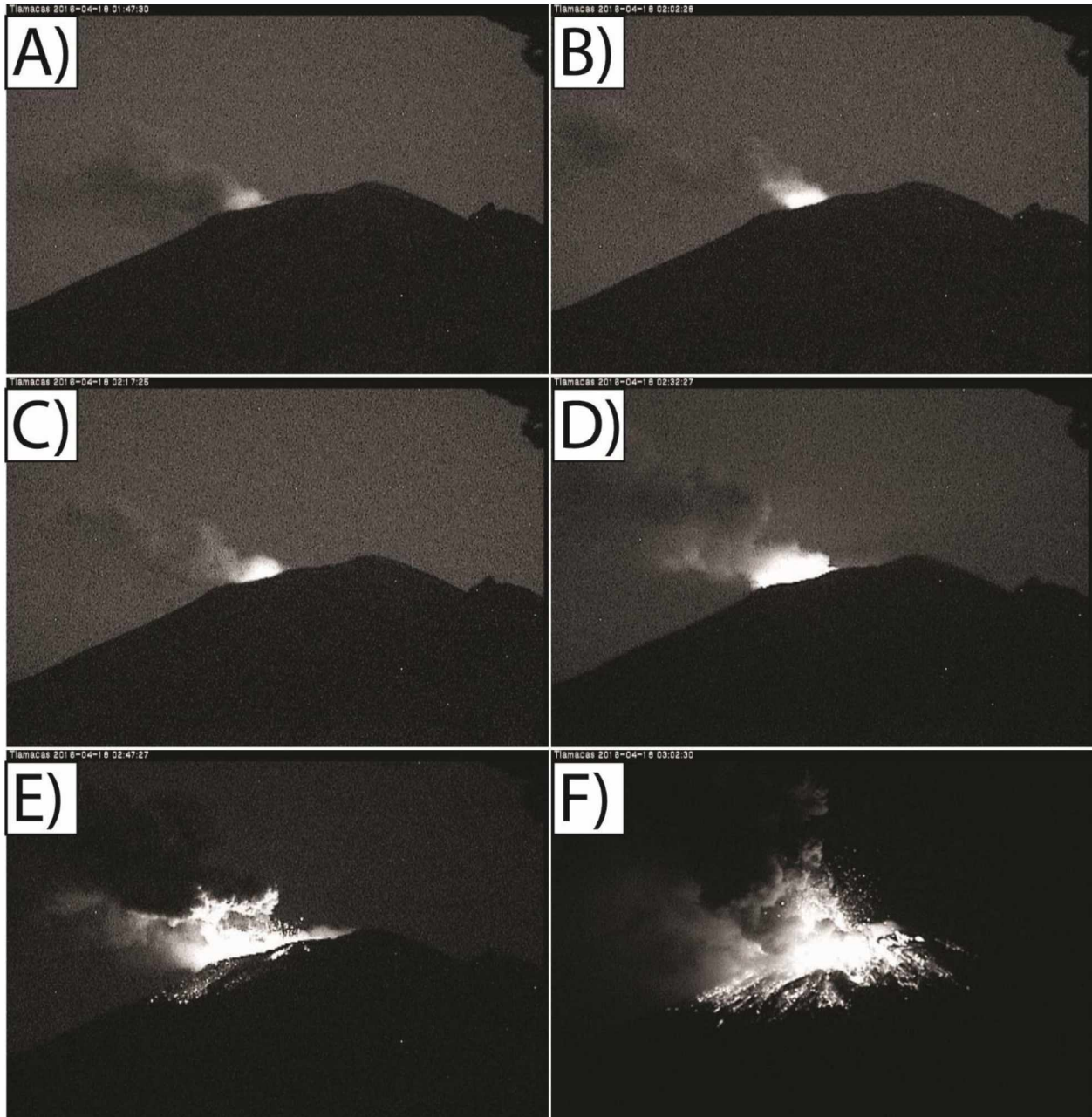


Figure 4.7. Example images demonstrating the increase in thermally incandescent activity in fifteen minute intervals leading up to a large paroxysm event on April 18, 2016. A) 06:47:30 UTC, B) 07:02:26 UTC, C) 07:17:25 UTC, D) 07:32:27 UTC, E) 07:47:27 UTC, and F) 08:02:30 UTC

Although several periods of increased activity were identified throughout the analyzed period, the HAW 2.0 algorithm isn't always completely effective and sometimes fails in specific regions. Figure 4.8 shows a selection of FP and FN images for Popocatepetl volcano. Figures 4.8A, 4.8B, and 4.8C (January 11, 2016 at 01:39:22 UTC, January 14, 2016 at 08:12:23 UTC, and February 3, 2016 at 01:48:46 UTC respectively) demonstrate examples from the webcam data where the statistical, edge, and GMM tests falsely identify activity respectively. The primary cause of FP detections is the significant level of image noise in each image, which for each test, erroneously causes the HAW algorithm to flag activity as being present. This camera performs relatively poorly when capturing in lower light conditions, compared to the other cameras examined so far in this study, and as such has a relatively high FP detection rate. Additional noise reduction techniques such as image smoothing could potentially mitigate this issue, but would dramatically reduce the likelihood of detecting smaller events. Unfortunately, the only true way to improve the image quality at this volcano would be to purchase a new digital webcam to replace the current camera, which increases overall costs slightly, but the data quality would be significantly improved, as well as the ability to detect and inform the local observatory of changes in volcanic activity.

Figure 4.8D shows an example image from January 11, 2016 at 11:12:24 UTC, where both the edge and GMM tests were successful, but the statistical test failed. This is likely due to several erroneous 'dead' pixels in the ROI affecting the mean and standard deviation, resulting in an extremely high threshold that the statistical test cannot breach. Again, the only real way to deal with this issue is to improve the original quality of the image by replacing the camera. Figure 4.8E shows an example image from January 17, 2016 at 05:23:26 UTC, where both the edge and statistical tests were successful, but the GMM test failed. This is likely a combination of severe image noise present in the image, and ambient light reflecting off the snow on the flanks of the volcano. Although there is a relatively significant pixel brightness signal resulting from thermally incandescent activity at the summit, these two contributing factors likely distort the boundary between the two unique Gaussian distributions, causing the GMM test to fail.

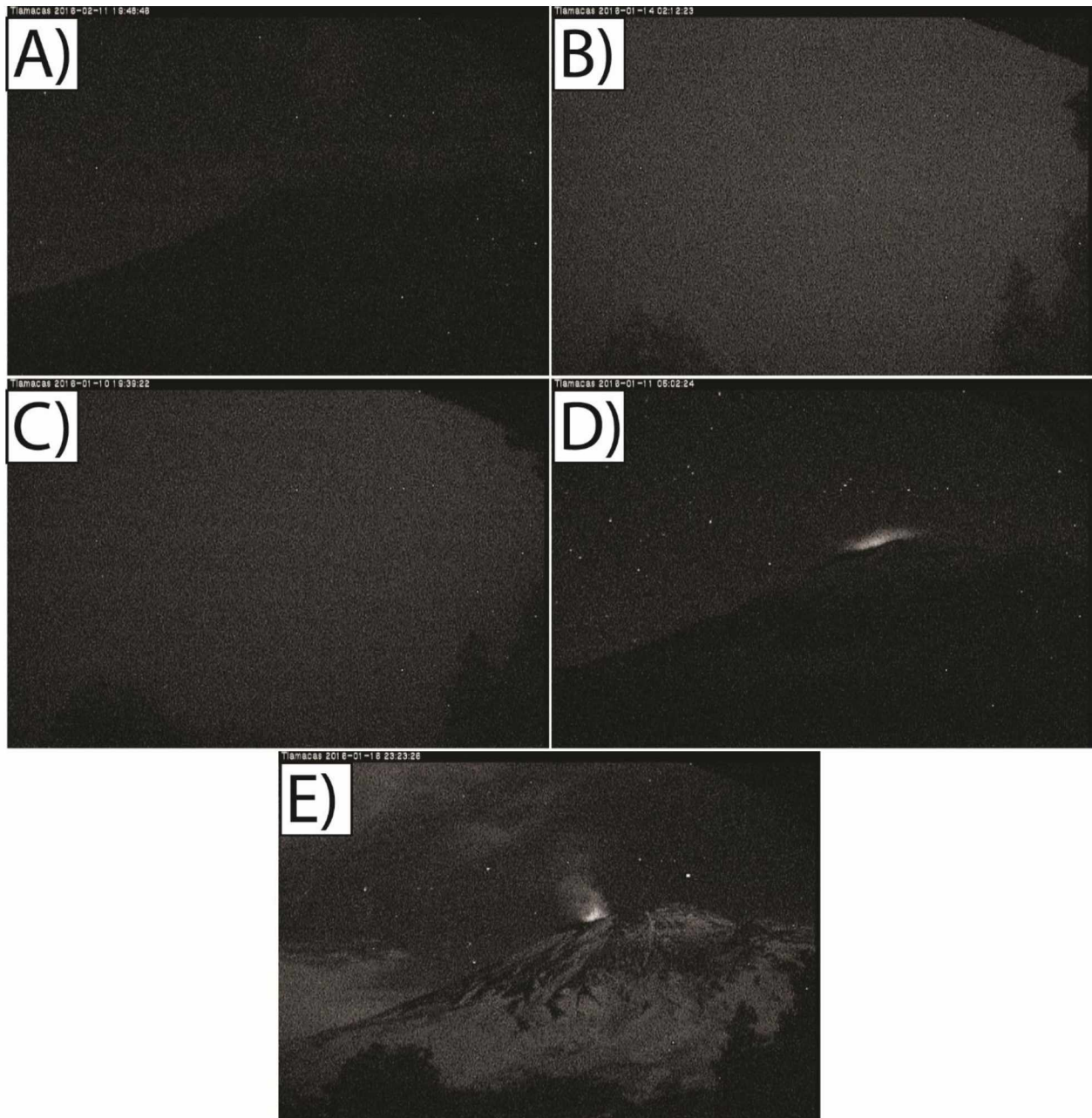


Figure 4.8. Example images where the algorithm falsely detected activity or failed to detect activity at Popocatepetl volcano. A) A false detection from the statistical test, B) a false detection from the edge test, C) a false detection from the GMM test, D) a failure to detect activity by the statistical test, and E) a failure to detect activity by the GMM test. In images D) and E), both additional tests successfully identified incandescent activity.

4.3.4 Stromboli Volcano, Italy

4.3.4.1 Image Acquisition

The webcam data used here is the same dataset as the manual analysis performed in Section 3.2.1. A permanently installed TIR camera on the northern flank of Stromboli, approximately 1,440m (4,724 ft) from the active crater terrace (Figure 1.1), is streamed live online (http://193.206.127.20/Stromboli_Monitoring/DatabasePNG/snap_ROC.jpg) courtesy of the University of Florence, Italy. Only 8-bit .jpg images are available for download, instead of the raw TIR data. The camera has a relatively poor resolution of 352 by 288 pixels and is captured every four seconds, resulting in approximately 20,600 images a day, with an individual file size of ~20 KB, totaling almost 120 GB of data per month.

4.3.4.2 Results

An eight-day period from March 30 to April 7, 2014 was manually analyzed to identify all images containing a unique eruption, which are summarized in Table 4.1. This comparison of the manual, by a trained observer, and automated, from the HAW 2.0 algorithm, analyses follows the work conducted in Section 3, where the capabilities of the original HAW algorithm were evaluated and demonstrated its high accuracy in detecting events and non-active imagery, and low errors in missed events and false detections. As shown in Section 3.3.3.1, several different methods exist to determine the overall success of the HAW 2.0 algorithm detections from a TP, TN, FP, and FN confusion matrix, which are summarized in Table 4.2.

Table 4.1. Confusion matrix results from the comparison of a manual analysis and the HAW algorithm analysis for thermal camera webcam images at Stromboli volcano from March 30 to April 7, 2014.

		Manual analysis	
		Event	No event
HAW algorithm analysis	Event	Type: True Positive (TP) Number of images: 675	Type: False Positive (FP) Number of images: 4,458
	No event	Type: False Negative (FN) Number of images: 647	Type: True Negative (TN) Number of images: 86,960

The overall accuracy (Equation 4.1) of the automatic analysis of eruptive events at Stromboli was 0.945, or 94.50%, which is relatively high, but overwhelmingly controlled by the large number of TN events identified.

$$\frac{TP+TN}{TP+TN+FP+FN} \quad (\text{Equation 4.1})$$

The frequency of detected events (Equation 4.2) was calculated as 0.06, or 6.0%, which means only 6% of the total images contain automatically detected activity. This is expected because of the high temporal resolution these images are collected at.

$$\frac{TP+FP}{TP+TN+FP+FN} \quad (\text{Equation 4.2})$$

The POD (Equation 4.3) was calculated as 0.511, or 51.1%, which determines the likelihood of an event being detected by the HAW 2.0 algorithm. This value demonstrates approximately one in two actual events will be identified. This value is relatively low compared to the HAW algorithm's other POD rates with different cameras, and the cause is related to the

operational wavelength of this camera. During the manual analysis stage of this investigation, all slight changes in pixel brightness were flagged as an event. However, because the camera captures in the thermal infrared, which is then converted to an RGB image, instead of brightness for the NIR cameras, hot steam plumes and potentially erroneous atmospheric clouds ‘heated’ by natural sunlight were likely included. A complete re-analysis of the manually inspected dataset would likely result in a significantly higher probability of detection, as several instances of ‘sun warmed steam plumes’ were identified in images after the HAW algorithm had performed its analysis.

$$\frac{TP}{TP+FN} \tag{Equation 4.3}$$

The FAR (Equation 4.4) was calculated as 0.869, or 86.9%, which is extremely high compared to the HAW 2.0 algorithm’s FAR’s with different cameras, and the cause relates to the wavelength this camera operates at. As the TIR camera images hot material, compared to incandescent material for NIR cameras, when an event occurs, this hot material is usually visible around the crater rim for several images after the event while it cools to background levels. Although every effort was made to ensure the ROI only incorporated the area directly above the active crater terrace, examples of automatically flagged hot material in several images resulting from one unique eruption are apparent. An example of such an image sequence is shown in Figure 4.9, where one individual event is falsely identified as five unique events due to the thermally hot material from the original eruption surrounding the crater. This issue causes the number of FP images to increase, as thermally anomalous activity is identified, but it is not manually identified as a unique event. To mitigate this issue, either the manual analysis could be conducted again to identify individual images where activity is present, compared to unique eruptions, or the automatic analysis could have a temporal filter added to not identify any activity in several subsequent images, after a unique event is detected.

$$\frac{FP}{TP+FP} \tag{Equation 4.4}$$

The CSI (Equation 4.5) was calculated as 0.117, or 11.7%; this value is particularly low as it is affected by both sources of error mentioned above. To mitigate this issue, a temporal filter needs to be incorporated into the HAW 2.0 algorithm for this camera, and the manual analysis needs to be conducted again, and this time identify images that contain thermally incandescent activity, compared to only unique events.

$$\frac{TP}{TP+FP+FN} \quad \text{(Equation 4.5)}$$

Table 4.2. Statistical classification results from the HAW algorithm analysis for thermal camera webcam images at Stromboli volcano from March 30 to April 1, 2014.

	Equation	Value (Percentage)
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$	0.945 (94.5%)
Frequency	$\frac{TP + FP}{TP + TN + FP + FN}$	0.06 (6.0%)
Probability of detection (POD)	$\frac{TP}{TP + FN}$	0.511 (51.1%)
False alarm rate (FAR)	$\frac{FP}{TP + FP}$	0.869 (86.9%)
Critical Success Index (CSI)	$\frac{TP}{TP + FP + FN}$	0.117 (11.7%)

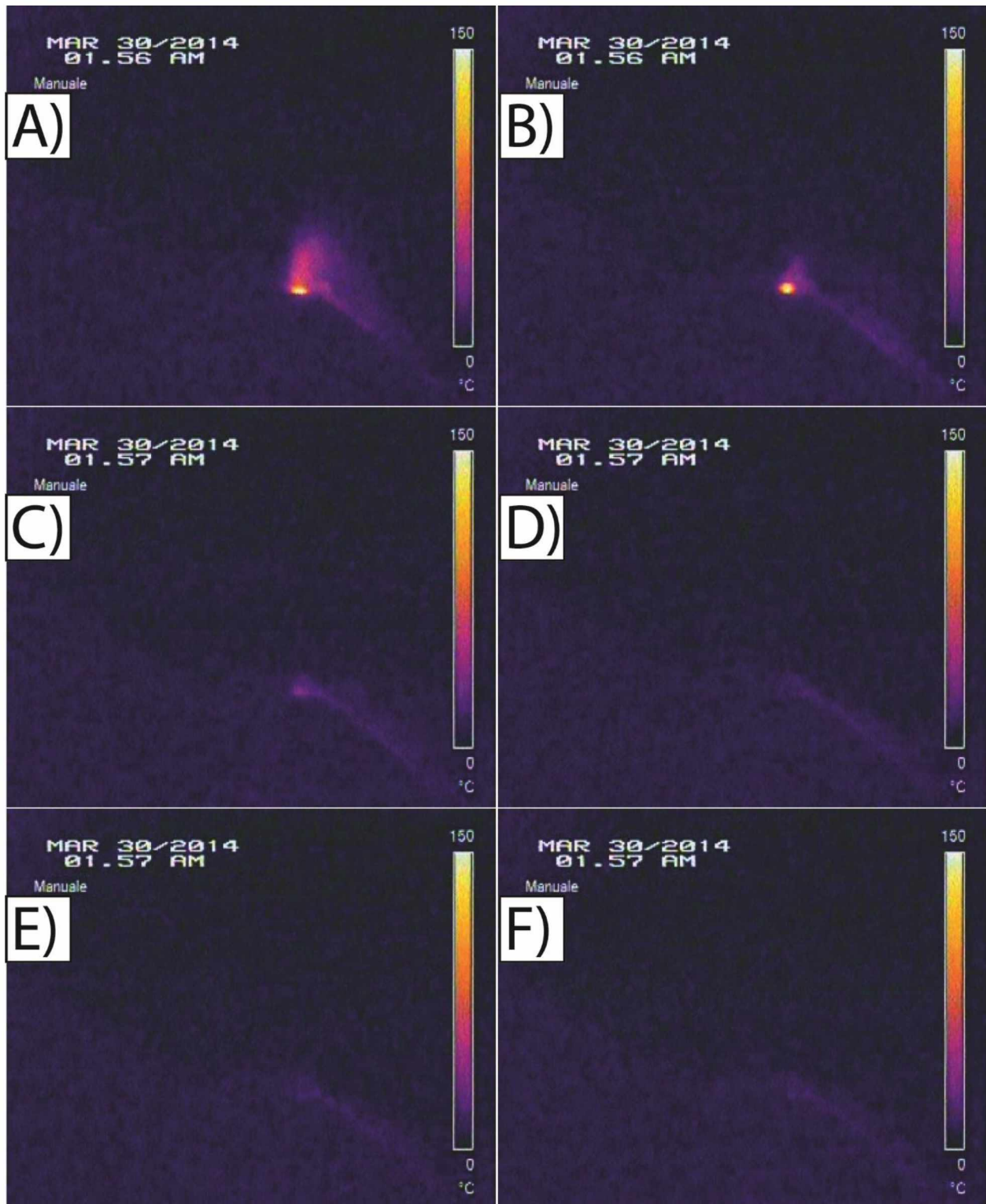


Figure 4.9. Sequence of images showing an individual event at Stromboli volcano that is falsely identified as five unique events by the HAW algorithm on March 30, 2014. A) 01:56:49 UTC, B) 01:56:57 UTC, C) 01:57:05 UTC, D) 01:57:13 UTC, E) 01:57:21 UTC, and F) 01:57:30 UTC

4.4 Conclusions

4.4.1 Volcano Tests

Overall, the HAW 2.0 algorithm performed better than expected when analyzing NIR webcam data from several different volcanoes. At Shiveluch volcano, the automatic tests correctly identified most images within the studied time period (February 2015) and successfully detected a unique period of increased activity beginning on February 17, 2015 (Figure 4.1A). Confirming this period of activity, several hotspots were detected in NOAA AVHRR sensor data the day prior, and shortly after the eruption, and a VONA released by KVERT detailed a large eruption, which ejected an ash plume to 4,000 m (13,123 ft) ASL. Satellite hotspots were not identified in the two images prior to the eruption, and were only detected approximately 40 minutes after the eruption had occurred. Similarly, the ash plume remained undetected for 2.5 hours until a NOAA 18 AVHRR image became available, but was successfully identified at the beginning of the large eruption together with a sustained increased level of activity in the HAW 2.0 algorithm analysis. Another eruption also occurred on February 1, 2015, which ejected an ash plume to 5,000 m (16,404 ft) ASL and drifted 780 km (485 miles) before being detected by a NOAA 18 AVHRR image. Several hotspots were detected the previous day, but February 1 was extremely cloudy and the final detected hotspot prior to the occurrence of the eruption was approximately eight hours prior. Again, this event was successfully identified by the HAW algorithm, combined with a period of sustained increased activity. At this volcano, this automatic testing of thermally incandescent images performed extremely well, and represents a valuable tool which could be used in the near real-time analysis of webcam images to detect increases in volcanic activity.

At Fuego volcano, the HAW 2.0 algorithm successfully identified two unique periods of increased activity resulting from sustained thermal incandescence. The period of increased activity which occurred on March 1, 2016 (Figure 4.3A) was identified by the automatic system at 01:10:57 UTC (Figure 4.3B), which closely matched seismic records provided by INSIVUMEH and signaled the beginning of a large paroxysm event at Fuego volcano. The number of events detected here was relatively high compared to the other volcanoes analyzed as the camera has a direct line of sight to the extremely active vent, and is very sensitive to increases in thermally incandescent

light. However, the number of false detections at Fuego volcano is also relatively high compared to the Shiveluch analysis due to the presence of strong light pollution in the lower left corner of the images. Unfortunately, the only way to mitigate this problem would be to physically move the camera, or turn off the light. While this is a relatively trivial problem, it can cause a significant detrimental effect on the overall accuracy of the HAW 2.0 algorithm. This volcano and camera was actually chosen in part due to this issue, because testing the performance of an automatic detection algorithm under conditions that don't reflect real world situations is relatively meaningless.

At Popocatepetl volcano, the HAW 2.0 algorithm successfully identified a unique period of increased activity resulting from sustained thermal incandescence at the summit from April 17 to April 19, 2016 (Figure 4.6A), and was identified in the webcam data at approximately 08:00 UTC (Figure 4.6B). This large event was identified as the beginning of a paroxysm event at Popocatepetl, where Strombolian activity ejected material up to 1.6 km (1 mile) from the active vent. CENAPRED reported there was little to no warning of the beginning of this event, but a distinct increase in levels of thermal incandescence is visible in Figure 4.7, 90 minutes prior to the eruption, and increasing up to the start of the main paroxysm event. However, although this individual eruption was successfully identified by the HAW 2.0 algorithm, the false detection rate of this camera is also relatively high, resulting from overall image degradation. When capturing at night in the automatically selected grayscale mode, the camera has an extremely high level of image noise, and numerous 'dead' pixels, likely resulting from the age and length of sustained operation. Unfortunately, this detrimentally affects the HAW 2.0 algorithm's rate of success, but represents a real-world example where data feeds aren't always perfect. Image smoothing may lower the levels of false detections, but it would also lower the levels of positive detections as small events would be smoothed out. An additional way to remove these 'dead pixels' would be to develop an additional filter test for the algorithm, which would identify and remove the pixels, which are constantly in the same location. However, the only true way to mitigate this issue and improve the rate of detection for Popocatepetl, would be to replace the camera.

At Stromboli volcano, the algorithm successfully identified 51.1% of all manually identified eruptions (POD), but had an extremely high false alarm rate of 86.9%, resulting in an overall critical success index of only 11.7%. The value is very low compared to the other volcanoes

analyzed in this investigation, and is likely due to the type of images examined and the manual analysis performed by the trained observer. The camera at Stromboli is a TIR camera that images in raw temperature values, but saves the file online as a conventional pixel brightness .jpg image. By examining the data, two main sources of significant error were identified. Firstly, the manual analysis was extremely specific about the number of events identified as activity, which may have erroneously included regular steam plumes from the crater terrace or atmospheric clouds that appear warmer than the background. To mitigate this issue, a second manual analysis of the data is necessary to confirm or dispute the original findings. Secondly and most significantly, the original manual analysis identified the number of unique events occurring from any of the three craters. However, the HAW 2.0 algorithm attempts to identify the number of unique images that contain thermally incandescent activity, which is a slightly different classification approach, compared to the manual analysis. This error is compounded by hot material from an eruption remaining sufficiently hot in several subsequent images to be automatically identified as an event (Worden et al., 2014), whereas the manual analysis would only identify the initial eruption (Figure 4.9). To mitigate this issue and lower the extremely high false detection rate; 1) the manual analysis could be conducted again to identify the number of images that contain activity, instead of the number of unique events, 2) the HAW 2.0 algorithm could have additional temporal tests developed to stop images being flagged as a unique event if one of the previous five images was already flagged, and 3) a combination of the two as the number of unique events is useful for automated alerting, while the number of unique images can be useful to classify the type, size, and intensity of the activity. This would significantly reduce the false alarm rate as it is not uncommon for hot material to remain significantly above background temperatures for at least five images after a unique event. In addition, this temporal filter approach would be useful for real-time analysis and volcano monitoring as it would flag only new eruptive events or unique events, rather than alerting for all images where incandescence material is detected. This approach has been used in remote sensing algorithms for volcano monitoring in the North Pacific (Dehn and Harris, 2015) and in tasking National Aeronautics and Space Administration (NASA) higher spatial resolution data (Duda et al., 2009).

4.4.2 Image Tests

For the HAW 2.0 algorithm three unique tests (Edge, GMM, and Statistical) are performed on every image to attempt to identify thermally incandescent activity resulting from increased pixel brightness. The statistical test uses the least computational power and is the fastest performing test of the three, with the highest overall accuracy rating and the second lowest false detection rate. Due to the completion time and accuracy levels, this test is determined to be the best performing test in the HAW 2.0 algorithm. The test with the second-best performance is the edge detection method. It has a similar completion time to the statistical test, but has a slightly lower accuracy rating and a higher missed event rate. However, the edge detection test has the lowest rate of false detections for all three tests. Finally, the GMM test was deemed as the lowest performing test. It still performed relatively well, but the combination of a lower accuracy rating, higher missed event rate, and lengthy completion time resulted in a relatively low performance compared to the other two tests. The accuracy levels are not significantly lower than the other two tests, but the completion time is extremely detrimental to the overall algorithm, which could be reduced by approximately 90% by not including this test. This is a significant time saving when analyzing a large dataset of images as shown here, but it should not be an issue when analyzing new images in real time if the HAW 2.0 algorithm can finish running before a new image is uploaded.

4.4.3 Future Work

Future work for the HAW algorithm would revolve around improving the overall accuracy rating, lowering the false detection and missed event levels, and performing the analysis of images in near real-time to then implement its detection methods into an operating volcano observatory. To reduce the major restriction of computational power requirements, the local camera capturing the images could upload the files to a cloud computing system, such as those provided by Amazon (Elastic Computing, EC2) or Microsoft (Azure) to name two examples, where the algorithm is constantly running and analyzing new images, which removes the need for expensive processing

equipment in the local area. However, this would require the local organization to maintain a cloud computing account and support this technique long term.

Alternatively, as the algorithm performance is limited by the computer's central processing unit (CPU) speeds, instead of the random-access memory (RAM) speeds, a migration towards graphics processing unit (GPU) analysis would dramatically increase the performance levels (Gregg and Hazelwood, 2011). GPU's can process information at often hundreds or thousands of times faster than a CPU, which would dramatically decrease the performance time for this algorithm. However, GPU processing is not currently very common and code adaptations would likely be necessary. In addition, the GMM test requires access to the brightness values for all the pixels to assign them to a specific Gaussian distribution, so it would not be possible to divide the image into sectors and then parallel process the different tests in the HAW algorithm.

With a marked increase in performance, the individual tests could be altered to improve the overall accuracy of the HAW algorithm. For example, instead of analyzing a ROI, which is relatively small to save computational processing time, the entire image could be analyzed (minus areas of known false detections such as street lights or image timestamps) which would likely improve the overall accuracy as the image being analyzed is substantially more representative of the actual real life situation. Furthermore, the GMM test specifically could be improved by using a sample rate of 100% instead of just 10%, and further increasing the number of iterations. Both of these alterations would likely improve the accuracy ratings of the GMM test, but would only be possible for near real-time monitoring of an active volcano if a dramatic increase in computer performance is achieved.

Additionally, the implementation of a moving ROI test would greatly benefit the overall success of the algorithm. Many falsely detected events from all volcanoes analyzed here result from erroneous bright atmospheric background being included in the ROI when the camera moves slightly throughout the year. This is particular apparent at Popocatepetl volcano, where the camera moves drastically every month or so. By implementing an automatic test to identify the outline of the volcano from a clear image within the last few days, a moving ROI could be used instead of the current static one.

It should also be noted that the HAW 2.0 algorithm assigns the same outcome to events with one bright pixel as an event with thousands of saturated pixels. Therefore, it is difficult to

determine the difference between a small dome growth glow and a large thermally incandescent eruption. Currently, a rudimentary way of identifying larger activity is to determine the number of HAW 2.0 tests that flagged a particular image. Generally, when all three tests flag, the thermal incandescence in the image is significantly stronger or larger than if only one test flags it. However, a future improvement could incorporate additional tests to further classify images into eruption strength and size categories, based up pixel brightness, the number of saturated pixels, and their image spatial distribution.

Section 5: Conclusions

5.1 Summary of Findings

Throughout this investigation, ground based remote sensing of volcanic activity using NIR cameras to detect increases in thermal incandescence has been shown to be a successful method to effectively identify volcanic eruptions. The method proposed here focuses on developing a relatively robust technique of volcanic eruption monitoring, while maintaining a high level of realism relating to financial constraints and field deployment ability. For example, other methods of volcano monitoring such as satellite remote sensing (Webley et al., 2009a; Meyer et al., 2015; Thomas and Prata, 2011), seismic (Lyons and Waite, 2011), GPS and tiltmeters (Larson et al., 2010), and infrasound (Arámbula-Mendoza et al., 2013) can potentially provide a more accurate event detection system to identify eruption precursors (seismic, GPS), or capture data that encompasses a larger spatial extent (space borne data), than is possible from ground based remote sensing, but the financial costs of such systems are vast compared to the method proposed here.

Initially, this investigation focused upon attempting to classify volcanic eruption temperatures by directly comparing incandescent targets to their corresponding pixel brightness values. This approach was initially conducted throughout the extensive laboratory spectral response experiments (Section 2.4) and attempted during a field campaign to Stromboli volcano, Italy (Section 2.5). Although the method showed promise with a clear correlation between true target temperature and image pixel brightness, several issues were encountered that limited the methods applicability. For example, during the Stromboli field campaign tests, comparisons between eruptive object temperatures and pixel brightness were possible, but due to the extreme temperatures at an active volcano, the NIR cameras were almost always saturated. With a saturated pixel, the true correlation between brightness values and object temperature remained unknown. It was possible to determine that the imaged eruption must be above a specific temperature, the known temperature that causes the camera to saturate at that specific distance, but the absolute temperature of the eruption remained elusive. Therefore, this investigation pivoted towards detecting volcanic events by comparing pixel brightness values relative to previous captured

images at a specified volcano, instead of an absolute comparison to additional TIR cameras or thermally incandescent targets of a known temperature.

The development and application of the HAW algorithm, original in Section 3, version 2.0 in Section 4, to four volcanoes around the world (Shiveluch, Russia; Stromboli, Italy; Fuego, Guatemala; Popocatepetl, Mexico), demonstrated the effectiveness of monitoring volcanoes using inexpensive ground based NIR cameras. Three unique tests (statistical, edge, and GMM) were implemented into a detection algorithm that analyzed large temporal datasets at each volcano and attempted to identify periods of increased activity. The developed HAW algorithm was very successful when compared to manually analyzed webcam data from the same period and successfully identified several periods of heightened eruptive activity which corresponded with alternative datasets or monitoring reports by local volcano observatories.

5.2 Method Development

As mentioned above, the initial stages of this investigation focused primarily on identifying eruption temperatures by comparing object temperatures with collocated pixel brightness values. However, after finding this approach to be severely limited by the capabilities of the NIR cameras, see analysis in Section 2, a change in direction toward creating an algorithm to identify increases in volcanic activity in near real-time for operational purposes was made.

The original HAW algorithm, see Section 3, was developed to attempt to identify increases in volcanic activity, by detecting changes in image pixel brightness values relating from thermally incandescent material at the volcano. Essentially, an increase in volcanic activity will cause an increase in temperature, which in turn will cause an increase in pixel brightness and can be identified to alert an operational system of the hazard potential. To achieve this, three uniquely individual tests were developed that were designed to isolate and identify sources of incandescent activity and assign a test score accordingly:

- 1) Statistical calculations (Statistical test) - The initial test of the HAW algorithm calculates the mean and standard deviation of a specific region of interest (ROI) within the webcam image and determines if an event has occurred by testing to see if the maximum pixel value in the ROI is

greater than the mean plus a specified number of standard deviations of the ROI. If so, the test concludes the target image contains thermally incandescent activity, and assigns a score of one.

2) Canny edge detection (Edge test) - The second test in the algorithm is a Canny edge detection (Figure 3.6), which first smooths the target image with a Gaussian filter to remove any potentially erroneous source of image noise, before analyzing the gradient between neighboring pixel values to identify if there are any unique edges in the image. If an edge is detected, the system will flag the image as thermally active and assign a score of one.

3) Gaussian Mixture Model (GMM test) - The third and final test in the algorithm is a Gaussian Mixture Model test, which applies a probabilistic approach to identifying unique clusters of pixels within the target image. The test is designed to search for two unique Gaussian distributions within the image, with one belonging to the darker background pixels in the image, and the second belonging to the brighter pixels resulting from thermally incandescent activity. Initially, all pixels in the image are randomly assigned to either Gaussian distribution, known as the initial Expectation step. The first Maximization step, then calculates the probability that each pixel belongs to the distribution it was initially assigned to, using a variety of statistical parameters including distance from the center of the Gaussian distribution and standard deviation. These new probabilities are then assigned in the second Expectation step, before being re-evaluated in the second Maximization step. This iterative approach continues for a predefined number of iterations until this test can determine an outcome (Figure 3.7). If the likelihood that each cluster of pixels belongs to an appropriate Gaussian distribution, true to a specified number of decimal places, then the test flags the image as containing thermally incandescent activity, and assigns a score of one.

All three tests are conducted on every image within the specified dataset time window, with an overall outcome of an image test score value ranging from zero to three. A zero score shows that none of the tests successfully identified any level of activity and as such, the image likely contains only dark background pixels with no volcanic activity present. A positive score of one or more, shows that a number of tests identified the image as containing some level of thermally incandescent activity. A score of one represents an image only detected by one of the tests, whereas an image classified as a three, was detected by all tests. This scoring system of zero to three can be used as a proxy for event classification and increases in activity levels. For example, a single image with a score of three may be thermally incandescent, but is not likely to represent a

significant increase in activity. However, a sustained increase in test score values from zero to one, to two, to three, could indicate a heightened level of activity and the potential for a large eruptive event in the near future.

5.3 Method Application

The HAW algorithm (original and version 2.0) was applied to five individual datasets from four different volcanoes around the world (Shiveluch, Stromboli, Fuego, and Popocatepetl), see Sections 3 and 4 for a detailed analysis. Each volcano was chosen because of the varied geolocation and eruptive style that they represent, ranging from the sub-polar regions of the North Pacific to the tropical regions of Central America to the islands of the Mediterranean and volcanologically, from Strombolian to sub-Plinian eruptive styles. The algorithm was initially applied to a fifteen-minute temporal resolution image dataset from 2015 at Shiveluch volcano, Section 3. A total of 29,648 images were manually analyzed by a trained observer and automatically by the original HAW algorithm to determine the overall level of accuracy. The camera at Shiveluch automatically changes into a grayscale nighttime capture mode when low levels of light are detected, which meant 11,699 grayscale images were analyzed for this stage of the investigation. An overall success accuracy of 96.0% was achieved by the original HAW algorithm, which is defined as the percentage of successfully identified events and non-events which matched the manual analysis results (Equation 3.2). The original algorithm also had a POD rate of 84.0%, which is defined as the number of successfully detected events divided by the sum of the successfully detected events and the missed events (Equation 3.4). If an algorithm is performing perfectly, it would achieve a POD score of 1, or 100%. Overall, the algorithm has a CSI value of 76.7% which is defined as the number of successfully detected events divided by the sum of the successfully detected events, the false alarms, and the missed events (Equation 3.6). Generally, the images that were falsely detected were flagged because of either image noise or cloudy data detrimentally affecting the image pixel values, and the events that were missed were generally extremely small events amounting to a small glow at the summit of the volcano. It should be noted that all sizeable events where an actual eruption occurred were successfully identified by the original HAW algorithm for this dataset.

The algorithm was then subsequently upgraded to version 2.0, with some new tests and evaluation criteria to improve accuracy and reduce processing time and applied to four additional datasets (Section 4). However, no manual analysis has yet been performed on these datasets so the confusion matrix performance of these tests is quantitatively unknown. Nevertheless, by examining several specific time periods identified by significant sustained increases in activity, several examples were identified which demonstrated the effectiveness of the HAW 2.0 algorithm identifying increases in activity shortly prior to a large eruptive event.

5.3.1 Shiveluch Volcano

For Shiveluch volcano, data from February 2015 were analyzed again, but with a significantly higher temporal resolution of one image every minute. From this, a unique cluster of increased activity was identified (Figure 4.1A), which corresponded to the identification of thermally anomalous hotspots in NOAA AVHRR data and the release of a VONA by KVERT. Several hotspots were identified in the several days prior to this large eruption, but the most recent detection prior to the event was February 16, 2015 at 04:46 UTC, 36 hours prior to the eruption, with an image 2.5 prior to the eruption showing no sign of increased thermal activity. Additionally, the VONA detailed a 4,000 m (13,123 ft) high ash plume that was identified at 20:07 UTC by NOAA satellite 18 AVHRR sensor data. However, from the HAW 2.0 algorithm analysis of the ground based NIR remote sensing data, the eruption was initially identified at 17:40:48 UTC (Figure 4.1B), approximately 40 minutes prior to the initial hotspot detection and a full 2.5 hours prior to the satellite data. However, it should be noted that the hotspot identification did not show a significant increase in thermal activity and was very similar to all other hotspots identified at this extremely active volcano. A similar example was identified on February 1, 2015 where the HAW algorithm initially detected a significant eruption at 14:09:31 UTC, but any precursory activity for this event was not identified by satellite remote sensing methods due to cloudy conditions, with the last detection occurring approximately eight hours prior to the eruption. Additionally, the VONA released by KVERT at 19:50 UTC detailed a 5,000 m (16,404 ft) high ash plume which had drifted 780 km (485 miles). These two examples from only one month of data demonstrate how effective the HAW 2.0 algorithm could be if incorporated into every day volcano monitoring.

5.3.2 Fuego Volcano

At Fuego volcano, data from February to May 2016 were analyzed, with a specific ten-day period from February 26 to March 6, 2016 selected to evaluate the algorithms success. Two significant groups of increased activity were identified from February 27 to February 28, and March 1 to March 4, 2016 (Figure 4.3A). During this time, significant thermally incandescent anomalies were detected leading up to an explosive event on March 1, 2016, identified by the HAW 2.0 algorithm at 01:10:57 UTC (Figure 4.3B). This identification is supported by a seismogram provided by INSIVUMEH showing a large eruption occurring at approximately 01:09 UTC. The sustained seismic signal seen for the following hours also supports the sustained eruptive activity identified by the HAW 2.0 algorithm.

5.3.3 Popocatepetl Volcano

At Popocatepetl volcano, data from January to April were analyzed with a specific focus between April 14 and April 20, 2016. During this seven-day window, CENAPRED reported an eruption that ejected ash to 8,000 m (26,247 ft) ASL and thermally incandescent material up to 1.6 km (1 mile) from the vent itself. This eruption confirms an increase in thermally incandescent activity detected by the HAW algorithm (Figure 4.6A), which culminated into a large explosion at 08:02:30 UTC (Figure 4.6B). CENAPRED reported that the large paroxysm event gave little to no warning, but the HAW 2.0 algorithm successfully identified a sustained increase in image pixel brightness levels in the 90 minutes leading up to the main eruptive phase (Figure 4.7).

5.3.4 Stromboli Volcano

At Stromboli volcano, an eight-day period from March 30 to April 7, 2014 was analyzed, where the images were captured with a four second temporal resolution. This same period was also manually analyzed by a trained observer, see Section 3, allowing a direct comparison between

the two methods to determine the overall accuracy. The main difference with this dataset, apart from the increased temporal resolution, is the type of images that are captured. The camera at Stromboli is a permanently installed TIR camera that streams images online in a conventional .jpg format, instead of the original raw temperature data format. Therefore, the same pixel image brightness tests are performed on the temperature color scaled images instead. For these data, the HAW 2.0 algorithm had an overall accuracy of 94.5%, a POD value of 51.1%, a FAR of 86.9%, and a CSI of 11.7% (Table 4.2). These numbers are particularly disappointing and demonstrate that the HAW 2.0 algorithm in its current form did not perform very well on the high temporal TIR image data from Stromboli volcano. The high false detection and low success rate are primarily due to a high number of false events detected by the algorithm, which identified each individual image with activity, compared to the manual analysis that identified each unique event. This is significant because one event can eject material that remains hot and thermally incandescent for several seconds to minutes (Worden et al., 2014) and is detected by the HAW 2.0 algorithm as multiple events (Figure 4.9). To improve upon these values, the manual analysis will need to be conducted again, to exactly match the specific criteria used by the HAW 2.0 algorithm, or the algorithm could be updated to contain a temporal filter which, when applied, can determine the difference between unique events and unique images with elevated activity levels.

Overall, the HAW algorithm, both in its original form and version 2.0, was very successful when analyzing the four NIR camera datasets, with an almost 96% success rate compared to the manually analyzed data, and successful identification of significant eruptions when compared to alternate volcano monitoring methods, sometimes even identifying activity several hours in advance. However, applying the HAW 2.0 algorithm to thermal images with a high refresh rate needs further work to accurately determine the success rate at Stromboli volcano. This method of identifying increases in volcanic activity by relatively comparing image pixel brightness to identify thermally incandescent activity is very effective and would prove to be a useful additional tool at a volcano observatory to improve their monitoring capabilities.

5.4 Future Work

Performing another manual analysis of the Stromboli TIR camera data would be a main priority, which would allow a direct comparison of the HAW 2.0 algorithm to a manually inspected dataset to determine the algorithm's success rate in its current form. Additionally, performing a manual analysis of the images from the other three volcanoes would also allow flaws in the code to be identified and further improvements applied to increase the overall success rate of the HAW 2.0 algorithm. Any updates to the algorithm would lead to version 2.1 or 3.0, depending on the scale and success of the implemented changes.

Improving the algorithm efficiency when analyzing the images would also be a top priority for any future investigations. Primarily, this is targeted at the GMM test which requires high levels of computational performance. By increasing the test's performance, a higher level of accuracy can be obtained using an increased number of iterations and a higher sample rate that is used by the algorithm to analyze each image. However, the issue of computational limitations is only apparent when attempting to analyze large datasets all at once. If the HAW 2.0 algorithm, and any updated versions, was incorporated into real time monitoring, analyzing a new image every minute would easily be possible even at the current levels of computational efficiency.

Related to the previous point, the time taken for both versions of the HAW algorithm to execute on large datasets can also be improved by altering the system on which the analysis is performed. For example, the code is currently analyzed on a Linux server with 16 CPUs and 32 GB of RAM, but the code currently only runs on one CPU. Therefore, either a code adaptation to allow an implementation of parallel processing on more than one CPU, or the migration to GPU processing would dramatically increase the algorithm's execution speed (Gregg and Hazelwood, 2011). Modern GPUs are significantly faster at processing information than CPUs and could provide an extremely fast method of analyzing high volumes of data in near real-time.

Further improving the accuracy of the HAW algorithm is always a high priority, and the main way to achieve this is to compare the results to other forms of volcano monitoring datasets. For example, a direct comparison between the levels of thermally incandescent activity detected by the algorithm to the strength of seismic energy from nearby seismic stations, could provide an additional level of 'truthing'. Additionally, high temporal resolution satellite remote sensing data,

such as the Japanese Meteorological Agency's Himawari satellite and Advanced Himawari Imager (AHI), or in the future, the NOAA / NASA Geostationary Operational Environmental Satellite (GOES) R series satellite and Advanced Band Imager (ABI), could provide an additional data stream to compare the algorithm's effectiveness to detect volcano activity and improve its overall accuracy. Here, temporal correlations would be used to evaluate the timings of the events seen from the webcam data as compared to the other data streams. Again, statistical classification tests, such as the probability of detection, false alarm rate, and critical success index could be used as well as additional tests developed for the inter comparison of time series datasets. The different temporal frequencies of the data would need to be taken into consideration as this work progressed.

In its current form, the HAW 2.0 algorithm can only analyze nighttime images to identify thermal incandescence, as this type of activity is rarely identifiable during the day due to the high levels of natural illumination. However, an additional algorithm could be implemented to attempt to identify daytime explosive eruptions which are primarily ash rich, by performing almost the exact opposite of the current test, and daytime effusive eruptions such as lava flows on the flanks of a volcano, that may require a slightly adapted version of the HAW algorithm. For example, in its current form, three tests identify thermally incandescent bright pixels compared to a dark nighttime background. However, for daytime imagery, the method could be reversed to detect a dark cluster of pixels' representative of an ash plume compared to a bright atmospheric background for the explosive eruptions, and on the volcanic flanks it would be a bright flow or incandescent material on a darker landscape. Additionally, there would be data in the all three RGB channels, which could be examined to classify activity type, size, and intensity directly from webcam data.

Finally, the overall implementation of the HAW algorithm into a volcano observatory and their active monitoring techniques would be the ultimate success. The entire Python code could be compiled and packaged into an easily installable Python module, that could be downloaded and run anywhere in the world with a few minor alterations by individual end users. The ease at which the HAW algorithm can be shared and executed, combined with the extremely low level cost of the entire monitoring system, demonstrates that it has the potential to be an invaluable monitoring tool at volcano observatories around the world. The HAW algorithm would allow researchers and decision makers to take webcam data and move from the qualitative world into quantitative analysis and add the imagery and derived products into their decision-making process.

Section 6: References

Ajadi, O.A., Meyer, F.J. and Webley, P.W., 2016. Change Detection in Synthetic Aperture Radar Images Using a Multiscale-Driven Approach. *Remote Sensing*, 8(6): 482.

Alaska Volcano Observatory (AVO) website, 2016. <https://www.avo.alaska.edu/activity/report.php?type=8&id=349921&mode=hans>.

Arámbula-Mendoza, R., Valdés-González, C. and Martínez-Bringas, A., 2010. Temporal and spatial variation of the stress state of Popocatepetl Volcano, Mexico. *Journal of Volcanology and Geothermal Research*, 196(3): 156-168.

Arámbula-Mendoza, R., Valdés-González, C., Varley, N., Juárez-García, B., Alonso-Rivera, P. and Hernández-Joffre, V., 2013. Observation of vulcanian explosions with seismic and acoustic data at Popocatepetl volcano, Mexico. *Monitoring of Volcanic Activity: Methods and Results*, edited by Z. Vyacheslav: 13-33.

Arrighi, S., Rosi, M., Tanguy, J.C. and Courtillot, V., 2004. Recent eruptive history of Stromboli (Aeolian Islands, Italy) determined from high-accuracy archeomagnetic dating. *Geophysical Research Letters*, 31(19).

Aspinall, W., Auker, M., Hincks, T., Mahony, S., Nadim, F., Pooley, J., Sparks, R. and Syre, E., 2011. Volcano hazard and exposure in GDRFF priority countries and risk mitigation measures- GFDRR Volcano Risk Study. Bristol: Bristol University Cabot Institute and NGI Norway for the World Bank: NGI Report, 20100806(2011): 3.

Auker, M.R., Sparks, R.S.J., Siebert, L., Crosweller, H.S. and Ewert, J., 2013. A statistical analysis of the global historical volcanic fatalities record. *Journal of Applied Volcanology*, 2(1): 2.

Barberi, F., Rosi, M. and Sodi, A., 1993. Volcanic hazard assessment at Stromboli based on review of historical data. *Acta Vulcanol*, 3: 173-187.

Belousov, A., Belousova, M. and Voight, B., 1999. Multiple edifice failures, debris avalanches and associated eruptions in the Holocene history of Shiveluch volcano, Kamchatka, Russia. *Bulletin of Volcanology*, 61(5): 324-342.

Bertagnini, A., Di Roberto, A. and Pompilio, M., 2011. Paroxysmal activity at Stromboli: lessons from the past. *Bulletin of Volcanology*, 73(9): 1229-1243.

Blackburn, E., Wilson, L. and Sparks, R.S.J., 1976. Mechanisms and dynamics of strombolian activity. *Journal of the Geological Society*, 132(4): 429-440.

Bullard, F.M., 1980. *Volcanoes of the Earth*. University of Texas Press.

Calvari, S., Lodato, L. and Spampinato, L., 2004. Monitoring active volcanoes using a handheld thermal camera, Defense and Security. *International Society for Optics and Photonics*, pp. 199-209.

Canny, J., 1986. A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*(6): 679-698.

Carling, G.T., Radebaugh, J., Saito, T., Lorenz, R.D., Dangerfield, A., Tingey, D.G., Keith, J.D., South, J.V., Lopes, R.M. and Diniega, S., 2015. Temperatures, thermal structure, and behavior of eruptions at Kilauea and Erta Ale volcanoes using a consumer digital camcorder. *GeoResJ*, 5: 47-56.

Carr, M.J., Patino, L.C. and Feigenson, M.D., 2007. Petrology and geochemistry of lavas. *Central America: geology, resources and hazards*, 1: 565-577.

Casadevall, T.J., 1994. The 1989–1990 eruption of Redoubt Volcano, Alaska: impacts on aircraft operations. *Journal of Volcanology and Geothermal Research*, 62(1-4): 301-316.

Chesner, C.A. and Halsor, S.P., 2006. The Escuintla and La Democracia debris avalanche deposits, Guatemala: Constraining their sources. *Geological Society of America Special Papers*, 412: 105-120.

Crisp, J.A., 1984. Rates of magma emplacement and volcanic output. *Journal of Volcanology and Geothermal Research*, 20(3-4): 177-211.

Davis, J.C. and Sampson, R.J., 1986. *Statistics and data analysis in geology*.

Dehn, J. and Harris, A.J., 2015. Thermal anomalies at volcanoes, *Monitoring Volcanoes in the North Pacific*. Springer, pp. 49-78.

Dirksen, O., Humphreys, M., Pletchov, P., Melnik, O., Demyanchuk, Y., Sparks, R. and Mahony, S., 2006. The 2001–2004 dome-forming eruption of Shiveluch volcano, Kamchatka: Observation, petrological investigation and numerical modelling. *Journal of volcanology and geothermal research*, 155(3): 201-226.

Duda, K.A., Ramsey, M., Wessels, R. and Dehn, J., 2009. *Optical Satellite Volcano Monitoring: A Multi T Sensor Rapid Response System*.

Erdem, J.E. and Waite, G.P., 2013. Temporal changes in eruptive behavior identified with coda wave interferometry and seismo-acoustic observations at Fuego Volcano, Guatemala. *Geological Society of America Special Papers*, 498: 107-123.

Francis, P., 1993. *Volcanoes. A planetary perspective*. Volcanoes. A planetary perspective., by Francis, P.. Clarendon Press, Oxford (UK), 1993, 452 p., ISBN 0-19-854452-9, ISBN 0-19-854033-7 (paper). 1.

Furukawa, Y., 2010. Infrared thermography of the fumarole area in the active crater of the Aso volcano, Japan, using a consumer digital camera. *Journal of Asian Earth Sciences*, 38(6): 283-288.

Gasparini, C., Iannaccone, G., Scandone, P. and Scarpa, R., 1982. Seismotectonics of the Calabrian arc. *Tectonophysics*, 84(2-4): 267-286.

Gillot, P. and Keller, J., 1993. Radiochronological dating of Stromboli. *Acta Vulcanol*, 3: 69-77.

Gorshkov, G. and Dubik, Y.M., 1970. Gigantic directed blast at Shiveluch volcano (Kamchatka). *Bulletin of Volcanology*, 34(1): 261-288.

Gregg, C. and Hazelwood, K., 2011. Where is the data? Why you cannot debate CPU vs. GPU performance without the answer, Performance Analysis of Systems and Software (ISPASS), 2011 IEEE International Symposium on. IEEE, pp. 134-144.

Gregg, C.E., Houghton, B. and Ewert, J.W., 2015. Chapter 67 - Volcano Warning Systems A2 - Sigurdsson, Haraldur, The Encyclopedia of Volcanoes (Second Edition). Academic Press, Amsterdam, pp. 1173-1185.

Grundland, M. and Dodgson, N.A., 2005. The decolorize algorithm for contrast enhancing, color to grayscale conversion. University of Cambridge, Technical Report UCAM-CL-TR-649.

Gudmundsson, M.T., Thordarson, T., Höskuldsson, Á., Larsen, G., Björnsson, H., Prata, F.J., Oddsson, B., Magnússon, E., Högnadóttir, T. and Petersen, G.N., 2012. Ash generation and distribution from the April-May 2010 eruption of Eyjafjallajökull, Iceland. Scientific reports, 2: 572.

Guffanti, M., Casadevall, T. and Mayberry, G., 2004. Reducing encounters of aircraft with volcanic-ash clouds, 2nd International Conference on Volcanic Ash and Aviation Safety.

Guffanti, M., Mayberry, G.C., Casadevall, T.J. and Wunderman, R., 2009. Volcanic hazards to airports. Natural Hazards, 51(2): 287-302.

Hanstrum, B. and Watson, A., 1983. A case study of two eruptions of Mount Galunggung and an investigation of volcanic eruption cloud characteristics using remote sensing techniques. Aust. Met. Mag, 31: 131-177.

Harris, A., Pirie, D., Horton, K., Garbeil, H., Pilger, E., Ramm, H., Hoblitt, R., Thornber, C., Ripepe, M. and Marchetti, E., 2005. DUCKS: low cost thermal monitoring units for near-vent deployment. *Journal of volcanology and geothermal research*, 143(4): 335-360.

Heidorn, P.B., 2008. Shedding light on the dark data in the long tail of science. *Library Trends*, 57(2): 280-299.

Higgins, J. and Harris, A., 1997. VAST: a program to locate and analyse volcanic thermal anomalies automatically from remotely sensed data. *Computers & Geosciences*, 23(6): 627-645.

Hort, M., Seyfried, R. and Vöge, M., 2003. Radar Doppler velocimetry of volcanic eruptions: theoretical considerations and quantitative documentation of changes in eruptive behaviour at Stromboli volcano, Italy. *Geophysical Journal International*, 154(2): 515-532.

Janssen, V., Roberts, C., Rizos, C. and Abidin, H.Z., 2002. Low-cost GPS-based volcano deformation monitoring at Mt. Papandayan, Indonesia. *Journal of Volcanology and Geothermal Research*, 115(1): 139-151.

Jonkman, S.N., 2005. Global Perspectives on Loss of Human Life Caused by Floods. *Natural Hazards*, 34(2): 151-175.

Kamchatka Volcanic Eruption Response Team (KVERT) website, 2016a. <http://www.kscnet.ru/ivs/kvert/van/index.php?n=2015-70>

Kamchatka Volcanic Eruption Response Team (KVERT) website, 2016b. <http://www.kscnet.ru/ivs/kvert/van/index.php?n=2015-40>

Kokelaar, P. and Romagnoli, C., 1995. Sector collapse, sedimentation and clast population evolution at an active island-arc volcano: Stromboli, Italy. *Bulletin of Volcanology*, 57(4): 240-262.

Kuznetsov, P. and Koulakov, I.Y., 2014. The three-dimensional structure beneath the Popocatépetl volcano (Mexico) based on local earthquake seismic tomography. *Journal of Volcanology and Geothermal Research*, 276: 10-21.

Larson, K.M., Poland, M. and Miklius, A., 2010. Volcano monitoring using GPS: Developing data analysis strategies based on the June 2007 Kīlauea Volcano intrusion and eruption. *Journal of Geophysical Research: Solid Earth*, 115(B7): n/a-n/a.

Lyons, J.J. and Waite, G.P., 2011. Dynamics of explosive volcanism at Fuego volcano imaged with very long period seismicity. *Journal of Geophysical Research: Solid Earth*, 116(B9).

Lyons, J.J., Waite, G.P., Rose, W.I. and Chigna, G., 2010. Patterns in open vent, strombolian behavior at Fuego volcano, Guatemala, 2005–2007. *Bulletin of Volcanology*, 72(1): 1.

Martin, D.P. and Rose, W.I., 1981. Behavioral patterns of Fuego volcano, Guatemala. *Journal of Volcanology and Geothermal Research*, 10(1-3): 67-81.

Melekestsev, I., Volynets, O., Ermakov, V., Kirsanova, T. and Masurenkov, Y.P., 1991. Shiveluch volcano. *Active volcanoes of Kamchatka*, 1: 84-92.

Meyer, F., McAlpin, D., Gong, W., Ajadi, O., Arko, S., Webley, P. and Dehn, J., 2015. Integrating SAR and derived products into operational volcano monitoring and decision support systems. *ISPRS Journal of Photogrammetry and Remote Sensing*, 100: 106-117.

Miller, T., McGimsey, R., Richter, D., Riehle, J., Nye, C., Yount, M. and Dumoulin, J., 1998. Catalog of the historically active volcanoes of Alaska. US Geological Survey Open-File Report, 98: 582.

Miller, T.P. and Casadevall, T.J., 2000. Volcanic ash hazards to aviation. *Encyclopedia of volcanoes*: 915-930.

Moon, T.K., 1996. The expectation-maximization algorithm. *IEEE Signal processing magazine*, 13(6): 47-60.

Nadeau, P.A., Palma, J.L. and Waite, G.P., 2011. Linking volcanic tremor, degassing, and eruption dynamics via SO₂ imaging. *Geophysical Research Letters*, 38(1): n/a-n/a.

Newhall, C.G. and Self, S., 1982. The volcanic explosivity index (VEI) an estimate of explosive magnitude for historical volcanism. *Journal of Geophysical Research: Oceans*, 87(C2): 1231-1238.

Parker, C.F., 2015. Complex Negative Events and the Diffusion of Crisis: Lessons from the 2010 and 2011 Icelandic Volcanic Ash Cloud Events. *Geografiska Annaler: Series A, Physical Geography*, 97(1): 97-108.

Paskievitch, J., Read, C. and Parker, T., 2006. Remote telemetered and time-lapse cameras at Augustine Volcano. *The*: 285-293.

Patrick, M.R., Harris, A.J., Ripepe, M., Dehn, J., Rothery, D.A. and Calvari, S., 2007. Strombolian explosive styles and source conditions: insights from thermal (FLIR) video. *Bulletin of Volcanology*, 69(7): 769-784.

Patrick, M.R., Kauahikaua, J.P. and Antolik, L., 2010. MATLAB tools for improved characterization and quantification of volcanic incandescence in webcam imagery: Applications at Kīlauea Volcano, Hawai'i. *US Geological Survey Techniques and Methods 13–A1*: 16.

Patrick, M.R., Orr, T., Antolik, L., Lee, L. and Kamibayashi, K., 2014. Continuous monitoring of Hawaiian volcanoes with thermal cameras. *Journal of Applied Volcanology*, 3(1): 1.

Patrick, M.R., Orr, T.R., Lee, L. and Moniz, C.J., 2015. A multipurpose camera system for monitoring Kīlauea Volcano, Hawai'i. 2328-7055, *US Geological Survey*.

Patrick, M.R., Swanson, D. and Orr, T., 2016. Automated tracking of lava lake level using thermal images at Kīlauea Volcano, Hawai'i. *Journal of Applied Volcanology*, 5(1): 6.

Pavolonis, M.J., Feltz, W.F., Heidinger, A.K. and Gallina, G.M., 2006. A daytime complement to the reverse absorption technique for improved automated detection of volcanic ash. *Journal of atmospheric and oceanic technology*, 23(11): 1422-1444.

Pevzner, M., Babansky, A., Tolstykh, M. and Kononkova, N., 2013. Reconstruction of the magmatic system in the Shiveluch volcanic massif as a result of large-scale collapses of its edifice in the late Pleistocene-early Holocene, *Doklady Earth Sciences*. Springer, pp. 35-37.

Ponomareva, V.V., Pevzner, M.M. and Melekestsev, I.V., 1998. Large debris avalanches and associated eruptions in the Holocene eruptive history of Shiveluch volcano, Kamchatka, Russia. *Bulletin of Volcanology*, 59(7): 490-505.

Prata, F. and Rose, B., 2015. Chapter 52 - Volcanic Ash Hazards to Aviation A2 - Sigurdsson, Haraldur, *The Encyclopedia of Volcanoes (Second Edition)*. Academic Press, Amsterdam, pp. 911-934.

Ripepe, M., Rossi, M. and Saccorotti, G., 1993. Image processing of explosive activity at Stromboli. *Journal of Volcanology and Geothermal Research*, 54(3): 335-351.

Rizzo, A.L., Federico, C., Inguaggiato, S., Sollami, A., Tantillo, M., Vita, F., Bellomo, S., Longo, M., Grassa, F. and Liuzzo, M., 2015. The 2014 effusive eruption at Stromboli volcano (Italy): Inferences from soil CO₂ flux and ³He/⁴He ratio in thermal waters. *Geophysical Research Letters*, 42(7): 2235-2243.

Robey, F.C., Fuhrmann, D.R., Kelly, E.J. and Nitzberg, R., 1992. A CFAR adaptive matched filter detector. *IEEE Transactions on aerospace and electronic systems*, 28(1): 208-216.

Rodríguez, L.A., Watson, I.M., Rose, W.I., Branam, Y.K., Bluth, G.J., Chigna, G., Matías, O., Escobar, D., Carn, S.A. and Fischer, T.P., 2004. SO₂ emissions to the atmosphere from active volcanoes in Guatemala and El Salvador, 1999–2002. *Journal of Volcanology and Geothermal Research*, 138(3): 325-344.

Romero, R., 2004. The international airways volcano watch (IAVW), *Proceedings of 2nd International Conference on Volcanic Ash and Aviation Safety*, Virginia, USA. Citeseer.

Rose, W., Self, S., Murrow, P., Ernst, G., Bonadonna, C. and Durant, A., 2007. Pyroclastic fall deposit from the October 14, 1974 eruption of Fuego Volcano, Guatemala. *Volcanol*, 70: 1043-1067.

Rosi, M., Bertagnini, A. and Landi, P., 2000. Onset of the persistent activity at Stromboli volcano (Italy). *Bulletin of volcanology*, 62(4): 294-300.

Rozenberg, G., 1963. Twilight phenomena, their nature, and use for atmospheric research. *Soviet Physics Uspekhi*, 6(2): 198.

Schaefer, J.T., 1990. The critical success index as an indicator of warning skill. *Weather and Forecasting*, 5(4): 570-575.

Sentman, D.D., McNutt, S.R., Stenbaek-Nielsen, H.C., Tytgat, G. and DeRoin, N., 2006. Imaging observations of thermal emissions from Augustine Volcano using a small astronomical camera. *The*: 569-577.

Siebe, C., Abrams, M., Macías, J.L. and Obenholzner, J., 1996. Repeated volcanic disasters in Prehispanic time at Popocatepetl, central Mexico: Past key to the future? *Geology*, 24(5): 399-402.

Siebert, L., Simkin, T. and Kimberly, P., 2011. *Volcanoes of the World*. Univ of California Press.

Small, C. and Naumann, T., 2001. The global distribution of human population and recent volcanism. *Global Environmental Change Part B: Environmental Hazards*, 3(3): 93-109.

Smithsonian Institution Global Volcanism Program (SIGVP, 2016. <https://volcano.si.edu/showreport.cfm?doi=GVP.WVAR20160413-341090>

Spampinato, L., Calvari, S., Oppenheimer, C. and Boschi, E., 2011. Volcano surveillance using infrared cameras. *Earth-Science Reviews*, 106(1–2): 63-91.

Sparks, R.S.J., Biggs, J. and Neuberg, J.W., 2012. Monitoring Volcanoes. *Science*, 335(6074): 1310-1311.

Speranza, F., Pompilio, M., D'Ajello Caracciolo, F. and Sagnotti, L., 2008. Holocene eruptive history of the Stromboli volcano: constraints from paleomagnetic dating. *Journal of Geophysical Research: Solid Earth*, 113(B9).

Stahl, C. and Schoppmann, P., 2000. Advanced automatic target recognition for police helicopter missions, pp. 61-68.

Stevenson, J.A. and Varley, N., 2008. Fumarole monitoring with a handheld infrared camera: Volcán de Colima, Mexico, 2006–2007. *Journal of Volcanology and Geothermal Research*, 177(4): 911-924.

Taddeucci, J., Edmonds, M., Houghton, B., James, M.R. and Vergnolle, S., 2015. Chapter 27 - Hawaiian and Strombolian Eruptions A2 - Sigurdsson, Haraldur, *The Encyclopedia of Volcanoes (Second Edition)*. Academic Press, Amsterdam, pp. 485-503.

Thomas, H.E. and Prata, A., 2011. Sulphur dioxide as a volcanic ash proxy during the April-May 2010 eruption of Eyjafjallajökull Volcano, Iceland. *Atmospheric Chemistry and Physics*, 11(14): 6871.

Tibaldi, A., 2001. Multiple sector collapses at Stromboli volcano, Italy: how they work. *Bulletin of Volcanology*, 63(2-3): 112-125.

Twigg, J., 2002. The human factor in early warnings: risk perception and appropriate communications. *Early warning systems for natural disaster reduction*: 19-26.

Vallance, J., Schilling, S., Matías, O., Rose Jr, W.I. and Howell, M., 2001. Volcano hazards at Fuego and Acatenango, Guatemala. 2331-1258, US Geological Survey.

van Manen, S.M., Blake, S. and Dehn, J., 2012. Satellite thermal infrared data of Shiveluch, Kliuchevskoi and Karymsky, 1993–2008: effusion, explosions and the potential to forecast ash plumes. *Bulletin of volcanology*, 74(6): 1313-1335.

Webley, P., Dehn, J., Lovick, J., Dean, K., Bailey, J. and Valcic, L., 2009a. Near-real-time volcanic ash cloud detection: Experiences from the Alaska Volcano Observatory. *Journal of Volcanology and Geothermal Research*, 186(1): 79-90.

Webley, P., Stunder, B. and Dean, K., 2009b. Preliminary sensitivity study of eruption source parameters for operational volcanic ash cloud transport and dispersion models—A case study of the August 1992 eruption of the Crater Peak vent, Mount Spurr, Alaska. *Journal of Volcanology and Geothermal Research*, 186(1): 108-119.

Webley, P.W., Dean, K., Peterson, R., Steffke, A., Harrild, M. and Groves, J., 2012. Dispersion modeling of volcanic ash clouds: North Pacific eruptions, the past 40 years: 1970–2010. *Natural Hazards*, 61(2): 661-671.

Worden, A., Dehn, J., Ripepe, M. and Delle Donne, D., 2014. Frequency based detection and monitoring of small scale explosive activity by comparing satellite and ground based infrared observations at Stromboli Volcano, Italy. *Journal of Volcanology and Geothermal Research*, 283: 159-171.

Zobin, V.M. and Martínez, A., 2010. Quantification of the 1998–1999 explosion sequence at Popocatepetl volcano, Mexico. *Journal of Volcanology and Geothermal Research*, 194(4): 165-173.