UiT
THE ARCTIC
UNIVERSITY
OF NORWAY

Faculty of Science and Technology
Department of Physics and Technology

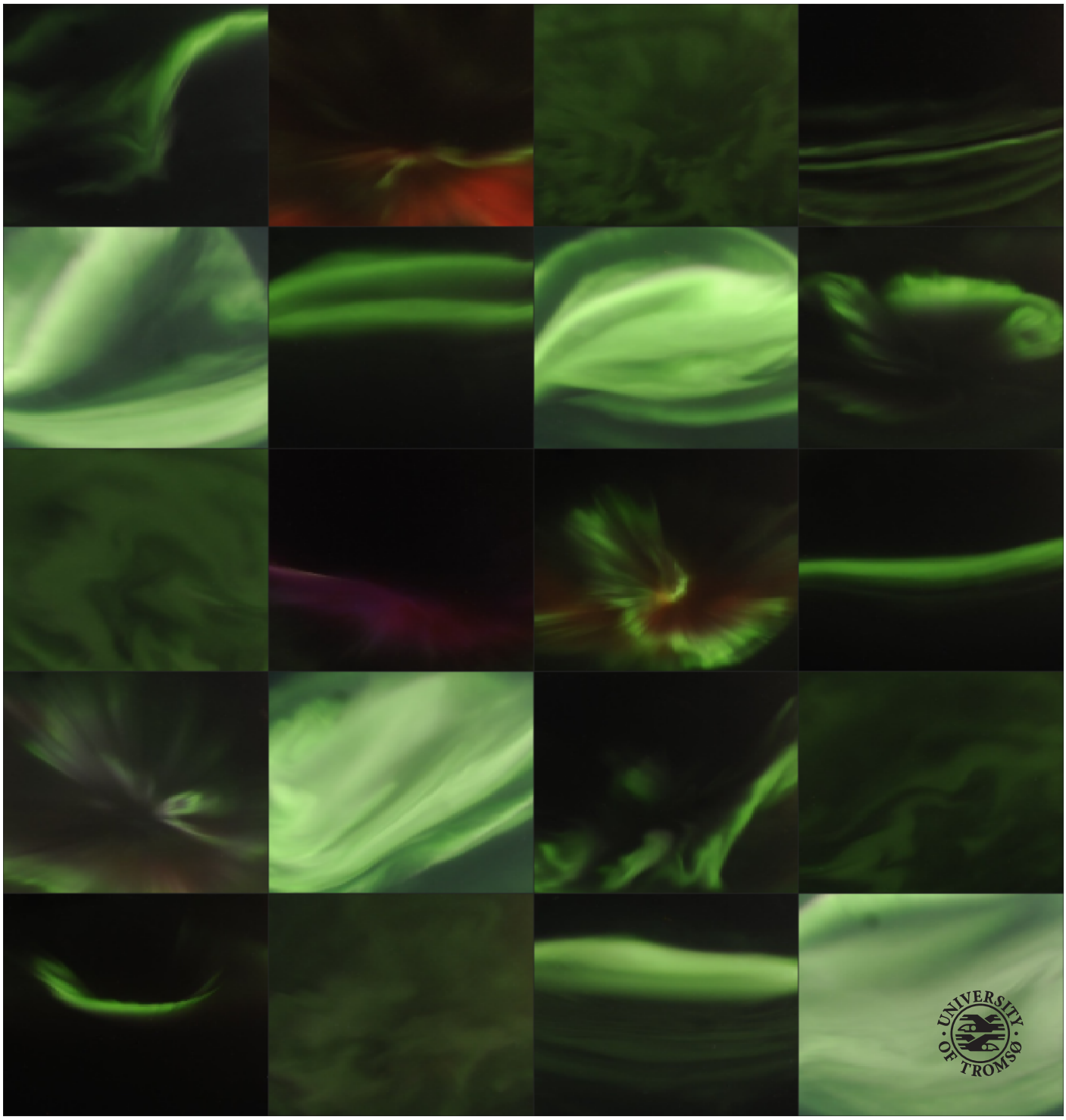
Auroral Image Processing Techniques

Machine Learning Classification and Multi-Viewpoint Analysis

—

Andreas Kvammen

A dissertation for the degree of Philosophiae Doctor – July 2021



Abstract

Every year, millions of scientific images are acquired in order to study the auroral phenomena. The accumulated data contain a vast amount of untapped information that can be used in auroral science. Yet, auroral research has traditionally been focused on case studies, where one or a few auroral events have been investigated and explained in detail. Consequently, theories have often been developed on the basis of limited data sets, which can possibly be biased in location, spatial resolution or temporal resolution.

Advances in technology and data processing now allow for acquisition and analysis of large image data sets. These tools have made it feasible to perform statistical studies based on auroral data from numerous events, varying geophysical conditions and multiple locations in the Arctic and Antarctic. Such studies require reliable auroral image processing techniques to organize, extract and represent the auroral information in a scientifically rigorous manner, preferably with a minimal amount of user interaction. This dissertation focuses on two such branches of image processing techniques: *machine learning classification* and *multi-viewpoint analysis*.

Machine learning classification: This thesis provides an in-depth description on the implementation of machine learning methods for auroral image classification; from raw images to labeled data. The main conclusion of this work is that convolutional neural networks stand out as a particularly suitable classifier for auroral image data, achieving up to 91 % average class-wise accuracy. A major challenge is that most auroral images have an ambiguous auroral form. These images can not be readily labeled without establishing an auroral morphology, where each class is clearly defined.

Multi-viewpoint analysis: Three multi-viewpoint analysis techniques are evaluated and described in this work: triangulation, shell-projection and 3–D reconstruction. These techniques are used for estimating the volume distribution of artificially induced aurora and the height and horizontal distribution of a newly reported auroral feature: Lumikot aurora. The multi-viewpoint analysis techniques are compared and methods for obtaining uncertainty estimates are suggested.

Overall, this dissertation evaluates and describes auroral image processing techniques that require little or no user input. The presented methods may therefore facilitate statistical studies such as: probability studies of auroral classes, investigations of the evolution and formation of auroral structures, and studies of the height and distribution of auroral displays. Furthermore, automatic classification and cataloging of large image data sets will support auroral scientists in finding the data of interest, reducing the needed time for manual inspection of auroral images.

Contents

Preface

During the dark and cold winter-season in Northern Norway, you learn to appreciate and wonder over the spectacular auroral displays that suddenly appear in the night sky. It was therefore a great privilege to be offered a PhD position in space physics at UiT – The Arctic University of Norway, where I could study the auroral phenomena in detail.

Initially, when I started the doctoral project in January 2017, I was intrigued by ionospheric modification experiments, where auroral emissions are artificially induced by ground-based facilities. This interest, supervised by Björn Gustavsson and Juha Vierinen, led to the work presented in Paper [IV] of this thesis.

Later, after taking a course on auroral substorms at UNiS – The University Centre in Svalbard, I was involved in a project where a previously unreported auroral feature was discovered. The analysis of this feature, named Lumikot[1] aurora, led to Paper [III] of this thesis.

In the latter part of my doctoral studies, I became interested in machine learning and its largely untapped potential in auroral science. This motivated me to study pattern recognition and the previous work in this field of research. During this time, I started discussions on deep neural networks and auroral science with Kristoffer Wickstrøm from the Machine Learning Group at UiT, Derek M^{c}Kay, then at NORCE – The Norwegian Research Centre and Björn Gustavsson from the Space Physics Group at UiT. These discussions initiated a collaborative project, together with Noora Partamies from the Department of Arctic Geophysics at UNiS, that ultimately led to Papers [I] and [II] of this thesis.

This dissertation is the result of the work presented in Papers [I], [II], [III] and [IV]. Chapter 1 provides an overview of the Aurora Borealis and artificial aurora, introducing the theoretical context of the included papers. Chapter 2 presents an extended introduction to auroral image classification with machine learning methods; from raw images to labeled data. Chapter 3 provides a description of multi-viewpoint analysis techniques, where the overall goal is to retrieve volumetric information from auroral images. Finally, Chapter 4 summarizes the presented work with concluding remarks and a discussion on future projects.

[1] Singular: Lumikko. Plural: Lumikot, after the Finnish word for weasel

List of Papers

This dissertation consists of a subject introduction and the following peer-reviewed papers:

I **Kvammen, A.,** Wickstrøm, K., McKay, D., and Partamies, N.: *Auroral Image Classification with Deep Neural Networks,* Journal of Geophysical Research: Space Physics, 125, e2020JA027 808, 2020.
URL https://doi.org/10.1029/2020JA027808.

II McKay, D., and **Kvammen, A.**: *Auroral classification ergonomics and the implications for machine learning,* Geoscientific Instrumentation, Methods and Data Systems, 9, 267-273, 2020.
URL https://doi.org/10.5194/gi-9-267-2020.

III McKay, D., Paavilainen, T., Gustavsson, B., **Kvammen, A.**, and Partamies, N.: *Lumikot: Fast auroral transients during the growth phase of substorms,* Geophysical Research Letters, 46, 7214-7221, 2019.
URL https://doi.org/10.1029/2019GL082985.

IV **Kvammen, A.**, Gustavsson, B., Sergienko, T., Brändström, U., Rietveld, M., Rexer, T., and Vierinen, J.: *The 3–D distribution of artificial aurora induced by HF radio waves in the ionosphere,* Journal of Geophysical Research: Space Physics, 124, 2992-3006, 2019.
URL https://doi.org/10.1029/2018JA025988.

In the text, these papers will be referred to by their Roman numerals.

Additional Publications

The following publications (listed alphabetically, by first author) were completed during the duration of the PhD programme, but are not included in the dissertation.

1. Dey, L., Valtonen, M. J., Gopakumar, A., Zola, S., Hudec, R., Pihajoki, P., ... **Kvammen, A.,** ... & Zielinski, P. (2018). Authenticating the presence of a relativistic massive black hole binary in OJ 287 using its general relativity centenary flare: Improved orbital parameters. *The Astrophysical Journal, 866*(1), 11. URL `https://doi.org/10.3847/1538-4357/aadd95`.

2. Goyal, A., Zola, S., Marchenko, V., Soida, M., Nilsson, K., Ciprini, S., ... **Kvammen, A.,** ... & Giroletti, M. (2018). Stochastic modeling of multiwavelength variability of the classical BL Lac object OJ 287 on timescales ranging from decades to hours. *The Astrophysical Journal, 863*(2), 175. URL `https://doi.org/10.3847/1538-4357/aad2de`.

3. Reed, M. D., Armbrecht, E. L., Telting, J. H., Baran, A. S., Østensen, R. H., Blay, P., **Kvammen, A.,** ... & Jeffery, C. S. (2018). K2 Campaign 5 observations of pulsating subdwarf B stars: binaries and super-Nyquist frequencies. *Monthly Notices of the Royal Astronomical Society*, 474(4), 5186-5198. URL `https://doi.org/10.1093/mnras/stx3133`.

Acknowledgements

First and foremost, I would like to thank my supervisors: Juha Vierinen, Björn Gustavsson and Derek McKay, for all your help throughout this project. Juha Vierinen, I am grateful for your support and for encouraging me to follow my scientific interests. In addition, thank you for taking me on much-needed skiing trips towards the hectic end of my doctoral period. Björn Gustavsson, I appreciate that you always find time to help me and I am grateful for all our discussions. Thank you for all you have taught me since I started this project. Derek McKay, your support has been crucial for this project and it is always a pleasure to work with you. Thank you for your assistance and guidance.

I would like to acknowledge the Swedish Institute of Space Physics (IRF) and the Finnish Meteorological Institute (FMI) for providing the auroral image data that has been used throughout this work. Special thanks to Urban Brandstöm for making the IRF data easily accessible and well-organized. I would also like to acknowledge the European Incoherent SCATter (EISCAT) community for conducting the ionospheric modification experiments. Many thanks to Michael Rietveld for your commitment to this interesting field of research.

To my fellow PhD students and colleagues at the Space Physics Group; thank you for making a great work environment. Special note to the camaraderie of Theresa Rexer and Derek McKay, many thanks for your support during the doctoral period. I am also grateful for the interesting and fruitful collaborations with Kristoffer Wickstrøm from the Machine Learning Group at UiT and Noora Partamies from the Department of Arctic Geophysics at UNiS.

In addition, I would like state my appreciation towards Roger Varney and SRI International, Menlo Park, for hosting my visit and for giving me the opportunity to do research among the great scientists at the Center for Geospace Studies. It was a very rewarding experience. I am also grateful towards UiT and the administration at the Department of Physics and Technology for the awarded travel grant and the bureaucratic assistance.

Last but not least, I would like to thank my family and friends. Special thanks to Christine Simonsen for all your support and comfort throughout this period.

This thesis was completed at UiT – The Arctic University of Norway. Contributions of co-authors to the included papers are acknowledged therein. All illustrations are made by the author unless otherwise stated. Any material from others is used with permission. This work has been funded by the Tromsø Research Foundation.

Chapter 1

Introduction

> *Man stod der undrende igjen, – sligt nordlys har jeg aldrig seet hverken før eller siden. Og der nede under os laa fjorden mørk og rolig.*
>
> – Fridtjof Nansen, *Paa ski over Grønland*

This dissertation focuses on processing techniques of auroral images. The topic is introduced by a description of the Aurora Borealis in Section 1.1 and an overview of artificial aurora in Section 1.2.

1.1 Aurora Borealis

The Aurora Borealis, commonly called the Northern lights, is a spectacular display that can be observed at Arctic latitudes. In the early days of science, the aurora was considered to be a meteorological phenomena, traditionally referred to as "weather lights" (Brekke, 2012). In the late 19th century however, it was proposed that the auroral displays typically form along annular belts encircling the magnetic poles and that the auroral activity is correlated with the sunspot number, indicating that the auroral driving mechanisms are dependent on both the Earth's magnetic field and solar conditions (Feldstein, 1986; Moss and Stauning, 2012). This notion was later supported by the Terrella[1] experiment of Kristian Birkeland at the beginning of the 20th century. Birkeland demonstrated that luminous rings appear around the Northern and Southern poles of the Terrella (a magnetized ball painted with a florescent layer) when bombarded by electrons in a vacuum chamber. Birkeland used the Terrella experiment as an analogy to argue that the auroral emissions around the magnetic poles of the Terra (the Earth) are caused by electrons streaming from the Sun.

Today, the coupling between the Sun and the Earth remains an active field of research and millions of scientific images are acquired annually to study the auroral phenomena. A brief summary of the auroral driving processes, as currently understood, is presented in this section. This overview is central for understanding why and how auroral images are classified in Papers [I] and [II], and understanding the

[1] Terrella: from Latin, meaning "little Earth"

analysis of the Lumikot aurora in Paper [III]. The section is organized chronologically, following the journey of the charged particles: from the ejection of particles from the Sun to the precipitation into the Earth's atmosphere.

The Solar Wind

The stream of charged particles that is constantly expanding from the Sun is called the *solar wind* and originates from the solar *corona*, the nonuniform aura of hot ($\sim 10^6$ K) particles that envelops the Sun (Gosling, 2014). The processes that heat the corona and accelerate the solar wind are poorly understood and still an active field of research. It is however proposed that charged particles can escape from the Sun through coronal holes, where the magnetic field is open, and from regions where the magnetic field is closed via rearrangement of the open-closed field boundary (Antiochos et al., 2011). Furthermore, large-scale bursts of particles are observed several times per day during active solar periods and approximately once every week during quiet periods. It is suggested that the large-scale ejections, called *coronal mass ejections*, are released as the solar atmosphere reconfigures itself in response to changes in the solar magnetic field (Gosling, 2014).

The solar corona can be observed directly during total solar eclipses, as seen in Figure 1.1, where the irregular structure indicates that the solar wind does not expand homogeneously, but is mainly released in streams. The plasma that streams

Figure 1.1: A photograph of the total solar eclipse over Ny-Ålesund, Svalbard, on 20 March 2015. The corona appears as an irregular aura that extends from $\sim$2000 km to 5 solar radii above the optical surface of the Sun, structured with radially extending rays (Brekke, 2012). Copyright Nathalie Grenzhaeuser.

from the Sun consists primarily of electrons and protons and is essentially fully ionized with a negligible resistance. The solar wind plasma therefore satisfies the *frozen in condition*, where the local magnetic field expands with the ejected plasma (and vice versa) (Gosling, 2014). The magnetic field that expands into the interplanetary space is known as the *Interplanetary Magnetic Field* (IMF). Overall, the solar wind velocity, the IMF magnitude and the IMF orientation determine the energy and particle input from the solar wind to the magnetic field of the Earth and therefore influence the auroral activity in the polar regions (Perreault and Akasofu, 1978; Rostoker et al., 1988).

Plasma Convection in the Magnetosphere

As the solar wind approaches the Earth, the solar particles and the embedded magnetic field interacts with the Earth's magnetic field, hereafter called the *geomagnetic* field, and the frozen in approximation ceases to be generally valid in the interaction region. The geomagnetic field is often approximated as a dipole in the near-Earth space. However, large-scale currents are induced when the solar wind interacts with the magnetic field that is surrounding the Earth, ultimately producing additional magnetic fields that combines with the geomagnetic field. The resulting magnetic field that is enclosing the Earth is the *magnetosphere* (Ganushkina et al., 2018; Luhmann and Solomon, 2007). The shape of the magnetosphere is dynamic, but can generally be described by a spheroid towards the dayside and by a highly elongated structure (the *magnetotail*) towards the nightside, as illustrated in Figure 1.2.

The solar wind is not fully deflected at the magnetopause. Both energy and particles from the solar wind may enter the magnetosphere when the solar wind interacts with the magnetosphere. The particle and energy transfer from the solar wind to the magnetosphere is closely related to the IMF orientation, as well as the IMF magnitude and the solar wind velocity (Rostoker et al., 1988; Tenfjord and Østgaard, 2013). Maximum transfer at the dayside magnetopause occurs when the IMF orientation is anti-parallel to the geomagnetic field, i.e. Southward relative to the Northward geomagnetic field at the dayside magnetopause. This corresponds to a negative magnetic field (z) component ($B_z < 0$) in the Geocentric Solar Magnetic (GSM) coordinate system (Laundal and Richmond, 2017).

During ($B_z < 0$) conditions, the IMF connects with the geomagnetic field at the dayside magnetopause and transfers energy and particles into the magnetosphere via *magnetic reconnection*. Magnetic reconnection is a physical process where magnetic fields with opposing orientation merge, forming a new magnetic field topology and releasing magnetic energy via plasma energization (Phan et al., 2000). Consequently, plasma is jetted away from the magnetic reconnection region, with a direction perpendicular to the original flow. The jetted plasma satisfies the frozen in condition and the merged magnetic field therefore flow with the plasma (Baumjohann and Treumann, 2012). These particles and the embedded magnetic field may undergo a cyclic convection within the magnetosphere, first proposed in Dungey (1961), and therefore named *the Dungey cycle*.

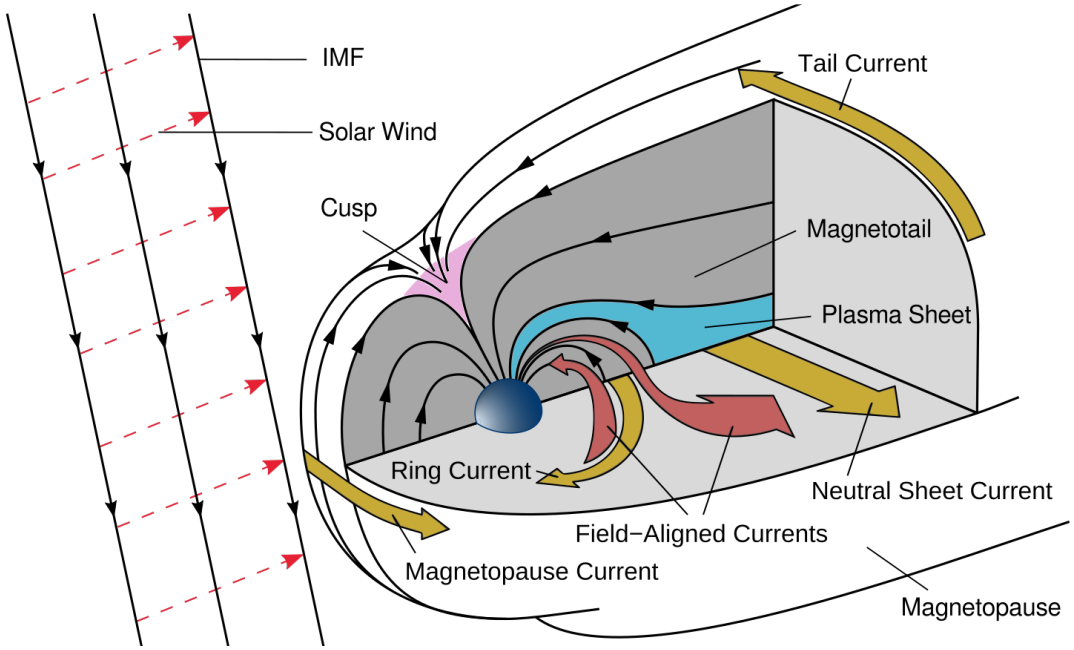


Figure 1.2: The Earth's magnetosphere with named regions and current systems. The interplanetary magnetic field (IMF) has a Southward orientation in the figure, i.e. anti-parallel to the Northward geomagnetic field at the dayside *magnetopause*. The magnetopause is the dynamic boundary between the magnetosphere and the interplanetary space, defined at the dayside by the balance of the kinetic pressure from the solar wind and the magnetic pressure from the magnetosphere. The magnetopause is typically located at ∼10 Earth radii from the Earth towards the dayside while the magnetic tail can extend up to hundreds of Earth radii towards the nightside (Baumjohann and Treumann, 2012). The Figure is adapted from Rexer (2015) with small additions.

The Dungey cycle timescale is roughly 1 hour, from the merging and opening of the magnetic fields in the dayside magnetopause to the re-merging and closing of the magnetic fields in the nightside magnetotail (Jackman et al., 2014). In this thesis, opened magnetic fields refer to a topology where the magnetic field has one footpoint connected to the Earth and the other connects with the IMF (outside the magnetosphere), while closed magnetic fields refer to a topology where both magnetic footpoints are connected to the Earth. A scheme of the Dungey cycle is depicted in Figure 1.3.

It should be noted that the Dungey cycle is an idealized steady-state model of the magnetospheric plasma convection during Southward IMF ($B_z < 0$) conditions. Steady state in this context refers to the situation where the reconnection rate, defined by the amount of magnetic flux merged per unit time, at the dayside magneopause (ϕ_D) is balanced with the reconnection rate in the nightside magnetotail (ϕ_N) (Baumjohann and Treumann, 2012). A steady state configuration is however rarely achieved in reality due to the variability of the solar wind, the convection time-lag and since the magnetotail both can accumulate energy during ($\phi_D > \phi_N$) conditions and unload the stored energy during ($\phi_D < \phi_N$) conditions (Zhang et al., 2015).

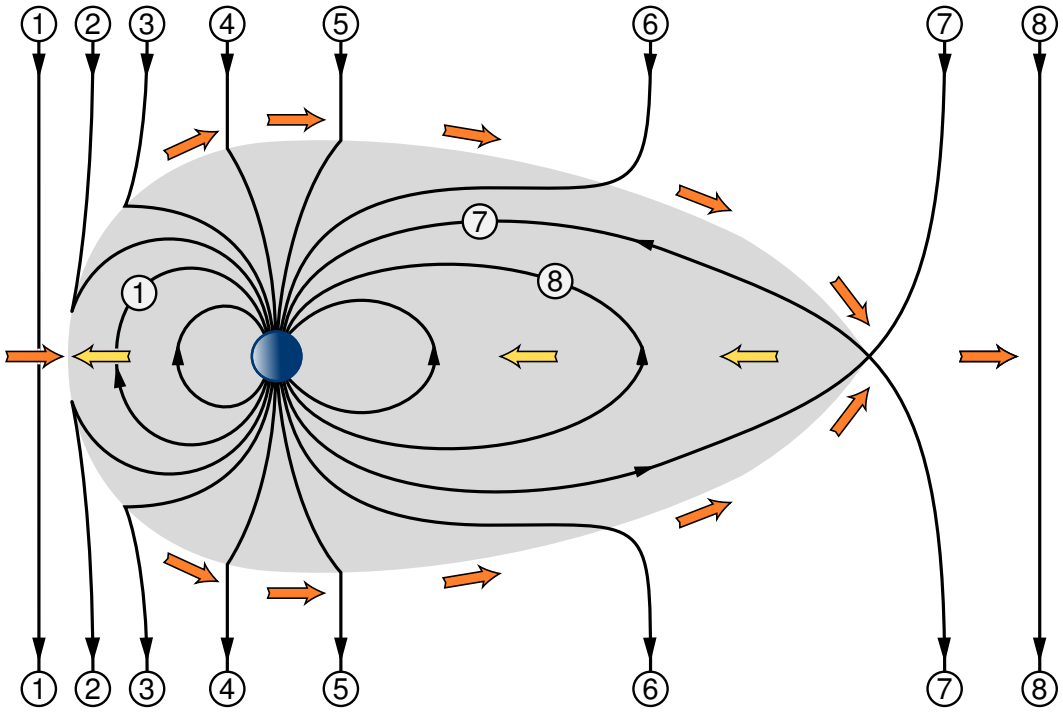


Figure 1.3: The diagram illustrates the plasma convection in the magnetopshere during Southward IMF ($B_z < 0$) conditions. 1; the Southward IMF reconnects with the Northward geomagnetic field ($B_z^E > 0$) at the magnetopause, resulting in a merged magnetic field with an open magnetic configuration. 2–6; the open magnetic field (and the frozen-in plasma) is "dragged" over the polar cap and towards the nightside magnetotail by the solar wind flow. 7; the open magnetic field from the Northern and Southern hemispheres meet and reconnects at the X-point. 8; plasma is jetted away from and towards the Earth by the released magnetic energy at the X-point. The plasma (and the embedded magnetic field) is eventually brought back to the dayside by the return flow, maintaining the equilibrium condition and completing the magnetospheric plasma convection cycle (Baumjohann and Treumann, 2012; Tenfjord et al., 2015). Note that the sketch is not to scale. During steady-state conditions, the tail reconnection region (the X point) is located in the distant magnetotail, ∼100–200 Earth radii towards the nightside (Baumjohann and Treumann, 2012). The Figure is reprinted from Rexer (2015).

The Magnetospheric Substorm

The *magnetospheric substorm* describes the loading and unloading of energy in the magnetotail by three distinct phases; *growth*, *expansion* and *recovery*. Typically, a magnetospheric substorm event has a duration of 3–4 hours, although another substorm may be initiated before the conclusion of the previous one (Akasofu, 2017).

The growth phase is characterized by a significant loading of energy in the magnetotail (Baker et al., 1985; McPherron, 1970). In an isolated substorm event, this can occur if the IMF quickly turns from a quasi-static Northward ($B_z > 0$) orientation with ($\phi_D \approx \phi_N$) to a quasi-static Southward ($B_z < 0$) orientation with ($\phi_D > \phi_N$). As energy is being loaded into the magnetotail, the tail cross-section increases (in

the xy-plane) and the neutral sheet current is enhanced. This leads to a thinning of the central region of the magnetotail, the *plasma sheet*, depicted in Figure 1.2. After $\sim$1 hour of energy loading into the magnetotail, perturbations appear within the thin plasma sheet. The magnetotail may then become unstable, triggering the second phase of the substorm; the expansion phase (Akasofu, 2013; Baumjohann and Treumann, 2012).

The expansion phase is initiated by magnetic reconnection in the near-Earth magnetotail (at $\sim$30 Earth radii) and a sudden enhancement in the magnetotail reconnection rate (Angelopoulos et al., 2008; Sergeev et al., 2011). During the expansion phase, the perturbed magnetotail starts unloading the accumulated energy ($\phi_D < \phi_N$) in a pulse-like manner by ejection of fast plasma flows. The magnetic reconnection in the near-Earth tail ceases after $\sim$45 minutes, initiating the final stage of the substorm; the recovery phase (Baumjohann and Treumann, 2012).

During the recovery phase, the reconnection region moves further out into the magnetotail and the stored energy in the magnetotail continues to unload ($\phi_D < \phi_N$). The recovery phase typically ends after $\sim$1–2 hours, when the reconnection region reaches the distant magnetotail and a stable configuration ($\phi_D \approx \phi_N$) is once again attained.

The Auroral Substorm

Hitherto, the plasma convection outside the Earth's atmosphere has been described. However, plasma also streams into and from the Earth's atmosphere, both during magnetic substorms and during quiet periods. More specifically, the charged particles flow into and from the *ionosphere*, the partly ionized upper atmosphere ($\sim$80–700 km), in field-aligned currents (Milan et al., 2017). The field-aligned currents connect the ionosphere to the plasma sheet and the ring current system, as illustrated in Figure 1.2 (Baumjohann and Treumann, 2012). The ionospheric footpoint of the field aligned currents encircle the magnetic poles and define the Northern and Southern *auroral ovals*, the annular belts where auroral displays most commonly appear. As the energetic particles enter the upper-atmosphere, they ionize, heat and excite neutral constituents, ultimately producing auroral displays that are characteristic for the magnetosphere–ionosphere coupling conditions in the associated mangetospheric region.

The sequence of auroral displays over the polar regions during the magnetospheric substorm was first described by Akasofu (1964) on a global scale and named the *auroral substorm*. Akasofu (1964) characterized the generic auroral substorm evolution after an extensive study of data from all-sky cameras and magnetometers in Alaska, Canada and Siberia. Note that in the initial description of the auroral substorm, only the expansion and recovery phases were defined, and the interval between isolated substorm events was named the quiet phase. Later, McPherron (1970) proposed to include the growth phase as the precursive stage of the substorm. Figure 1.4 depicts the idealized substorm evolution on a global scale, including examples of all-sky images that illustrates the characteristic signatures of the indicated substorm phase and location.

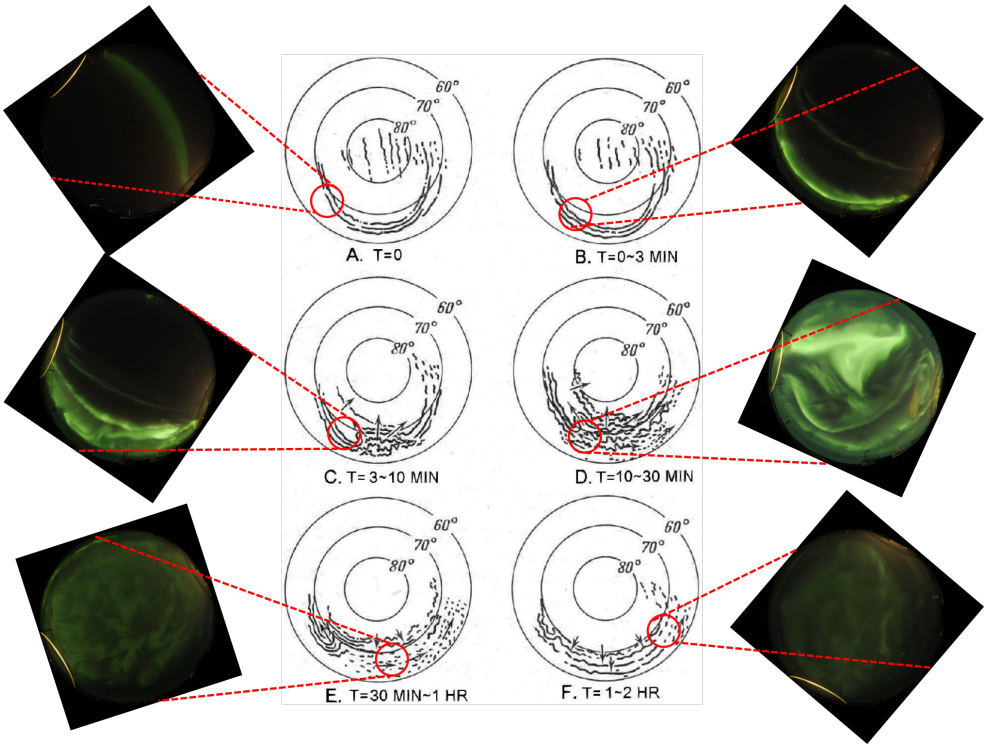


Figure 1.4: The Akosofu diagram illustrating the auroral substorm evolution including images that exemplify the characteristic signatures. The black circles indicate the magnetic latitudes with magnetic noon towards the top and magnetic dawn towards the right. The all-sky images are rotated to match the orientation of the field-of-view of the marked camera location, depicted by the red circles. A and B; a growth phase arc system is drifting Southwards and the Southernmost arc intensifies. C and D; the arc intensification initiates the rapid expansion of the bright and highly dynamic auroral bulge. E and F; after maximum expansion, the auroral substorm enters the recovery phase with pulsating auroral patches that drift towards the morning sector. The auroral displays slowly approach the quiet stage (A), concluding the auroral substorm. The Akosofu diagram is adapted from Feldstein et al. (2014) under the Creative Commons Attribution 3.0 Unported license, initially re-rendered from Akasofu (1964). The sample images are acquired by the all-sky camera in Kiruna, Sweden, operated by the Swedish Institute of Space Physics.

Figure 1.4 illustrates what Akasofu (1964) and McPherron (1970) suggested; the auroral emissions that appear in the ionosphere are linked to the magnetospheric substorm. The coupling between the ionosphere and the magnetosphere is an active field of research today. One of the goals is to understand the general mapping between the auroral displays and the associated magnetospheric production mechanisms. This topic is treated in Chapter 2; Auroral Image Classification, and is further discussed in Paper [I].

The Auroral Emissions

The charged high-energy particles that precipitate into the upper-atmosphere ionize the constituents in their path, ultimately modifying the ionospheric plasma.

The penetration depth of the charged particles is mainly dependent on their energy and the atmospheric neutral density. Figure 1.5 presents typical neutral density profiles in the Arctic upper-atmosphere, the ionization rate profiles for mono-energetic electron populations and typical electron density profiles under different ionospheric conditions.

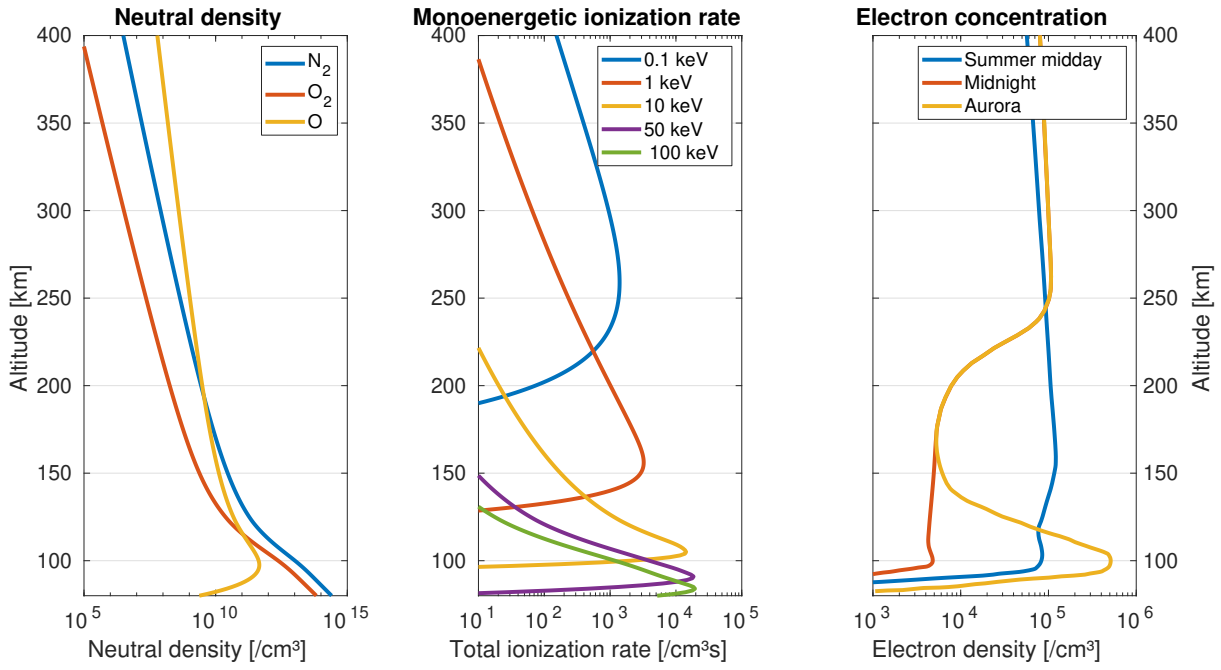


Figure 1.5: To the left, the main ionospheric constituents from the NRLMSISE–00 atmospheric model in the auroral altitude range (Picone et al., 2002). In the middle, the associated penetration depth for mono-energetic electron populations with an isotropic pitch angle distribution, calculated using the parameterized ionization rate model in Fang et al. (2008). To the right, the ionsopheric electron content under sunlit (summer midday), midnight and auroral conditions. The electron density data is based on Figure 4.2 from Brekke (2012).

During auroral events, the precipitating particles significantly modify the electron density at ∼90–150 km by ionization of neural constituents, an enhancement of two orders of magnitude is presented in Figure 1.5. The de-attached electrons produced in these collisions, the *secondary electrons*, have energies that typically match the peak of the cross-section energy range for various excited levels of atoms, molecules and ions in the upper atmosphere. Numerous emission lines, both visible and invisible to the naked eye, are therefore enhanced during auroral events.

The auroral intensity spectrum depends on the flux-energy distribution of the precipitating particles and the ionospheric background conditions (e.g. differences in the neutral density, temperature and solar zenith angle). A typical intensity spectrum, approximately in the visible wavelength region, during an auroral event is presented in Figure 1.6.

The main production mechanisms for three of the brightest optical emission lines;

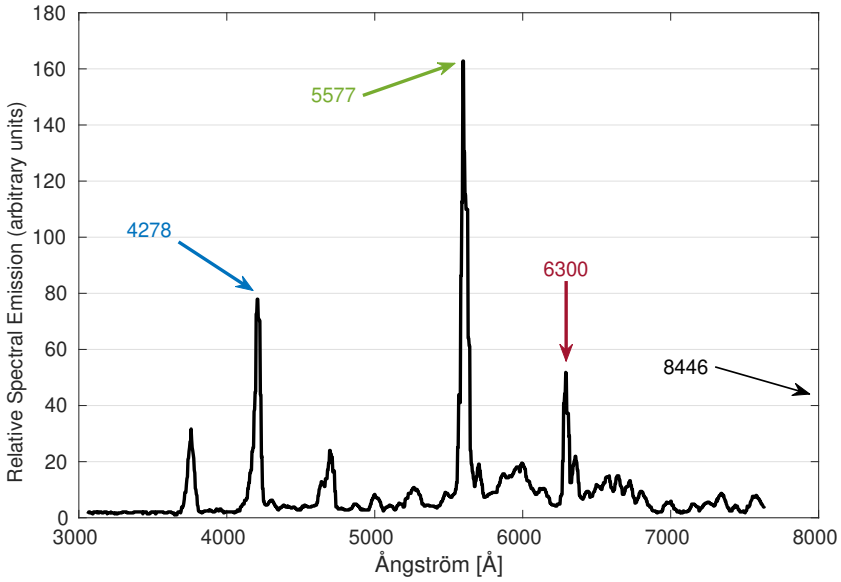


Figure 1.6: A typical auroral spectrum. The auroral emission spectrum was acquired at the Poker Flat Research Range in Fairbanks, Alaska, by the auroral spectral imager presented in Goenka et al. (2015) with courtesy of Jeff Baumgardner, CSP, Boston University. The re-printed spectrum values were extracted from Fig 6. in Goenka et al. (2015).

at 4278 Å, 5577 Å and 6300 Å, and the near infrared line, at 8446 Å, will be presented in this section as these emission lines are considered in Paper [IV]. It is however beyond the scope of this introduction to consider all emission lines and bands that are excited during auroral events. The review of the emission lines is inspired by the descriptions in Gustavsson (2000) and in Gustavsson and Eliasson (2008).

The Auroral Blue Line – 4278 Å

The auroral blue line is produced by de-excitation of excited ionized molecular nitrogen and can be observed towards the lower edge of bright auroral displays. More specifically, the 4278 Å photon is emitted by the transition of the excited $N_2^+(B^2\Sigma_u^+)$ state to the first vibrational level of the ground state of the molecular nitrogen ion, $N_2^+(X^2\Sigma_u^+)$ (Gustavsson, 2000). The main source of $N_2^+(B^2\Sigma_u^+)$ excitation is direct electron (e^-) impact:

$$N_2 + e^- \rightarrow N_2^+(B^2\Sigma_u^+) + 2e^- \tag{1.1}$$

The $N_2^+(B^2\Sigma_u^+)$ excitation state has a threshold energy of 18.75 eV (relative to the ground state of N_2) and emits a photon promptly (typically within 10^{-8} seconds) after excitation (Gilmore et al., 1992). The 4278 Å emission is mainly produced at lower altitudes (as compared to the auroral green line) and is therefore associated with precipitation of high energy electrons.

The Auroral Green Line – 5577 Å

The auroral green line is typically the brightest auroral emission. The 5577 Å photon is emitted when the second lowest excited state of atomic oxygen $O(^1S)$ relaxes

Figure 1.7: A bright auroral display with distinct colors. The apparent purple emission is mainly caused by the 4278 Å emission and is located towards the lower border of the bright auroral structure. The red auroral emission can be seen towards the top of the altitude-extending rays. Photo credit; Njål Gulbrandsen.

to the lowest excited state $O(^1D)$. The two main sources of $O(^1S)$ are:

- Direct electron collisions with atomic oxygen:

$$O + e^- \rightarrow O(^1S) + e^- \tag{1.2}$$

- Excitation by energy transfer from the $N_2(A^3\Sigma_u^+)$ state:

$$N_2(A^3\Sigma_u^+) + O \rightarrow N_2 + O(^1S) \tag{1.3}$$

The $O(^1S)$ state has a threshold energy of 4.17 eV and a radiative lifetime of approximately 0.7 seconds. The long lifetime leads to collisional de-excitation, also known as *quenching*, to become significant at altitudes below $\sim$100 km (Brekke, 2012). The $O(^1S)$ states that are quenched transfer their energy to the neural constituents by kinetic energy transfer and excitation of vibrational or rotational states, without emitting photons. The quenching rate is dependent on the collisional frequency and therefore increases as the neutral density increases at lower altitudes.

The Auroral Red Line – 6300 Å

The 6300 Å emission is induced by relaxation of the meta-stable $O(^1D)$, the lowest excited atomic oxygen state, to the ground state. The major ionospheric sources of the excited $O(^1D)$ state are:

- Excitation by dissociative recombination of an excited molecular oxygen ion (O_2^+) with an electron:
$$O_2^+ + e^- \rightarrow O + O(^1D) \tag{1.4}$$
- Excitation by direct electron collisions with atomic oxygen:
$$O + e^- \rightarrow O(^1D) + e^- \tag{1.5}$$

with smaller contributions from other processes, such as; energy transfer from excited atomic nitrogen states and cascading from $O(^1S)$.

$O(^1D)$ has a threshold energy of 1.96 eV and a long radiative lifetime of 107 seconds. Quenching of $O(^1D)$ is the dominant de-excitation process at typical auroral altitudes, due to the long lifetime, and most excited $O(^1D)$ states are quenched before emitting a photon during auroral events. Consequently, the red auroral emission is predominantly seen towards the top of auroral displays where the quenching rate is lower.

The Near Infra-Red Line – 8446 Å

The near infra-red emission is not visible to the naked eye but can be detected by auroral cameras. The 8446 Å emission is emitted by the transition $O(3p^3P) \rightarrow O(3s^3S)$ and the main source of $O(3p^3P)$ is direct electron impact on atomic oxygen:

$$O + e^- \rightarrow O(3p^3P) + e^- \tag{1.6}$$

The $O(3p^3P)$ excitation state has a threshold energy of 10.99 eV and emits a photon promptly after excitation (Gustavsson, 2000).

1.2 Artificial Aurora

The idea of artificially modifying the ionospheric F-region (at 150–800 km altitude) by stimulating the plasma with High-Frequency (HF) radio waves in the 3–10 MHz frequency range has been theoretically discussed since the 1960s (Farley Jr, 1963; LeLevier, 1969) and experimentally studied since the 1970s (Biondi et al., 1970; Utlaut, 1970). Biondi et al. (1970) first reported that optical emissions are artificially enhanced when HF radio waves are transmitted into the upper-atmosphere, and observations of the auroral emission lines has since been an important diagnostic tool for studying the modified plasma.

The enhanced emissions are often refereed to as *artificial aurora* when produced at auroral latitudes and *artificial airglow* when induced at sub-auroral latitudes.

The emissions appear at various wavelengths, with intensities and forms depending on both ionospheric conditions and the parameters of the transmitted radio waves, such as: power, frequency, polarization, beam pattern and the duration of the transmission. Field aligned rayed structures have for example been observed during very high-power transmission, above ∼400 MW effective radiative power (Kendall et al., 2010; Pedersen et al., 2009), while auroral patches, depicted in Figure 1.8, are typically observed at lower pump powers (Brändström et al., 1999; Grach et al., 2007; Gustavsson et al., 2001).

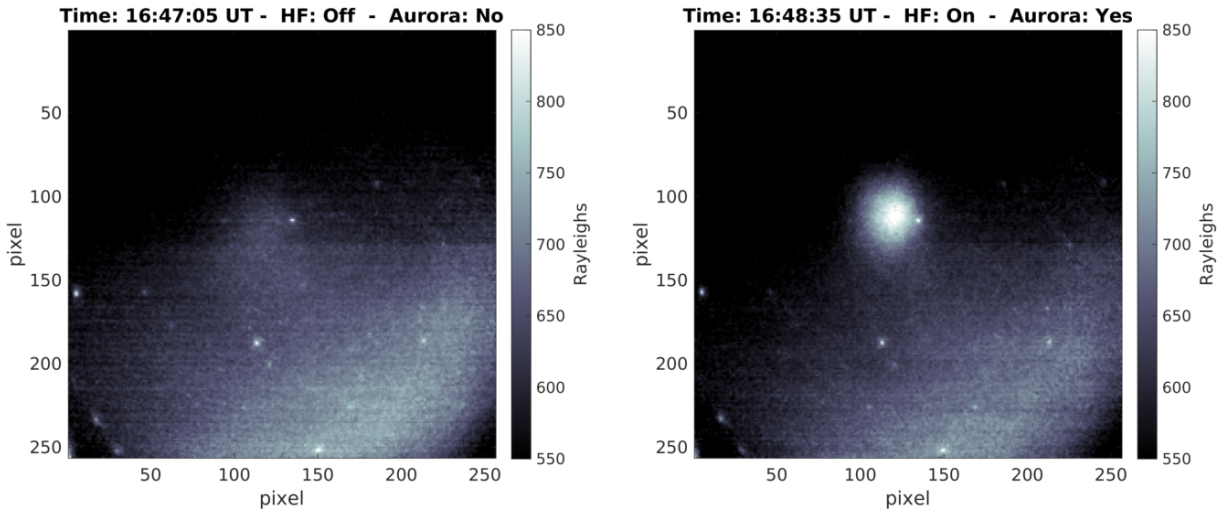


Figure 1.8: To the left, an image of the dusk sky, after sunset, near Abisko, Sweden, with a distance of approximately 140 km from the European Incoherent SCATter (EISCAT) Heating facility near Tromsø, Norway (Rietveld et al., 2016). To the right, a similar image but with ongoing HF transmission from the EISCAT Heater. The artificial aurora appears as a patch in the middle of the image. The images have a ($54^\circ \times 54^\circ$) field-of-view and are acquired by an Auroral Large Imaging System (ALIS) camera with a narrow-band filter, designed to capture emissions in the 6300 Å auroral line (Brändström, 2003).

The relationship between the enhanced emissions and the radio wave parameters is however not fully understood and still an active field of research.

An overview of ionospheric modification experiments and the current theoretical understanding is presented in this section. This summary introduces the topic of Paper [IV], where the 3–D distribution estimates of artificially induced auroral patches are presented.

Ionospheric Modification Experiments

Most commonly in ionospheric science, the upper atmosphere is studied by passive investigations, where the researcher observes the natural phenomena that appear or wait for a particular phenomena to occur. In contrast, during HF transmission experiments, the researcher is actively modifying a small volume of the upper-atmosphere and studies the response. These active investigations are therefore often refereed to as *ionospheric modification experiments*. The response can be

studied repeatably and under somewhat controlled conditions by adjustment of the transmitted wave parameters, thus turning the overhead ionosphere into a plasma laboratory (Streltsov et al., 2018).

Currently, there are three active facilities dedicated to conducting ionospheric modification experiments: the High-frequency Active Auroral Research Program (HAARP) facility near Gakona, Alaska, the Sura Ionospheric Heating Facility near Nizhny Novgorod, Russia and the EISCAT Heating facility near Tromsø, Norway. For technical details of these facilities, see for example Streltsov et al. (2018).

The EISCAT Heater is now the only ionospheric modification facility co-located with a functioning *Incoherent Scatter* (IS) radar, after the collapse of the Arecibo receiver platform on December 1, 2020. The experimental setup at EISCAT is illustrated in Figure 1.9, depicting the 32 meter IS radar and the EISCAT Heater beams. The great

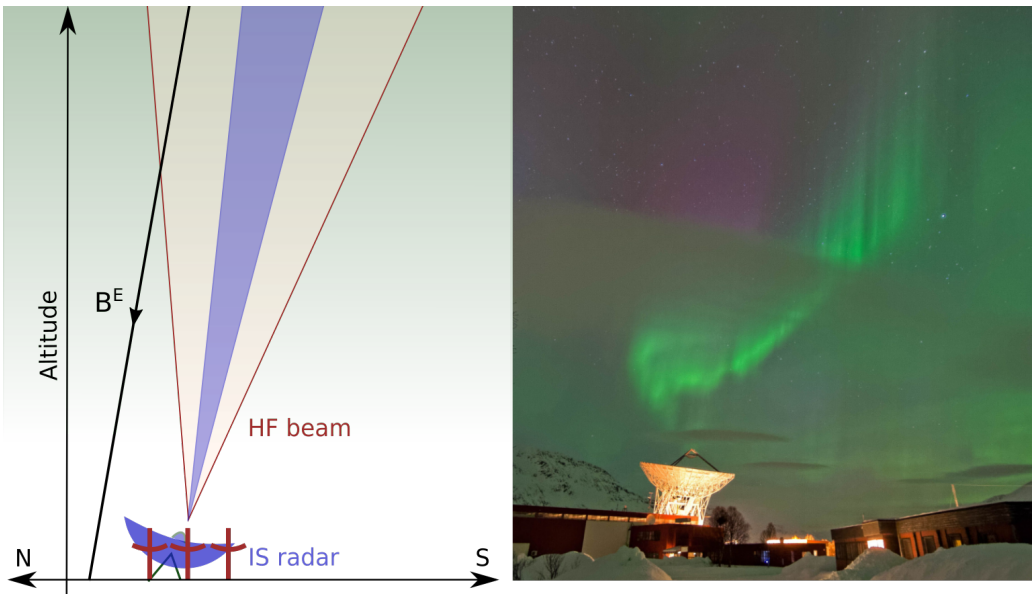


Figure 1.9: To the left, a diagram of the EISCAT heater beam and the IS radar beam in the meridional plane. The beams are not to scale and side-lobes are ignored for illustrative purposes. To the right, a photography of the 32 meter dish with aurora in the background. Note that the artificial aurora, induced by the EISCAT Heater is typically located 100–150 km above the peak-height of auroral arcs and is generally dimmer, smaller and less structured than natural auroral displays. Additionally, notice that the EISCAT Heater array is positioned ∼0.5 km from the dish antenna and is not depicted in the image. Illustration credit; Theresa Rexer, and photo credit; Njål Guldbrandsen.

advantage of the co-located IS radar is that the ionosphere can be monitored during HF transmission, allowing studies of the growth, decay and steady-state plasma response.

The IS radars are transmitting radio waves in the Very High (30–300 MHz) Frequency (VHF) and Ultra High (300–1000 MHz) Frequency (UHF) range into the upper-atmosphere and receive the signal that is scattered back from the free ionospheric electrons by the Thompson scattering process (Kudeki and Milla, 2010). The backscat-

tered signal is dependent on the collective state of the plasma within the observed volume. Thus, after processing, estimates of the plasma parameters (electron density, electron temperature, ion temperature and ion drift velocity) can be determined from the signal. See for example Kudeki and Milla (2010) for an introductory, but thorough description of the IS radar theory.

Figure 1.10 presents results from an ionospheric modification experiment at the EISCAT Heating facility, illustrating the electron temperature and the auroral emission enhancements. The ionospheric modifications, depicted in Figure 1.10, were induced by transmission of right-hand circular polarized radio waves, i.e. *Ordinary*-mode (O-mode) waves in the plasma. The O-mode waves, transmitted approximately anti-parallel to the magnetic field, generally induce the strongest artificial aurora response and is therefore the focus in this thesis. Yet strong *eXtraordinary*-mode (X-mode) response have been reported (Blagoveshchenskaya, 2020) and will likely be investigated further in future studies.

Additionally, the HF transmission triggers other observable phenomena in the ionosphere that are not treated in this thesis, such as: artificial ionization (Bernhardt et al., 2016; Pedersen et al., 2009), stimulated electromagnetic emissions (SEE) (Leyser, 2001), enhancement of the IS radar spectrum ion and plasma lines (Carlson et al., 1972; Rietveld et al., 2000), enhancement of the topside IS spectrum ion line (Isham et al., 1999; Rexer et al., 2018, 2021), wide altitude ion line enhancements (WAILEs) (Bazilchuk, 2019; Senior et al., 2013) and generation of Ultra, Extremely and Very Low Frequency (ULF/ELF/VLF) waves, see e.g. Streltsov et al. (2018) and references therein.

The Wave–Plasma Interaction

At auroral latitudes, when high frequency O-mode radio waves are transmitted anti-parallel to the magnetic field, the electromagnetic wave energy is dissipated by the ionospheric plasma, and electrostatic plasma waves are artificially excited (Bernhardt et al., 2016).

Most of the transmitted radio wave energy is dissipated within a small altitude region, referred to as the *resonance region* (Gurevich, 2007). The altitude of the resonance region is determined by the height where the transmitted radio waves are in resonance with naturally occurring ionospheric plasma waves. Assuming a cold plasma and heavy (stationary) ions, the upper border of the resonance region is defined by the height where the transmitted radio waves with frequency (f_0) are in resonance with the *Langmuir* waves with frequency (f_p):

$$f_p(z) = \frac{1}{2\pi}\sqrt{\frac{n_e(z)q_e^2}{\epsilon m_e}} \tag{1.7}$$

Where (z) is the altitude, $n_e(z)$ is the altitude dependent plasma density, (m_e) is the electron mass, (q_e) is the elementary electron charge and (ϵ) is the vacuum permittivity. Langmuir waves characterize the natural oscillations of electrons around their equilibrium position in a plasma. In a magnetized plasma, as in the iono-

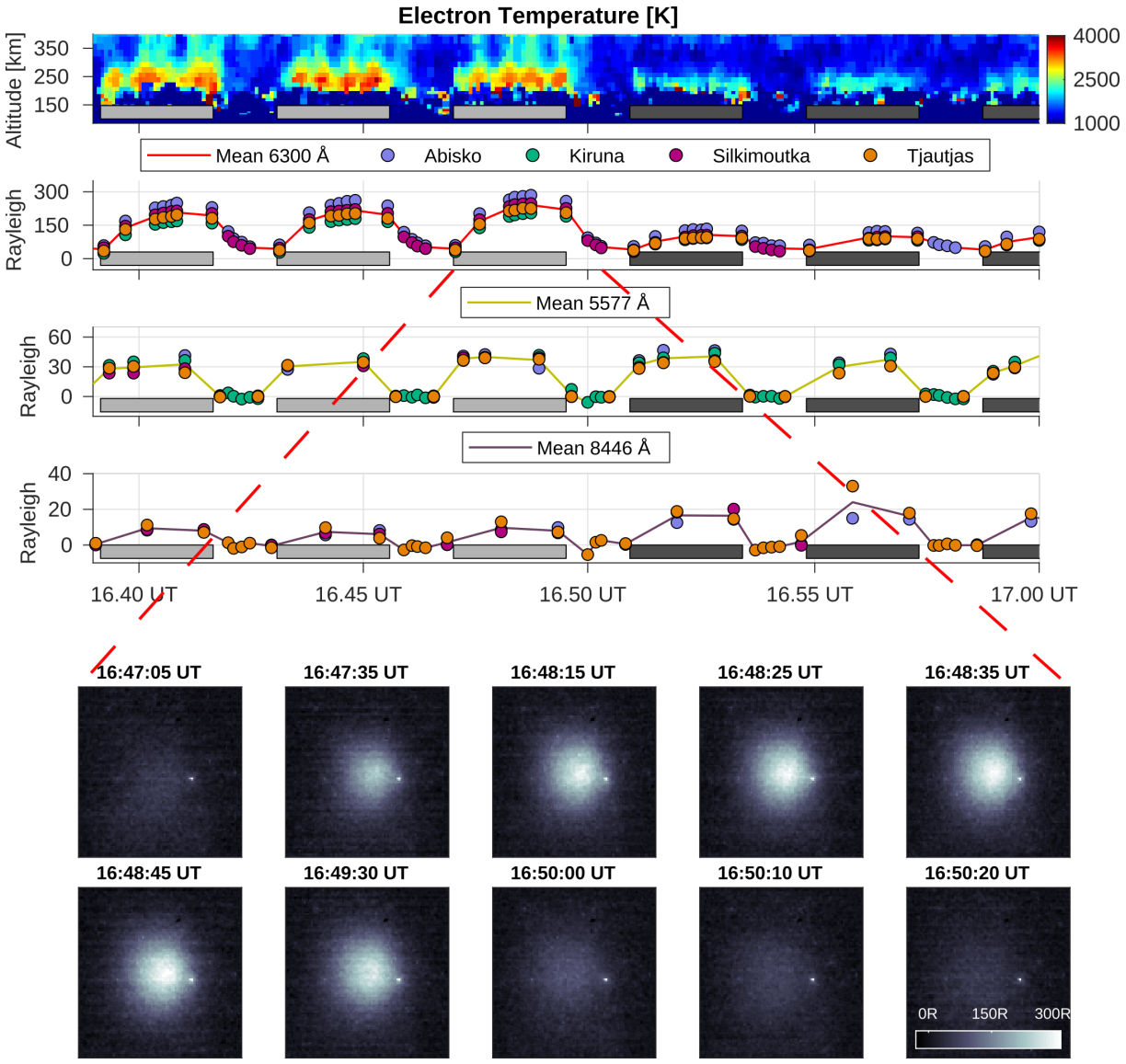


Figure 1.10: Results from an ionospheric modification experiment at EISCAT. The EISCAT HF transmitter, the EISCAT UHF IS radar and four ALIS imaging stations were operated simultaneously during this experiment. The EISCAT Heater transmitted radio waves in a pulsed modulation. The periods with ongoing transmission are indicated by the gray boxes, where the light gray boxes indicate 6.200 MHz transmission and the dark boxes indicate 5.423 MHz transmission. The IS radar electron temperature data is presented in the top panel and the average intensities in 6300 Å, 5577 Å and 8446 Å (within the heating patch pixels from images taken in Abisko, Kiruna, Silkimoutka and Tjautjas) are presented in rows 2–4. A distinct response in the electron temperature is observed and enhanced auroral emissions are seen during periods with ongoing HF transmission. Magnified and background subtracted 6300 Å images from the Abisko imaging station are presented at the bottom. Note that the rather slow growth and decay of the 6300 Å enhancement is due to the long radiative lifetime of the excited $O(^1D)$ state. The polar star can be seen as a bright speckle just right of the artificial aurora patch.

sphere, the electrons oscillate along the magnetic field with the Langmuir frequency (Chen et al., 1984).

The lower border of the resonance region, typically located 2–10 kilometers below the upper border, is determined by the altitude where the the transmitted radio waves are in resonance with the *Upper Hybrid* (UH) waves (Gurevich, 2007). The UH waves characterize the electron oscillation perpendicular to the magnetic field with frequency:

$$f_h(z) = \sqrt{f_p^2(z) + f_e^2(z)} \tag{1.8}$$

Where (f_e) is the electron *gyro-frequency*, the frequency of the circular motion of a charged single-particle in a magnetic field:

$$f_e(z) = \frac{|q_e| B^E(z)}{2\pi m_e} \tag{1.9}$$

Where (B^E) is the altitude dependent geomagnetic field strength. A frequency-altitude diagram, illustrating the relationship between the transmitted frequency (f_0), the Langmuir frequency (f_p), the upper-hybrid frequency (f_h) and harmonics of the gyro-frequency (f_e), is presented in Figure 1.11.

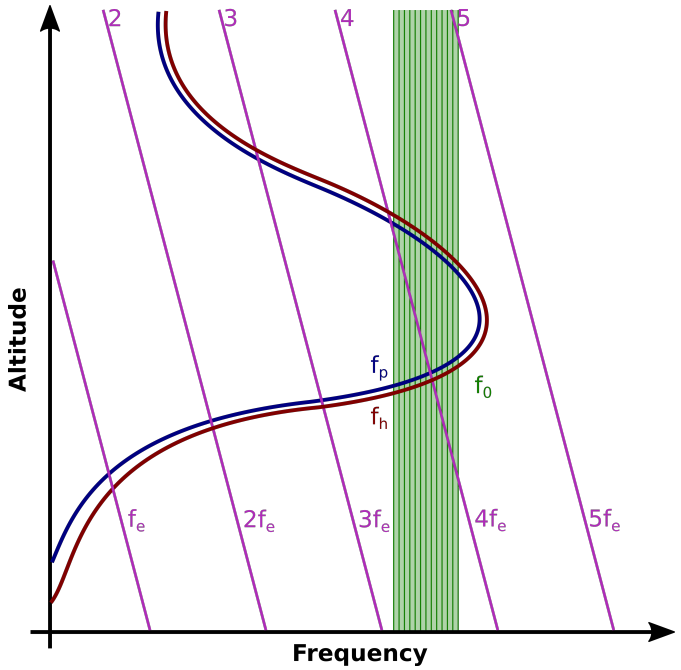


Figure 1.11: An illustration of the frequency-altitude relationship between the transmitted electromagnetic waves (f_0) and the essential electrostatic plasma waves; the Langmuir (plasma) frequency (f_p), the upper-hybrid frequency (f_h) and the elctron gyro-frequency (f_e). The strong wave–plasma interaction occurs generally within the resonance region, between the bottomside altitude where ($f_0 \approx f_u$) and ($f_0 \approx f_p$). Note that the transmitted wave frequency (f_0) can be adjusted, as indicated by the range of green frequencies, altering the height of the resonance region. Illustration credit; Theresa Rexer.

The ionosphere is often approximated as a layer of plasma enveloping the Earth with an altitude dependent density profile and a locally uniform density within

each infinitesimal height shell. However, natural perturbations within the plasma give rise to small-scale *field aligned striations* with depleted plasma densities (Kelley et al., 1982), illustrated in Figure 1.12. The field-aligned striations causes significant refraction of the electromagnetic waves, when the transmitted radio waves reach the resonance region and travel through the nonuniform plasma.

The refraction leads to self-focusing of the transmitted radio waves and an enhanced power flux is ducted into the depleted plasma regions (Gondarenko et al., 2005). This triggers an instability, known as the *resonance instability*, where the striations expand due to the increased power flux, thus leading to enhanced flux focusing and continued depletion growth (Gurevich, 2007). The small-scale striations with sizes $\sim$1 m perpendicular to the magnetic field will expand to $\sim$100–1000 m sized plasma cavities during the first 10–30 seconds of HF transmission before stabilizing, while the size along the magnetic field is $\sim$10 kilometers (Coster et al., 1985; Milikh et al., 2008). After that point, close to 100 % of the transmitted power flux is dissipated within the resonance region, provided that the transmitted power flux exceeds $\sim$30 μW/m^2 (Senior et al., 2012).

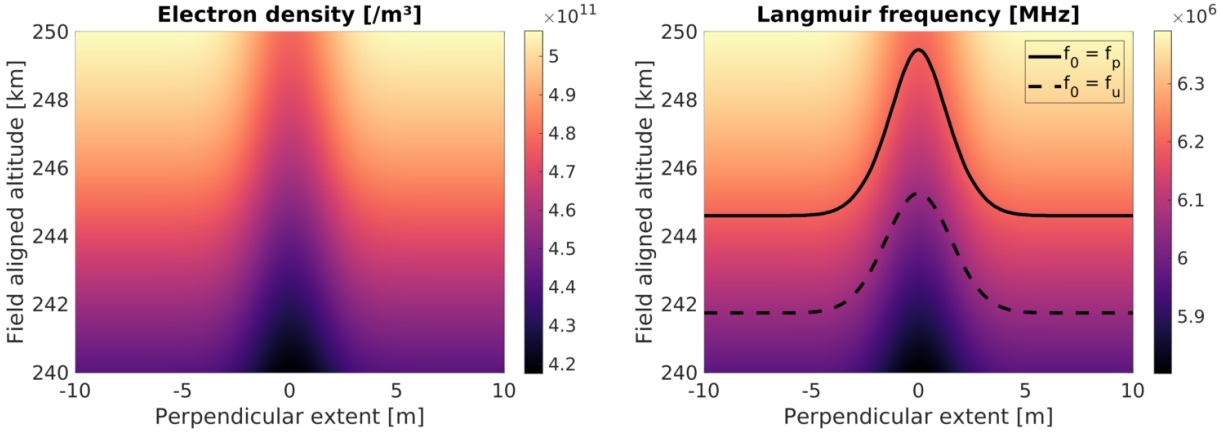


Figure 1.12: To the left, an illustration of the spatial profile of a field aligned plasma striation. To the right, the plasma striation in the Langmuir frequency space with the resonance height marked for a 6.200 MHz pump wave. Equation 1.7 is used to convert electron density to Langmuir frequency. The upper border of the resonance region is defined by the height where the transmitted radio waves are in resonance with the Langmuir waves (indicated by the top line). The lower border is determined by the height where the radio waves are in resonance with the upper hybrid waves (indicated by the dashed line). Note that the anti-parallel magnetic field direction (y-axis) is given in kilometers while the spatial extent in the plane perpendicular to the magnetic field (x-axis) is given in meters.

Observations of the IS radar spectrum during the first $\sim$100 milliseconds after HF transmission onset indicate that the resonance between the HF radio waves and the Langmuir waves dominate during this period (Djuth et al., 2004). Furthermore, simulations of electromagnetic waves in a slightly perturbed plasma suggest that Langmuir waves induce electric fields within the striations that efficiently energizes electrons, triggering the resonance instability (Gondarenko et al., 2005).

As the small-scale striations expand into larger cavities, the resonance between the transmitted radio waves and the upper hybrid waves generally becomes the dominant resonance process. Moreover, when the UH waves are excited within the cavities, *parametric decay instabilities* are activated (Najmi et al., 2016). The parametric decay instability is a cascading effect, where the primary UH waves are mode converted into secondary UH waves and lower hybrid waves, followed by excitation of electron Bernstein waves (Najmi et al., 2017).

It is beyond the scope of this thesis discuss the secondary waves and the associated mode conversion of the electrostatic waves. However, it is important to note that different electrostatic waves induce different compositions of electric fields within the cavities. These electric fields then energize the local electrons, and their resulting energy distribution is highly dependent on the population of electrostatic waves within the cavities (Najmi et al., 2017).

Energization of Ionospheric Electrons

In the recent review article on ionospheric modification experiments, Streltsov et al. (2018) stated that one of the outstanding issues for future ionospheric modification experiment research is to: *Investigate generation of supra-thermal electrons and their energy spectrum for different pump frequencies*.

It is known from observations of artificial aurora that electrons are accelerated to high energies ($>$10 eV) during ionospheric modification experiments (Gustavsson et al., 2005). Furthermore, simulations and observations imply that the electron energy distribution is strongly dependent on the relationship between the transmitted frequency and the harmonics of the electron gyro frequency (Gustavsson et al., 2006; Najmi et al., 2017).

A pronounced supra-thermal tail forms when the transmitted radio waves interact with the ionospheric plasma close to an harmonic (n) of the electro gyro frequency ($f_0 \approx nf_e$). The relationship between the pump wave frequency and the gyro-harmonics is depicted in Figure 1.11. It is suggested that the high-energy tail forms when the electron acceleration is dominated by the electric fields associated with the UH waves (Najmi et al., 2017).

In contrast, when transmitting radio waves away from the electron gyro-harmonics ($f_0 \not\approx nf_e$), electrons are mainly energized by electron Bernstein waves, which result in an energy distribution more similar to a thermal (Maxwell-Boltzmann) energy distribution. This notion is supported by observational evidence (Gustavsson et al., 2006), and is further discussed in Paper [IV]. However, additional studies and ionospheric modification experiments are needed to form a conclusive theory for the dependence of the electron energy on the HF frequency.

Chapter 2

Auroral Image Classification

> *As I write it has again begun to snow, and more thickly than a moment ago. I have been busily examining the little flakes*
>
> – Johannes Kepler, *On the Six-Cornered Snowflake*

Every snowflake and every galaxy is different. Similarly, every Aurora Borealis display is different. Nevertheless, the shapes of snowflakes, galaxies and Aurora Borealis are not random, but occur in certain patterns. Identifying and defining these reoccurring patterns, hereafter called classes, is the goal of morphological classification. The motivation for morphological classification is to understand the underlying conditions for formation and the generalized properties of the classes. Morphological classification is widely used for sorting objects, such as snowflakes by the Nakaya diagram in Figure 2.1, stars by the Hertzsprung–Russell diagram and galaxies by the Hubble sequence. These classification schemes have contributed to the science within the respective disciplines and have made it easier to communicate research and educate new researchers. Yet, there is no set of definitions or diagrams that are widely accepted for sorting auroral displays. In other words; there is no consensus of how many auroral classes exists and what they are. This makes interpretation of auroral displays difficult and is also a disadvantage for communicating auroral research. Without a clear consensus of the auroral classes, there is a great risk that scientists that read an article or talk together about an auroral class (e.g. named auroral breakup) have a different mental perception of the class. Naturally, this will lead to confusion or disagreement at some point and thus limit the exchange of knowledge. Aurora classification is therefore an inherently important part of auroral research.

Furthermore, after the breakthrough of the deep learning area, the value of labeled data has greatly increased. It has been firmly demonstrated that labeled images can be used to train deep neural networks for automatic classification of e.g. character (Wu and Chen, 2015), brain tumor (Amin et al., 2018) or facial images (Balaban, 2015). However, labeled auroral images have not been widely used for deep neural network classification of auroral displays. As the field of machine learning is rapidly progressing, it is to be expected that auroral image classification will become an important part of auroral science in the years to come.

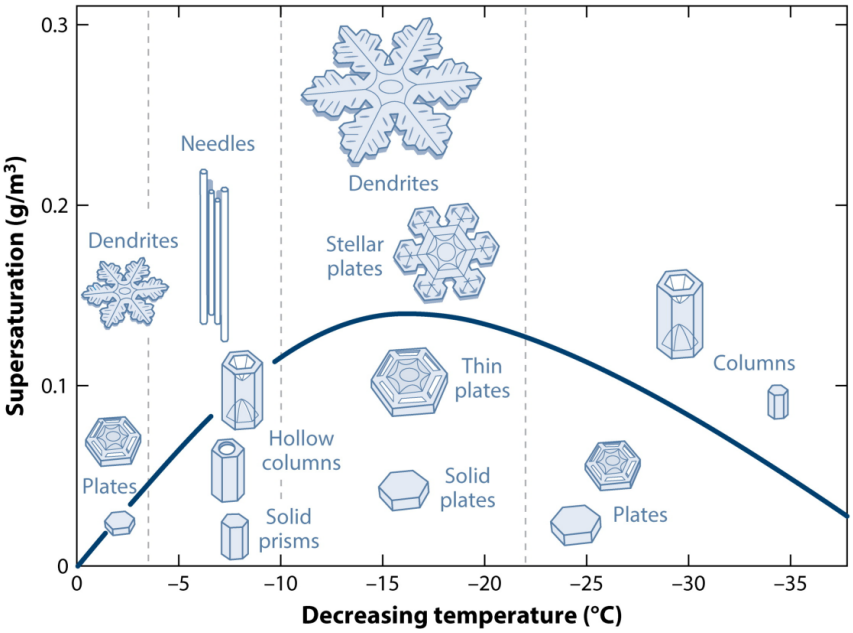


Figure 2.1: The Japanese physicist Ukichiro Nakaya made the first artificial snowflakes in the 1930s and sorted the snowflakes into a morphological diagram, known as the Nakaya diagram. Every snowflake is different, nevertheless, Nakaya revealed a general morphological dependence on temperature and supersaturation (humidity). There is no such diagram for the auroral morphology where auroral forms are related to solar wind, magnetospheric and/or ionospheric conditions. The Figure is reprinted from Libbrecht (2017), Copyright; Annual Reviews, all rights reserved.

Ideally, a machine learning classifier should be able to take any pre-processed auroral image as input and automatically output the label that best describes the depicted auroral form (e.g. as breakup, arc or patchy aurora). There are two different strategies for automatic classification using machine learning; supervised and unsupervised learning methods.

Supervised machine learning methods use manually labeled data to infer the function that maps the input (the auroral image) to the output (the auroral label). The supervised learning methods have the advantage that we can choose the labels being used by the classifier. The disadvantage is that manually labeled images are required for training the classifier and for testing its performance. Preferably a lot of images need to be manually labeled without significant contamination of biases and subjective interpretations. The needed amount of labeled data depends on the classification problem complexity, the desired accuracy, the method being used and the contamination of misclassified or ambiguous labels in the training and testing data.

In contrast, unsupervised methods do not use manually labeled data, but rather the information within the dataset (the auroral images) to divide the images into groups. The unsupervised algorithms groups images based on the similarity and dissimilarity with the other images in the dataset in a manner such that similar images are grouped together while each group is separated from the other groups.

The advantage of the unsupervised machine learning methods is that they are a more objective way of grouping data, using the data itself to group the images rather than classification based on human interpretation, which can be biased and unreliable. The disadvantage is that the resulting groups are not necessarily desirable for the classification problem at hand.

For auroral research, it is crucial the classifier groups together images that have a defined physical meaning and a scientific purpose, this can not be ensured by the unsupervised machine learning methods. Only supervised classification methods are therefore considered in this thesis. This chapter deals with the classification of all-sky auroral images. The chapter is structured didactically, from the pre-processing of the raw images to automatic classification into auroral labels. Section 2.1 presents a proposed auroral image pre-processing procedure that can be applied to to any all-sky image. In Section 2.2, the auroral morphology is discussed and appropriate auroral labels for machine learning classification are suggested. Section 2.3 deals with the extraction of auroral features from the images. Finally, Section 2.4, evaluates and describes three machine learning classification methods; K-nearest neighbor (K-NN), Support Vector Machine (SVM) and Convolutional Neural Network (CNN).

2.1 Pre-Processing for Auroral Classification

The goal of the pre-processing procedure is to standardize the input to the classifier such that similar auroral forms also appear similar in the pre-processed images, independently of the camera being used. Furthermore, the standardization reduces bias effects, makes the manual labeling easier to conduct and allows for adding together training and testing data from several studies and cameras. A description of a suggested pre-processing procedure is presented below with simultaneous image examples from two different cameras illustrating the pre-processing procedure in Figure 2.2:

1. Rotate and/or flip the images to direct geomagnetic North towards the top and East towards the left. A similar orientation will reduce bias effects from the camera orientation. In addition, the classifier might not be rotation invariant, making the orientation of the auroral form significant for the classification of an image into the auroral label. This is particularly important for orientation dependent classes, e.g. for discerning the E-W aligned arcs associated with the growth phase from the N-S arcs associated with the onset.

2. Normalize the intensity of the images in the data set so that the sensitivity is similar for all cameras and all exposure times being used. In addition, for classifiers using color (RGB) images, the spectral response can be normalized at the auroral emission lines in Red 6300 Å, Green 5577 Å and Blue 4278 Å. This requires sensor quantum efficiency information which can be obtained e.g. from Mauer and Wueller (2009). This step is important since the intensity of the auroral display is an important characteristic for many auroral forms.

3. Define the smallest square that encapsulates the entire all-sky image. The framing step is illustrated by the blue box in Figure 2.2. The square cropping ensures that the image ratios are similar, reducing bias effects and further increasing the robustness of the classifier.

4. Resize the squared image to (256×256) pixels. This step is necessary since most classifiers only accept an image of a certain size. In addition, for training/testing purposes, it is usually not desirable to use the full resolution images since these often are unnecessarily large and will cause a slow convergence of the classifier.

5. Filter the image using a (3×3) 2–D median filter. The median filtering reduces bias effects from the location of stars, remove bad pixels and reduce noise.

6. Crop the central (128×128) pixels image. The central image cropping is marked by the red squares in Figure 2.2. The central cropping reduces the number of image frames with auroral features towards the horizon, which often look similar due to small elevation angles and fish-eye distortion. Labeling the auroral forms towards the horizon therefore do not provide any additional information and might instead contaminate statistical studies. Focusing on the auroral forms in the center of the image also reduces the ambiguity of the auroral displays and therefore eases the manual labeling. Furthermore, the cropping reduces light pollution and cloud contamination which often is seen only at low elevation angles.

The pre-processing procedure above can be edited in many ways, depending on the study and the available set of cameras. The pre-processing procedure should be applicable on data from most cameras without requiring a lot of a priori information of the optical system parameters. However, there is a trade-off between robustness and the amount of pre-processing required. For example, for ionosphere-magnetosphere coupling studies, it might be desirable to classify magnetic longitude-latitude projections of the auroral forms. In addition, the longitude-latitude projection will correct some of the camera-dependent fish-eye distortions and allows for adding together classified images from a camera array system in a mosaic pattern. The projection step can be fitted between step 5 and 6 in the suggested pre-processing procedure. Note however that the longitude-latitude projection requires the field-of-view transformation matrix and an estimate of the auroral altitude, which might not be known in advance. An example of a longitude-latitude projection is shown in Figure 2.3, and the projection technique is further discussed in Chapter 3.

Note also that sky condition software that discerns images with auroral conditions and images that are polluted by clouds already exist, see sky condition classification for color images in Rao et al. (2014) and for black-and-white images in Clausen and Nickisch (2018). Furthermore, calculators for celestial objects, e.g. Rhodes (2019), provides accurate estimates of the Sun and Moon positions, thus images during twilight or moonlit periods can be automatically excluded at a pre-processing stage.

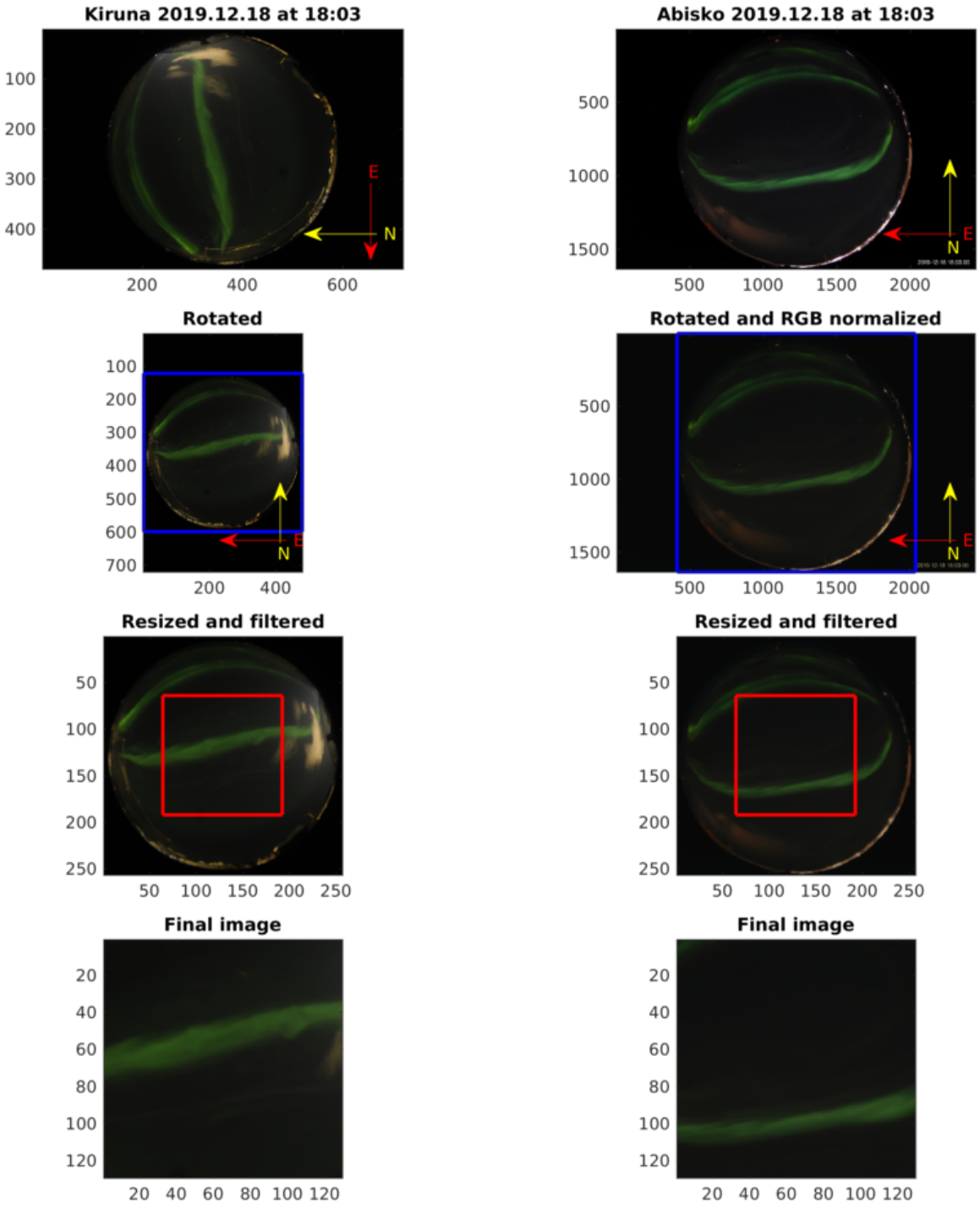


Figure 2.2: Demonstration of the pre-processing procedure with two different cameras, at Kiruna to the left and Abisko to the right. After the pre-processing, the output images have standardized orientation, pixel sensitivity and resolution. Training and testing of the classifier on the pre-processed images will enhance the overall performance and the robustness of the classifier. The red and yellow arrows indicate the image orientation with respect to the geomagnetic pole. The blue and red boxes indicate image cropping frames that are used during the pre-processing procedure. The Kiruna camera is operated by the Swedish Institute of Space Physics (IRF) while the Abisko camera is operated by the Hiroshima City University, Radio Communication and Science Laboratory, in collaboration with IRF.

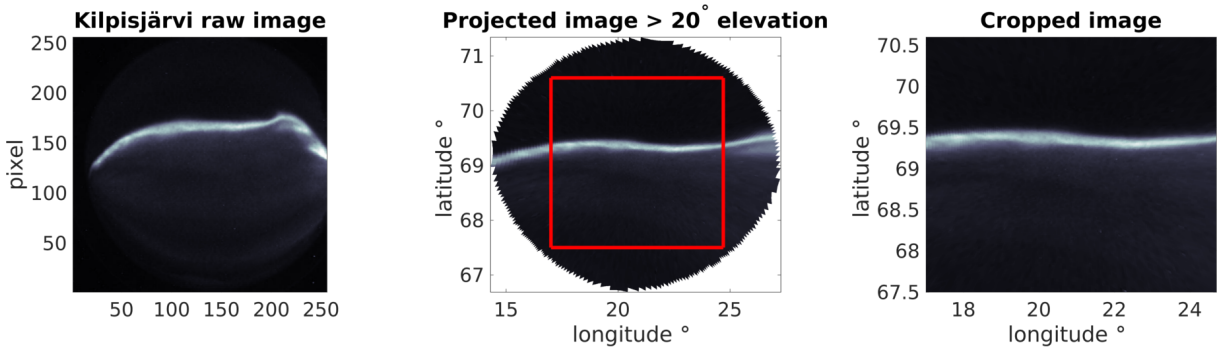


Figure 2.3: The figure illustrates an alternative pre-processing procedure where an all-sky image from Kilpisjärvi, Finland, is projected to a magnetic longitude-latitude shell at an altitude of 110 km, then cropped according to the red box. The additional projection step can be fitted between step 5 and step 6 in the suggested pre-processing procedure. Note however that the final step (the central image cropping) now must be defined by a longitude-latitude range or by the biggest central box that can be fitted within x° from zenith. Also, note that only the image pixels $> 20°$ is projected to the magnetic longitude-latitude shell in the figure.

Sky-condition classification and aurora classification during twilight and/or moonlit periods is therefore not treated in this thesis.

2.2 Labels for Auroral Image Classification

The supervised machine learning methods require a set of labeled images both to train the automatic classifier and to test its performance. The labeled data defines the mapping function from the image features to the associated label. Thus, great care is needed for defining appropriate labels. Each label should meet a set of criteria:

1. The auroral labels should be scientifically useful. This promotes labels that are associated with physical properties and characteristic production mechanisms. In addition, it is advisable to use auroral labels that have been previously studied, preferably by both imagers and satellites, and therefore have a broader appeal and scientific foundation.

2. All labels should be common with at least $\sim$1 % of the auroral images fitting into each label. The $\sim$1 % criteria ensures a somewhat balanced distribution of labeled images, which is important for efficient and reliable training of the classifier. Furthermore, the $\sim$1 % criteria promotes labels that describes the generic substorm, rather than auroral forms that are based on case studies.

3. All labels should be clearly defined and easy to recognize under different geomagnetic conditions and at different elevation angles. This reduces the label ambiguity and makes the manual labeling more reliable. Furthermore, elevation angle independence makes statistical studies easier to conduct.

In addition, it should be noted that the observed auroral displays are highly dependent on the field-of-view. This is clearly depicted in Figure 2.4. Generally when studying the global auroral substorm, the larger auroral forms are of interest and the small scale structures might not be relevant for the global dynamics. On the other hand, the narrow field-of-view cameras have been used for studies of for example the electron energy distribution of small-scale auroral structures (Dahlgren et al., 2008) and the the generation mechanism of flickering aurora (Whiter et al., 2010), which requires high spatial and temporal resolution. The work presented in this chapter and in Papers [I] and [II] focuses on images acquired by all-sky cameras and auroral classes that describes the larger auroral forms without including characterizations based on temporal behavior. However, the techniques presented here could be extended to include the temporal dynamics.

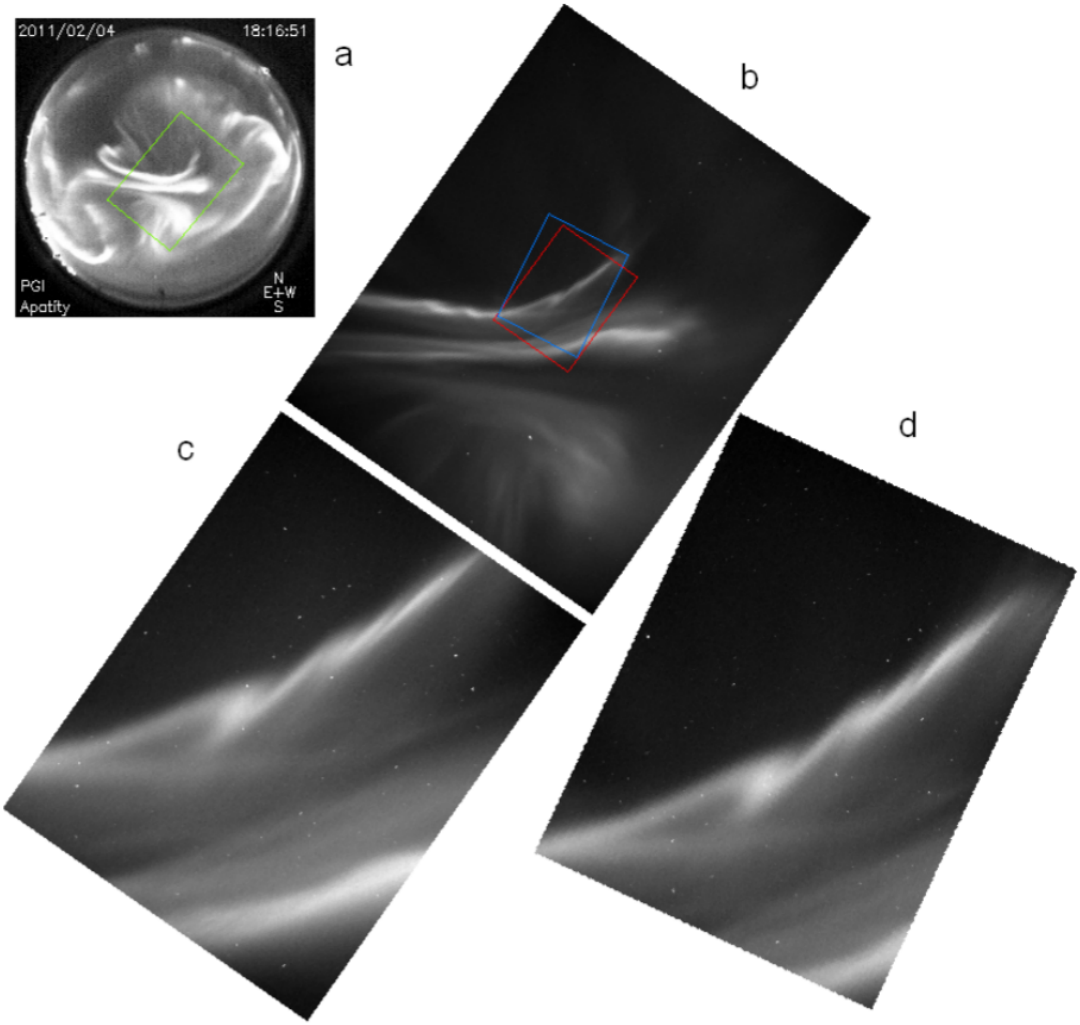


Figure 2.4: Simultaneous auroral images from 4 cameras in Apatity, Russia. The different fields-of-view reveal small scale drape structures within seemingly smooth auroral structures in the all-sky image. a) is an image captured by an all-sky camera, b) is taken by a 67° diagonal field-of-view camera while c) and d) are acquired by cameras with 18° diagonal field-of-view. The Figure is reprinted from Kozelov et al. (2012), under the Creative Commons Attribution 3.0 License.

Essential Auroral Labels

Hitherto, most of the night-time auroral classification studies have all used different pre-processing procedures and auroral labels, making it difficult to compare the results. In addition, it has not been common for the night-time auroral classification studies to use labels with a clear physical interpretation. In order to obtain information about geophysical conditions from auroral images, it is crucial to be able to associate the auroral displays to physical processes or conditions. Future studies should therefore opt for using auroral labels with a clear characteristic of each class and an associated physical meaning. In addition, accumulation of training and testing data from several studies and cameras could be achieved if a common pre-processing procedure and similar auroral labels were used. Accumulation of big training and testing sets from several studies will improve the applied machine learning tools and accelerate the progress in the field of auroral image classification.

After analysis of magnetometer data in relation to the auroral substorms (McPherron, 1970), most auroral studies have agreed upon that 3 substorm phases can be identified in an isolated substorm. Furthermore, simultaneous image and satellite observations have shown that each phase has characteristic auoral displays and generation mechanisms, e.g. (Akasofu, 2013; Donovan et al., 2008; Sergeev et al., 2011; Zou et al., 2010). In the simplest and most idealized case, for an observer located at high latitudes during an auroral substorm: An isolated auroral arc in the evening sector is associated with the growth phase; Auroral breakup (bright auroral bulge, coronal aurora and auroral spirals) near magnetic midnight is associated with the expansion phase; and patchy aurora in the morning sector is associated with the recovery phase (Akasofu, 2012). It is therefore natural to include auroral *arcs*, *breakup* and *patchy* to any labeling set as essential labels in addition to *faint aurora/clear sky*, when there is a quiet ionosphere-magnetosphere coupling along the field lines overhead. The essential labels meet all three criteria as they are scientifically useful and sound (criteria 1), common and descriptive of the generic substorm (criteria 2) and clearly defined (criteria 3). The relationship between the phases, displays, Ionosphere-magnetosphere (I-M) coupling and the magnetotail processes are summarized in the Table 2.1. Typical sample images of each auroral label are shown in black-and-white in Figure 2.5 and in color in Figure 2.6.

Successful classification of millions of auroral images even into just 4 classes (*arcs*, *breakup*, *patchy* and *faint aurora/clear sky*) will become highly useful for auroral researchers. Most importantly, labeled images will exclude the bulk of images that are not of interest for a particular auroral study and thus reduce the needed time for tedious image inspection. Note that a working auroral image classification tool is also useful for studies of auroral forms which are not included in the set of labels. For example, Lumikot (discussed in Paper [III]) are believed to only occur near growth-phase arcs, a labeled set of images therefore eases the data analysis by making it possible to only look for Lumikot in the images classified as arcs. In addition, a large labeled data set will facilitate statistical auroral studies, auroral evolution studies and studies for relating the auroral forms to magnetospheric and solar wind conditions.

Phase and display	I–M coupling	Magnetotail process
Quiet and Growth **Auroral arc** is characterized by auroral emissions distributed in East-West aligned structures spanning across the image. Auroral arcs are stable auroral forms that can last from a few to several minutes (Partamies et al., 2015) and span lengths up to hundreds of kilometers with widths ranging from a minimum of 50 m to a maximum of 10 km (Davis, 1978).	Auroral arcs forms by quasi-static particle acceleration in a region close to the ionosphere, and they magnetically map to the plasma sheet. Simultaneous satellite and ground imaging observations have related the stable auroral arc to a characteristic V-shaped electron energy spectrum (Colpitts et al., 2013), these arcs are therefore often called inverted-V aurora (Karlsson et al., 2020). The arcs run parallel to the auroral oval and the magnetic latitudes.	The growth phase is characterized by energy loading. Energy is loaded into the magneotail if the solar wind flux is considerable and if the intrinsic B_z is Southward. Energy can not be dissipated significantly into the ionosphere if the conductivity is low (if B_z suddenly turned Southward) and energy is accumulated into the tail. Consequently, the tail cross section is increasing and the near-Earth plasma sheet is thinning. The magneotail is stretched, causing the auroral oval to drift equatorwards.
Expansion **Auroral breakup** is characterized by bright and large auroral forms which covers most of the image frame. Auroral breakup displays are seen as the auroral bulge is expanding and includes a variety of different features such as spirals (often referred to as the Westward traveling surge) Elphinstone et al. (1996) and coronal aurora (Akasofu, 2012).	The sudden brightening of the most equatorward arc is considered as the initiation (the onset) of the auroral breakup. The onset is hypothesized to be triggered by the formation of a substorm current wedge, a near-Earth ($\sim$8 R_E) current disruption, linking the cross-tail current to the ionosphere (Lui, 2011). Various plasma flows and instabilities play a key role in the substorm current wedge formation, the process is however not fully understood and an active field of research.	When the accumulated energy reaches $\sim 5 \times 10^{22}$ ergs^{-1} and the conductivity in the ionosphere is high, plasma instabilities are developing and reconnection is triggered in the near-Earth magneotail (Akasofu, 2013). The energy is released into the ionosphere in a pulse-like manner by ejection of fast plasma flows. The plasma sheet is thickening and the tail cross section is decreasing after the initial reconnection bursts and as the expansion phase evolves.
Recovery **Patchy aurora** is characterized by diffuse aurora consisting of irregular shapes which cover large portions of the image. The intensity of the auroral emission is weak and associated with pulsating aurora. Each patch has a typical extent of tens of kilometers and a pulsation period of $\sim$10 seconds. (Lessard, 2012)	The diffuse patches are caused by pitch-angle scattering of energetic electrons to the ionosphere. THEMIS satellite observations have shown that whistler mode chorus waves in the magnetosphere, at 8–12 R_E, are correlated with pulsating aurora and thus also patchy aurora (Li et al., 2011).	The recovery phase is initiated by the disruption of the large-scale magnetic reconnection as the plasma sheet is thickening and the magnetosphere again is approaching a stable configuration. The ionospheric conductivity is however high and the magnetosphere continues to release energy into the ionosphere with a rate of $\sim 10^{18}$–10^{19} ergs^{-1} (Akasofu, 2013).

Table 2.1: The magnetospheric processes associated with the substorm phases and the essential auroral classes. The table is inspired by the labels in Paper [I] and descriptions in Akasofu (2013) and in Lui (1991).

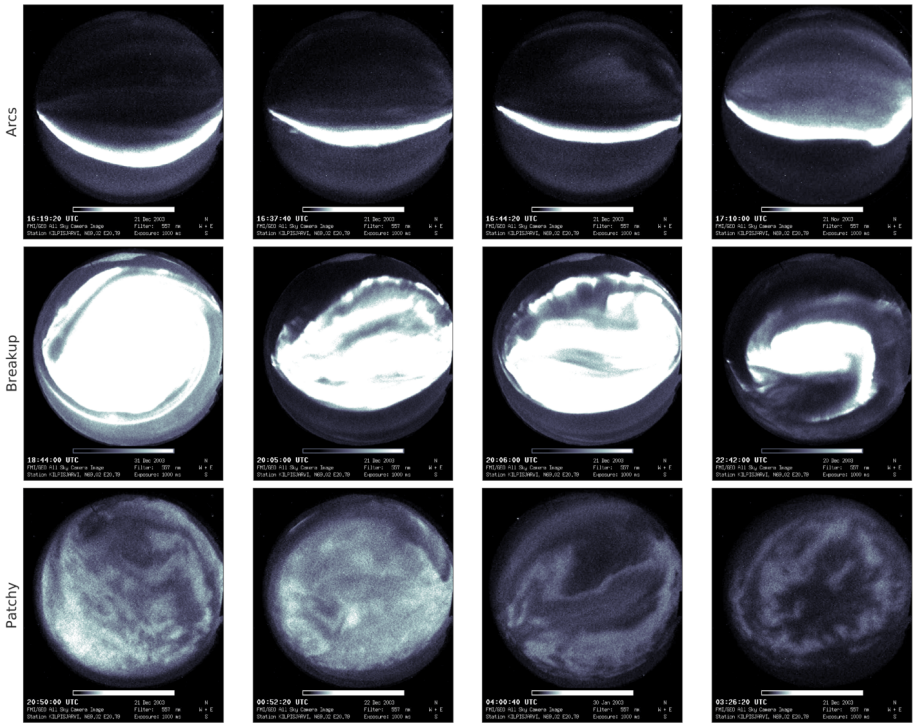


Figure 2.5: Black-and-white sample images from each auroral label (*arcs*, *breakup* and *patchy*). The all sky images are acquired by the MIRACLE imager near Kilpisjärvi, Finland (Sangalli et al., 2011b). Note that the intensity scaling in the breakup images is not the same as for the arcs and patchy images.

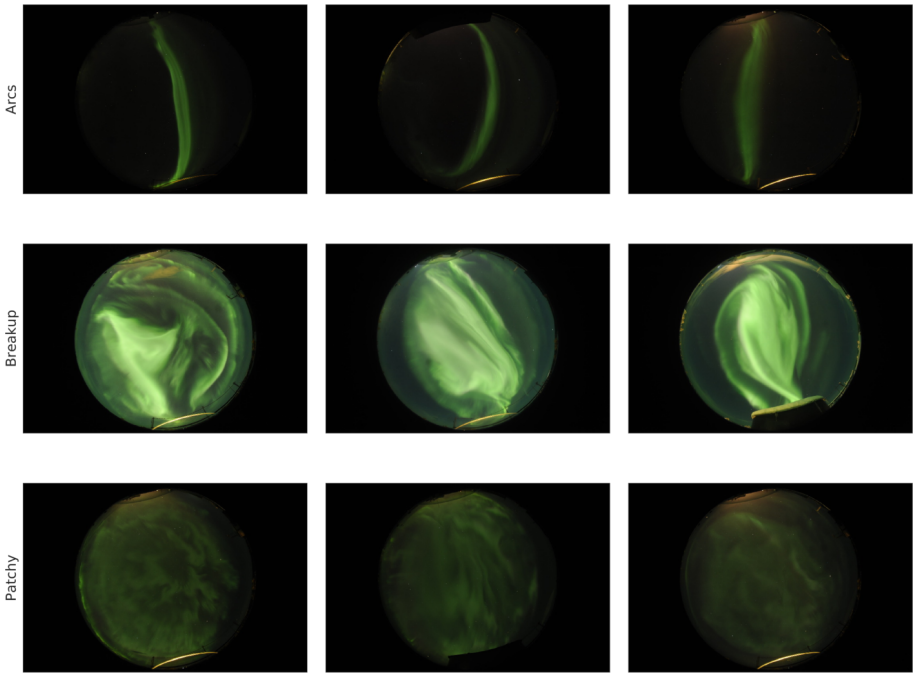


Figure 2.6: Color images from each auroral label (*arcs*, *breakup* and *patchy*). The all-sky images are taken in Kiruna, Sweden, and operated by the Swedish Institute for Space Physics. These images are among the raw images that were used for training and testing the classifiers in Paper [I].

Additional Auroral Labels

It is clear to those who have studied auroral images that most auroral displays do not fit into the canonical labels in Figures 2.5 and 2.6. This indicates that the displays shown in Figures 2.5 and 2.6 and the associated mechanism described in Table 2.1 does not fully represent the auroral activity in Arctic regions. Note that the essential labels in Table 2.1 primarily describe the evolution of nighttime aurora along the auroral oval. However, aurora also appears outside the nighttime auroral oval, but often then with a different morphology.

Dayside aurora, seen in polar regions during the winter season, is not driven by magnetotail processes, but is directly driven by the solar wind interaction with the dayside magnetosphere (Frey et al., 2019). The different driving mechanism is visually echoed in the dayside auroral display, which is dominated by drapery, radial and hot-spot aurora, as well as auroral arcs (Niu et al., 2018). Polar cap aurora is caused by precipitation along the open field lines at the polar cap and often appears as Sun-aligned arcs. Polar cap arcs are associated with a Northward interplanetary magnetic field and low magnetospheric activity (Reidy et al., 2018), completely in contrast to the auroral activity along the auroral oval. At subauroral latitudes, auroral forms such as the Stable Auroral Red (SAR) arcs and the newly reported Strong Thermal Emission Velocity Enhancement (STEVE) arcs can be observed (Gallardo-Lacourt et al., 2021). SAR arcs originates from thermal heating of atomic oxygen in the subauroral ionosphere as the inner edge of the ring current interacts with the plasmasphere during geomagnetic storms (Mendillo et al., 2016). The production mechanism of the STEVE arcs, depicted in Figure 2.7, is still an unresolved problem. However, simultaneous satellite and image data indicate that STEVE arcs are not caused by precipitation of energetic particles, but is produced locally in the ionosphere, and should therefore not be classified as aurora, but as an exotic ionospheric emission phenomena (Gallardo-Lacourt et al., 2018; MacDonald et al., 2018).

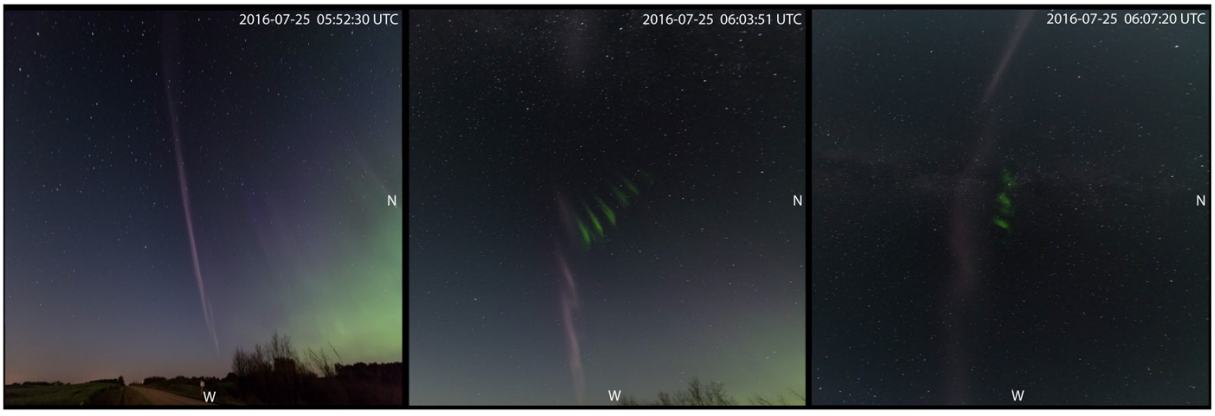


Figure 2.7: The STEVE phenomena is characterized by a thin East-West aligned arc with purple color at subauroral latitudes. The images are taken near Regina, Saskatchewan, Canada by Notanee Bourassa, and published in MacDonald et al. (2018). The Figure is reprinted under the Creative Commons NonCommercial license.

In this thesis, the auroral morphology associated with nighttime at the auroral oval is in focus. There are several labels that could be considered in addition to or as sub-labels to the labels described in Table 2.1. In this chapter, the labels used in Paper [I] are presented, discussed and employed as an example of an auroral label set that can be used for machine learning classification of auroral images. Thus, in addition to the 3 classes described in Table 2.1 and the *faint aurora/clear sky* class, the labels; *colored aurora*, *discrete aurora* and *edge aurora* were included after inspecting thousands of both black-and-white and color auroral images and after several iterations of trial and error with various label sets. A description of the display, physical meaning and comments on the scientific use of the additional labels are presented in Table 2.2. The labels described in Table 2.2 satisfy the criteria (1–3), listed in the introduction to this section. Typical examples of the auroral labels used in Paper [I] are presented in Figure 2.8.

Notice that the *auroral arcs* label includes multiple East-West aligned structures (sometimes refereed to as arc systems) in contrast to the single growth phase arc depicted in Figures 2.5 and 2.6. Furthermore, note that each of the additional labels include several sub-labels that have a more distinct physical interpretation. However, the data set that was used in the classification study reported in Papers [I] and [II] did not allow for further sub-divisions while maintaining an acceptable amount of training/testing data within each label. Further sub-division could be achieved by expanding the data set. An additional data inquiry was however not conducted in the initial study as time was limited and the the main goal was to demonstrate that machine learning methods can be used for reliable automatic auroral image classification, not to perform an extensive manual labeling study.

Auxiliary Auroral Labels and Bias Control

Even with the inclusion of the additional labels, most auroral images still do not fit clearly into any label or are a mixture of several labels. These images are ambiguous, and the appropriate label is a matter of subjective interpretation. The labeling of ambiguous auroral images is further complicated since there is no consensus within the auroral research community on how many auroral labels exist and what they are. Thus, it is unreasonable to expect the classifier to label these images correctly if not even auroral experts can agree on what the correct label is. Ambiguous images should therefore be omitted from training and testing of the machine learning classifiers.

In Paper [I], ambiguous auroral images were labeled as *unknown-complicated* and excluded from the training and testing data sets. Furthermore, the initial preprocessing stage did not exclude all images with clouds, light pollution and noise. These images were labeled as *rejected* and also excluded from the training and testing data. Sample images labeled as *unknown-complicated* and *rejected* are presented in Figure 2.9.

An additional step that can be added to reduce bias effects is to manually label the entire data set independently multiple times and only use the agreeing labels for training and testing. This procedure was implemented in the study presented in

Display	Physical description	Scientific Usage
Colored aurora An image is classified as colored aurora if the aurora, of any shape and form, has a prominent red, blue or purple emission. Colored aurora often appear as auroral rays, drapes or curled structures (see row 2, column 3 and 4), but also as East-West aligned structures (see column 1 and 2)	Distinct colored aurora occurs when the electron energy distribution has a pronounced low or high energy tail, changing the electron penetration depth into the ionosphere (Fang et al., 2008). Blue and purple auroral displays are usually seen when electrons penetrate deeper into the ionosphere, as compared to electrons causing the green aurora. Red aurora, however, is produced at higher altitudes and characterizes lower energy electron precipitation. In addition, ionospheric conditioning, such as auroral precipitation during a sunlit ionosphere, is known to produce distinct colored emissions (Shiokawa et al., 2019; Størmer, 1929)	The typical precipitation energy and the atmospheric composition in the altitude range of 90–130 km causes green (5577 Å) to be the dominant color in most auroral displays. As a result, most auroral cameras have a narrow band green filter that neglects the morphology of the distinct colored aurora. Studies of distinctly colored emissions might therefore be underrepresented, as indicated by the newly reported STEVE phenomena (MacDonald et al., 2018). Furthermore, distinctly colored aurora indicate precipitation with a skewed energy distributions or ionospheric conditioning, these auroral displays might be characteristic of traceable and interesting production mechanisms.
Discrete aurora Auroral emission appears as vortical structures, broken arcs and North-South aligned arcs or a combination of several discrete shapes. Many of these shapes are seen in association with highly dynamic aurora, see Colpitts et al. (2013) and pre-onset auroral forms, in particular Poleward Boundary Intensification (PBI) (Mende et al., 2011). Discrete-irregular auroral forms are not as bright and not as large as auroral breakup forms.	Simultaneous imaging and satellite observations have related discrete aurora with electron precipitation triggered by Alfvén waves in the magnetosphere (Colpitts et al., 2013). Furthermore, PBIs are hypothesized to be enhanced as plasma flow bursts from the open polar cap field crosses the nightside polar cap boundary into the closed field region (Lyons et al., 2011). Some of the PBIs develop into North-South aligned arcs, associated with channels of enhanced earthward flows within the plasma sheet (Nishimura et al., 2010).	Discrete-irregular aurora is of particular interest for studies of precipitation acceleration processes and triggering mechanisms of auroral onset. Nishimura et al. (2010) proposed a time sequence leading to auroral onset where the equatorward drift of the North-South aligned arcs (also called streamers) occurs prior to the auroral onset, suggesting that the onset is triggered after the streamer reaches the equatorward arc. Note however that not all streamers triggers an onset as plasma sheet conditions must be opportune (Lyons et al., 2011).
Edge aurora Images with auroral emission occurring only at the edge of the image are labeled as edge aurora. There is no particular form or shape associated with Edge aurora	Edge aurora occurs when there is precipitation along the field lines that are projected towards the edge of the image, there is no characteristic auroral appearance related to edge aurora and therefore no associated production mechanisms.	Edge aurora was defined as an independent class in order to make the classifier more robust since information about the depicted auroral forms is limited and there is a significant uncertainty of the auroral form outside the image frame. However, the location of the aurora in a time series of images is often valuable information for determining the auroral drift. Edge aurora was therefore added as an additional label.
Faint aurora/clear sky Dark images without a clearly visible aurora are labeled as faint aurora/clear sky.	Faint aurora/clear sky images are associated with a very weak electron precipitation.	Images without auroral emission indicates a quiet ionosphere–magnetosphere environment along the field lines overhead.

Table 2.2: An example of labels that can be used for classifying color auroral images.

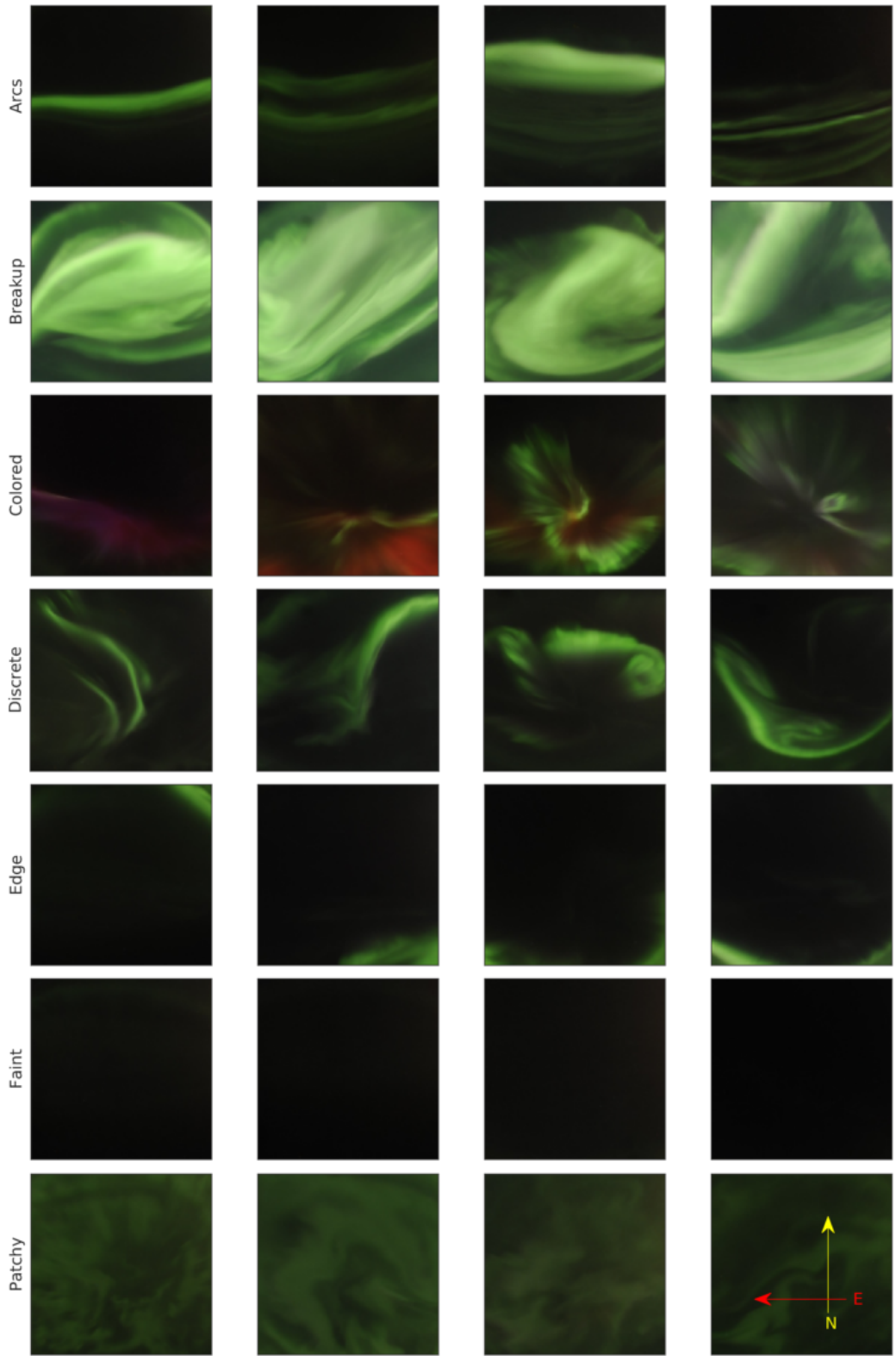


Figure 2.8: Four sample images of each auroral class. The auroral images have been pre-processed according to the 6-step pre-processing procedure in Section 2.1. The direction with respect to the magnetic pole is indicated by the arrows in the bottom right.

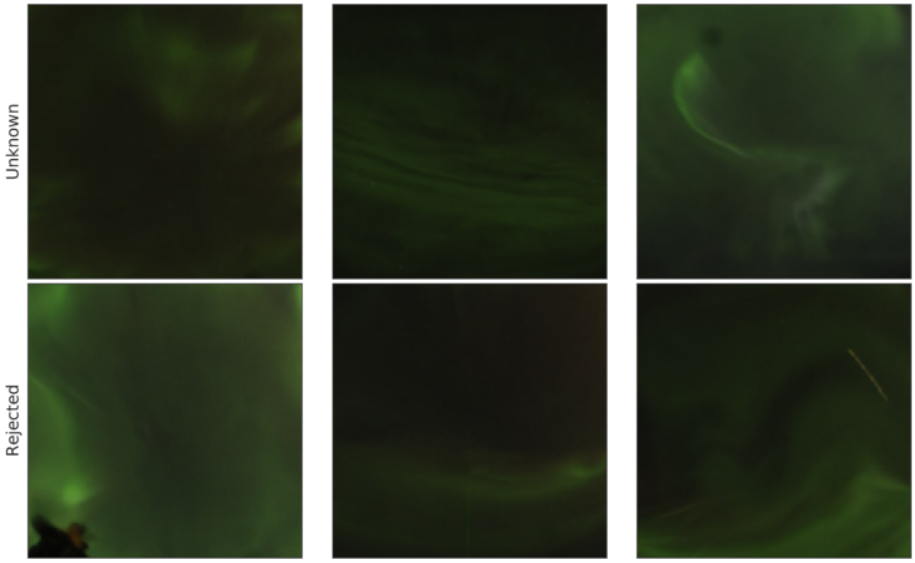


Figure 2.9: Sample images of excluded auroral images in Paper [I]. The top row illustrates aurora with ambiguous form labeled as *unknown-complicated*. The bottom row shows images labeled as *rejected*, these images were excluded from the training/testing data due to unwanted features, from left to right; a person in the bottom left corner, lidar emission and a passing aircraft. Note that these images have been pre-processed by the 6-step pre-processing procedure in Section 2.1.

Paper [I], and discussed in detail in Paper [II], where expert *One* labeled the images in consecutive order while expert *Two* labeled the images independently in random order. Only the agreeing labels of expert *One* and *Two* were considered to have an unambiguous auroral form suitable for training and testing of the supervised classifiers.

In the end, although 14030 auroral images were initially considered, only $\sim$4050 images had an auroral form that was considered unambiguous by both experts. $\sim$95 % (3846) of these images were labeled similarly by both experts and used as the final labeled data set. The 3846 labeled image data set was then randomly split into a training set (3000 images) and a testing set (846 images). Unless otherwise stated, the same data set is used in Sections 2.3 and 2.4 to exemplify the feature extraction techniques and the training- and testing of the supervised machine learning classifiers.

2.3 Feature Descriptors

Ideally, we want to use all the information within the labeled training images to develop a classifier. At the same time, we want to represent the information within the training images as robustly and compactly as possible in order to save computing time and to reliably train the classifier on only the significant image features. Feature descriptors are computer vision techniques that extracts the important features of an image (such as shape, texture and color) and represent these features in a vector with a reduced size, as compared to the size of the initial image. In the limiting cases; with no reduction, the full resolution image is used for training and testing, in contrast, with heavy reduction, the image is reduced to a simple feature vector with only a few global parameters that characterize the entire image, such as the mean standard deviation ($\bar{\sigma}$) and the mean intensity ($\bar{I}$) of the image pixels. In Syrjasuo and Partamies (2011), several feature extraction techniques were tested and compared for aurora detection in black-and-white all-sky images. It was concluded that global feature extraction techniques were outperformed by local feature extraction techniques. In contrast to the global techniques, the local feature extraction techniques divide the image into regions or key points and represent the entire image as a feature vector by combining the local features from all regions/key points. Note however that the appropriate feature extraction technique depends on the classification problem at hand, i.e. what part of the image information is relevant and necessary for adequately solving for the project task.

Global Feature Extraction: The Keogram Representation

In auroral science, a common global feature extraction technique is the keogram representation of images, a keogram example presented in Figure 2.10. The keogram summarizes the auroral activity of an entire day by slicing the central North-South column of the images and consecutively stacking them along the time-axis. Although the keogram representation successfully reduces the data size, (for example from an image matrix with dimension ($256\times256\times3$) to a feature vector of with length (256×3), it excludes a lot of the spatial information within the auroral images. Thus, the keogram feature extraction technique is not adequate for auroral classification into time-invariant subclasses, e.g. according to the labels described in Figure 2.8. The keogram representation might however be a good option, if the time dimension is included, for automatic classification of periods with aurora and possibly onset detection. Time-dependent auroral image classification is not the subject of this thesis, but is briefly discussed in Chapter 4.

Local Feature Extraction: The Histogram of Oriented Gradients

In this Section, the Histogram of Oriented Gradients (HOG) method, first presented in (Dalal and Triggs, 2005), is described as an example of a local feature descriptor that can be used for auroral image feature extraction. The basic idea of the HOG method is that objects in images often can be characterized by their directional gradient properties (e.g. for extracting edge and texture characteristics) and that these gradients can be represented as a collection of local histograms within image

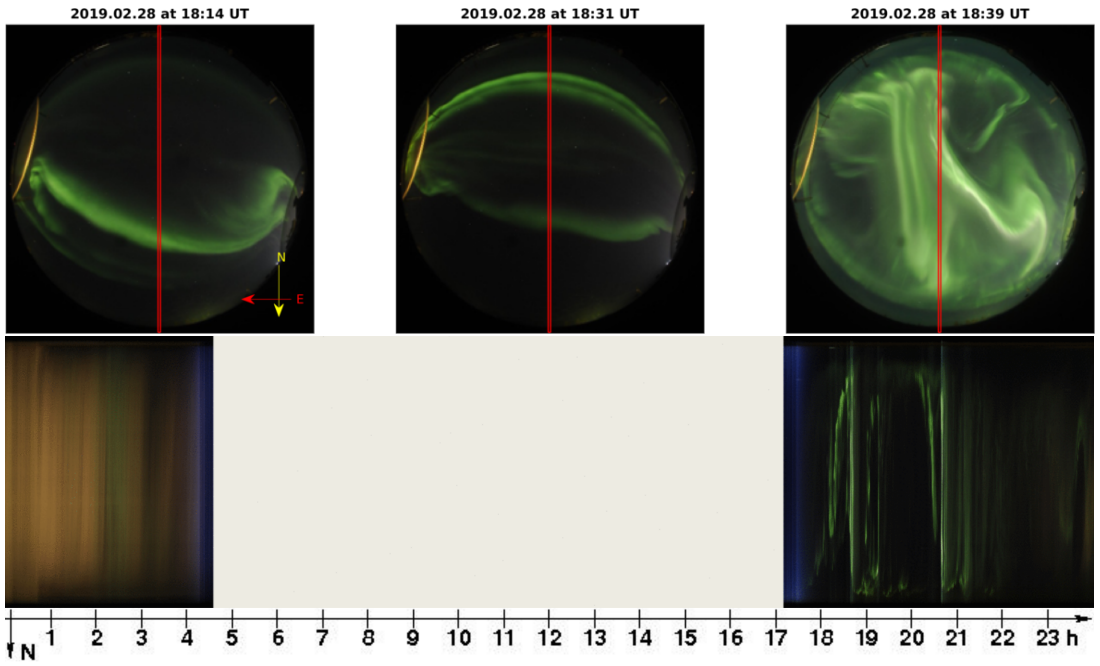


Figure 2.10: At the top, color images from the Kiruna all-sky camera. At the bottom, a keogram representation of the images acquired on the 28th of February, 2019. The red box in the all-sky images represents the N-S slices which are consecutively stacked along the x-axis (time-axis) to generate the keogram. Clouds are causing the beige color (∼00:00–04:20 UT). Twilight occurs after ∼04:20 and the gap in data acquisition between ∼04:40–17:10 is due to daylight. After ∼17:10, daylight is fading and a growth phase arc becomes visible (see all-sky image at 18:14) as it is drifting Southwards and intensifies (image at 18:31), triggering the auroral onset (breakup image at 18:39). Auroral activity continues throughout the night with a second onset at ∼20:40 UT.

regions. The HOG method was used for extracting the auroral image features in Paper [I]. Note that there are several local feature extraction techniques that could be considered, e.g. the Local Binary Pattern (LBP) method and the Scale-Invariant Feature Transform (SIFT) descriptor, both methods tested and compared in (Rao et al., 2014). However, it is beyond the scope of this thesis to present an overview of feature extraction techniques. Instead, the Histogram of Oriented Gradients method is described in detail and it is demonstrated how this method can be used to extract auroral features in all-sky images.

The HOG method is a commonly used feature extraction technique in computer vision. The descriptor was initially developed for detecting humans in images, but is a versatile tool that also has been used for various tasks such as vehicle detection (Mao et al., 2010), handwritten digit recognition (Ebrahimzadeh and Jampour, 2014) and text recognition in photographs (Bissacco et al., 2013). Furthermore, the method is not sensitive to changes in background illumination and is therefore robust for image classification during twilight, moonlit or light polluted periods, as long as the light source is not within the image frame. The HOG method can be described as a 4-step image processing procedure if no color normalization is ap-

plied:

1. **Calculate oriented gradients:** The image gradients are calculated by convolution with a simple 1–D [-1,0,1] mask producing the gradient along the x-axis (G_x) and the transposed mask provides the gradient along the y-axis (G_y) for each pixel (i, j). The magnitude G at each pixel is calculated by the $\boldsymbol{l^2}$ norm:

$$G(i,j) = \sqrt{G_x^2(i,j) + G_y^2(i,j)} \tag{2.1}$$

   And the orientation (θ) of the gradient at each pixel is calculated as:

$$\theta(i,j) = \tan^{-1}(G_y(i,j)/G_x(i,j)) \tag{2.2}$$

   Note that although 3 gradients are calculated for each pixel in color (RGB) images, only the maximum gradient and the associated orientation are used in the proceeding steps. The gradient magnitude $G(i, j)$ and orientation $\theta(i, j)$ are displayed for images with arc, discrete and patchy aurora in the second row of figure 2.11.

2. **Divide the image into cells and blocks** The next step is to divide the image into cells and blocks by defining the regions where the local features will be evaluated. Appropriate sizes can be selected by matching the cell and block sizes to the sizes of the features of interest in the images, here each cell is defined as a 16×16 pixel region and each block is defined as a 4×4 cell region (i.e. a 64×64 pixel region). The cell size is displayed in the middle display of the bottom row in Figure 2.11 as the red boxes and the block size is depicted as the blue box.

3. **Generate the histogram of oriented gradients** At each cell (16×16 pixel), the gradient magnitudes $G(i, j)$ are binned to form a histogram. The magnitude of the gradient determines the accumulated bin value and the associated bin is determined from the orientation $\theta(i, j)$. Here 8 orientation bins over (0–360°) were selected.

4. **Normalize the histograms** The pixel binning into the histogram representation is sensitive to brightness variations. Since auroral forms, for example an auroral arc, can appear with a variety of different intensities, it is advantageous to normalize the brightness sensitivity so that all arcs, independently of the intensity, have a similar feature vector. The 16 (4×4 cells) histograms in each block can be represented as a 128 dimension local feature vector (128 = 16 histograms × 8 bins in each histogram). The block vector is then normalized using the $\boldsymbol{l^2}$ norm. The normalized histogram in each cell is displayed in the third row of Figure 2.11 as an angle (rose) histogram where the length of each petal is scaled according to the bin value. The normalization processing is iterated for all blocks positions over the image with an overlap factor of (2 × 2) cells, i.e. 9 possible block positions.

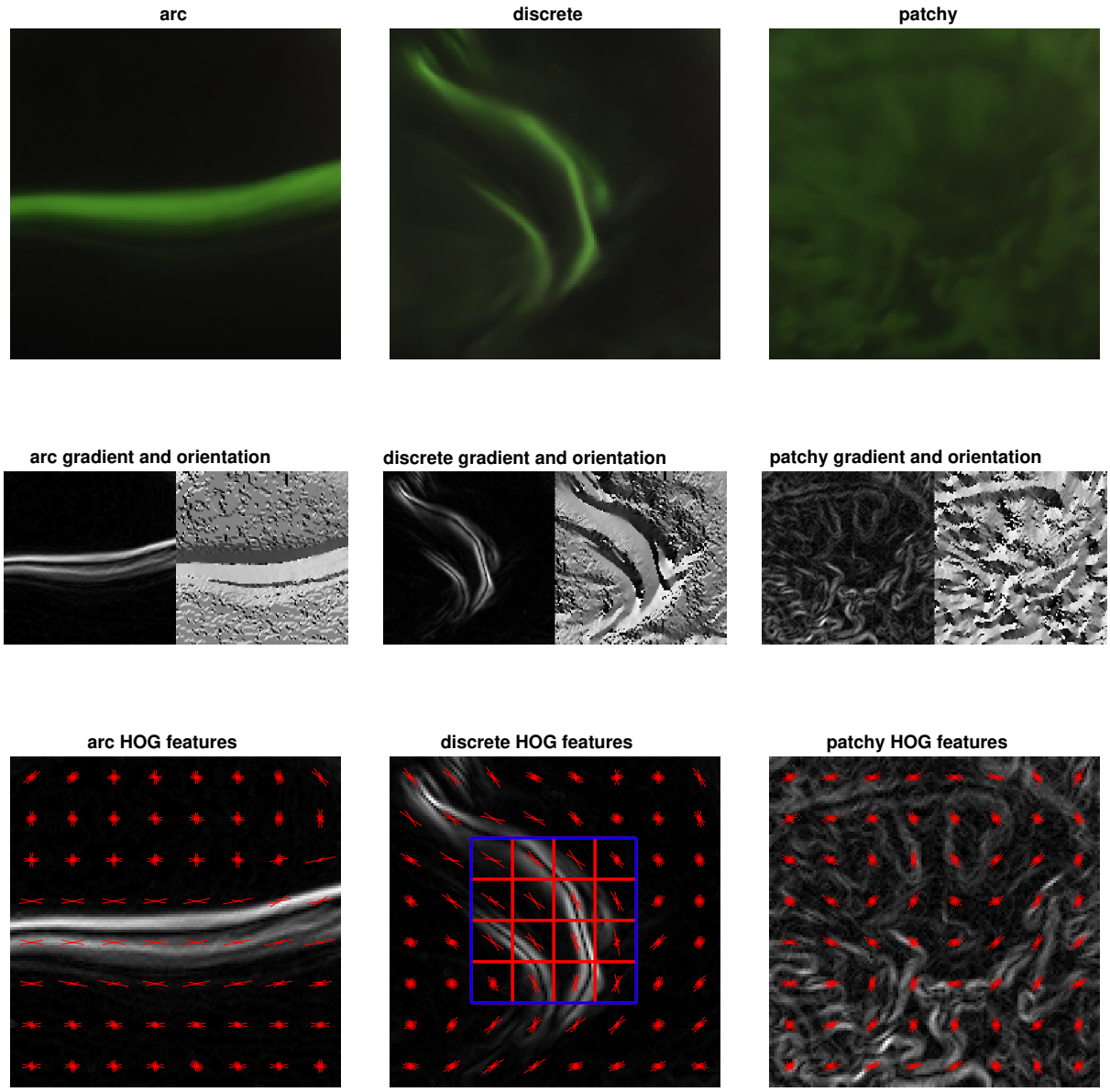


Figure 2.11: The figure illustrates how the auroral features are extracted using the Histogram of Oriented Gradients (HOG) method. Step 1 is to calculate the oriented gradients, displayed in the second row by the magnitude to the left and the orientation to the right. Note that only the maximum gradient magnitude of the 3 color channels at each pixel is used, thus the images become black-and-white after processing step 1. Step 2 is to divide the image into cells and blocks, as indicated by the red and blue boxes in the middle of row 3. Step 3 is to generate the histogram of oriented gradients at each cell. The histograms are thereafter normalized in step 4. The normalized histograms of oriented gradients are displayed in row 3 as angle (rose) histograms where a large horizontal petal symbolize a large gradient along the y-axis and a large vertical petal represent a large gradient along the x-axis.

The final feature vector output ($\boldsymbol{x}$) from the HOG method is a 1152 dimension vector (1152 = 128 dimension local block feature vector x 9 possible block positions). Thus, the HOG method reduced the data size of each image by 98 %, from an image with 49152 data points (128×128×3 pixel) to a feature vector representation of length 1152. Furthermore, the HOG method compactly and robustly extracted the oriented gradients of the image which often provide good discriminative power.

The testing data feature vector space can be studied in order to compare the global features and the local HOG features. The testing data is plotted to the left in Figure 2.12 as a function of the global features; the mean standard deviation ($\bar{\sigma}$) and the mean intensity ($\bar{I}$). In addition, although it is impossible to plot the testing data in the 1152 HOG feature vector space, it is however possible to do a dimension reduction using for example the t-distributed Stochastic Neighbor Embedding (t-SNE) method, see Van der Maaten and Hinton (2008) for details on the t-SNE technique. The t-SNE dimension reduced plot of the testing data in HOG-space is shown to the right of Figure 2.12.

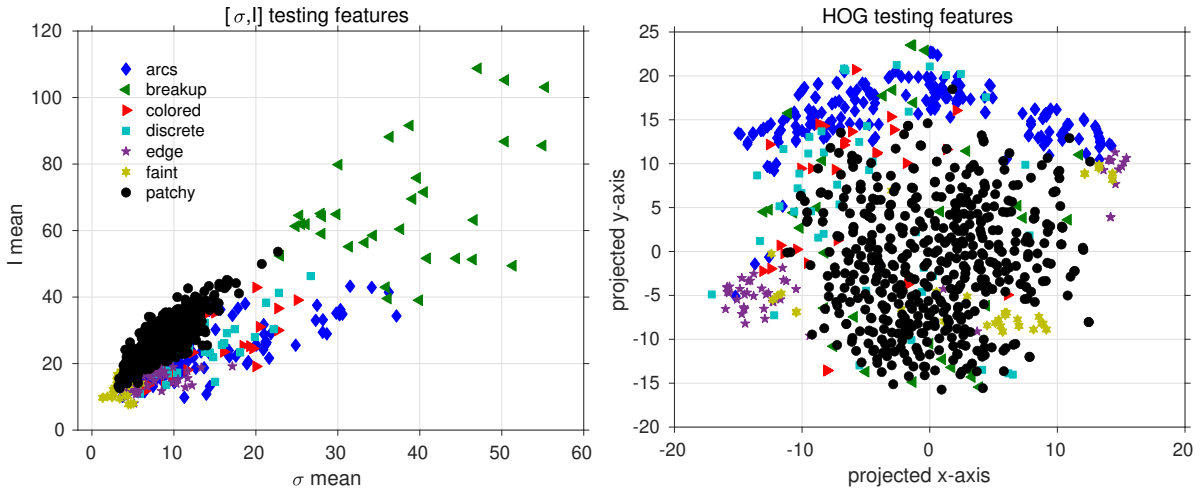


Figure 2.12: To the left, the labeled auroral test images plotted as a function of the mean standard deviation ($\bar{\sigma}$) and the mean intensity ($\bar{I}$). To the right, the test labels in the dimension reduced (t-SNE) HOG feature vector space. In general, the labeled images are better separated in the HOG feature vector space than in the global $[\bar{\sigma}, \bar{I}]$ feature vector space.

By comparing the distribution of labeled images in Figure 2.12, it is clear that the labels are better separated in the HOG feature vector space than in the $[\bar{\sigma}, \bar{I}]$-space, with one exception; *breakup*. The mean standard deviation ($\bar{\sigma}$) and the mean intensity ($\bar{I}$) of images labeled as *breakup* are considerably higher than the mean standard deviation and intensity of the other auroral forms, making the *breakup* $[\bar{\sigma}, \bar{I}]$ parameters a separable feature. However, the $[\bar{\sigma}, \bar{I}]$ features are less successful for separating other auroral labels. On the other hand, the images labeled as *arcs*, *edge*, *faint* and *patchy* appear as more or less separated clusters in the HOG feature vector space. A trained classifier is therefore expected to detect images with these auroral forms. However, the images labeled as *breakup*, *colored* and *discrete* are

not clearly clustered. This suggests that the *breakup*, *colored* and *discrete* auroral forms do not have distinguishable gradient features. Thus, other feature descriptors could be considered to better characterize these auroral forms. Note that an alternative feature extraction approach is considered in the convolutional neural network classification.

2.4 Machine Learning Classification

A direct approach for determining if an image with an unknown auroral form (A_U) is similar to an image with a known (labeled) auroral form (A_L) is to calculate the Euclidean distance:

$$d_E = \sqrt{\sum_{i,j,c} \left(A_U(i,j,c) - A_L(i,j,c)\right)^2} \tag{2.3}$$

where (i, j, c) are the pixel values and color channels.

The Euclidean distance can be used for image classification by calculating the distance to all labeled images in the training data. The unknown image (A_U) is then assigned to the same label as the image that minimized (d_E). However, comparing each test image to all training images in this manner is both computationally slow and not very robust to the natural variation between auroral images. Thus, other classification techniques should be considered. Here three techniques are evaluated and discussed for the auroral image classification task; K-Nearest Neighbor (K-NN), Support Vector Machine (SVM) and Convolutional Neural Network (CNN).

The classifiers are trained and tested on color images acquired by the all-sky camera located near Kiruna, Sweden and operated by the Swedish Institute for Space Physics, the same data set that was used in the study leading to Papers [I] and [II]. The data set consists of 3000 labeled training images and 846 labeled testing images. All images have been pre-processed according to the 6-step pre-processing procedure described in Section 2.1.

K-Nearest Neighbor (K-NN)

The K-Nearest Neighbor algorithm (K-NN) is a simple but popular supervised machine learning classifier (Theodoridis and Koutroumbas, 2009b), that has previously been used for both night-time aurora classification (Syrjäsuo and Donovan, 2004) and dayside aurora classification (Wang et al., 2010). The K-NN classifier uses the Euclidean distance, see Equation 2.3, (or any other appropriate distance measure) to calculate the distance between an unknown sample and all labeled samples in the training data. The approach is similar to the Euclidean distance (d_E) minimization with one exception: Instead of assigning the label of the unknown sample to the same label as the sample that minimized the Euclidean distance (the 1-Nearest Neighbor), the K-NN assigns the label to the majority of labels within the K-Nearest Neighbors of the unknown sample (the K number of samples with the smallest Euclidean distance).

In addition, it is common to use the K-NN classifier on extracted feature vectors for image classification problems rather than the full-resolution data. The feature vector is used in order to reduce computation time and increase robustness towards the natural variation between the auroral images. Figure 2.13 illustrates how an image with an unknown auroral form can be classified using the 5-Nearest Neighbors in the $[\bar{\sigma}, \bar{I}]$-space. Note that the appropriate number of neighbors (K) depends on the classification task and the available training/testing data. In general, increasing the number of nearest neighbors (K) makes the classifier less sensitive to the natural variation between samples (i.e. avoiding over-fitting the data). However, if K is too large, the classifiers smooths the decision boundary too much and starts under-fitting the classification boundary.

The 3-NN and the 5-NN classifiers, using both the global features $[\bar{\sigma}, \bar{I}]$ and the local features extracted by the HOG method, are considered in this section. The performance can be evaluated by studying the confusion matrices for K = 3 in Figure 2.14 and for K = 5 in Figure 2.15.

The confusion matrix summarizes the class-wise performance. The diagonal entries represents the accuracy, the proportion of correctly classified images for each class. Ideally, all diagonal entries should be 1 and all non-diagonal entries should be 0. The row entries mark the classified distribution of all images labeled as 'row'. For example, if all 'row' entries are 0 except the diagonal entry (which is 1), then all images are correctly classified. If the diagonal entry is 0.9 and another entry is 0.1 along the same 'row', then 90 % of the images are labeled correctly and 10 % of the images labeled (manually) as 'row' are classified erroneously according to the nonzero 'column'. By reading the row entries we can determine how the images labeled (manually) as 'row' are classified (automatically) by the algorithm. Moreover, the column entries represent the distribution of the true labels of all images predicted as class 'column'. By reading the column entries we can determine what kind of images (manually labeled as indicated by the associated 'row') end up as being classified (automatically) as 'column'. Overall, the confusion matrix summarizes which labels within the test data are confused (mixed) and which are well separated and identified by the classifier.

In addition, the precision, recall and F1-scores are calculated in order to compactly present and compare the overall performance of the classifiers. Precision is defined as:

$$\text{Precision} = \frac{\text{true positive}}{\text{true positive} + \text{false positive}} \tag{2.4}$$

and measures the classifiers ability for not labeling positive samples as negative. A true positive is when the model correctly predicts the positive class. A false positive is when the model incorrectly predicts the positive class. Recall is defined as:

$$\text{Recall} = \frac{\text{true positive}}{\text{true positive} + \text{false negative}} \tag{2.5}$$

and measures the classifiers ability to find positive samples. A false negative is

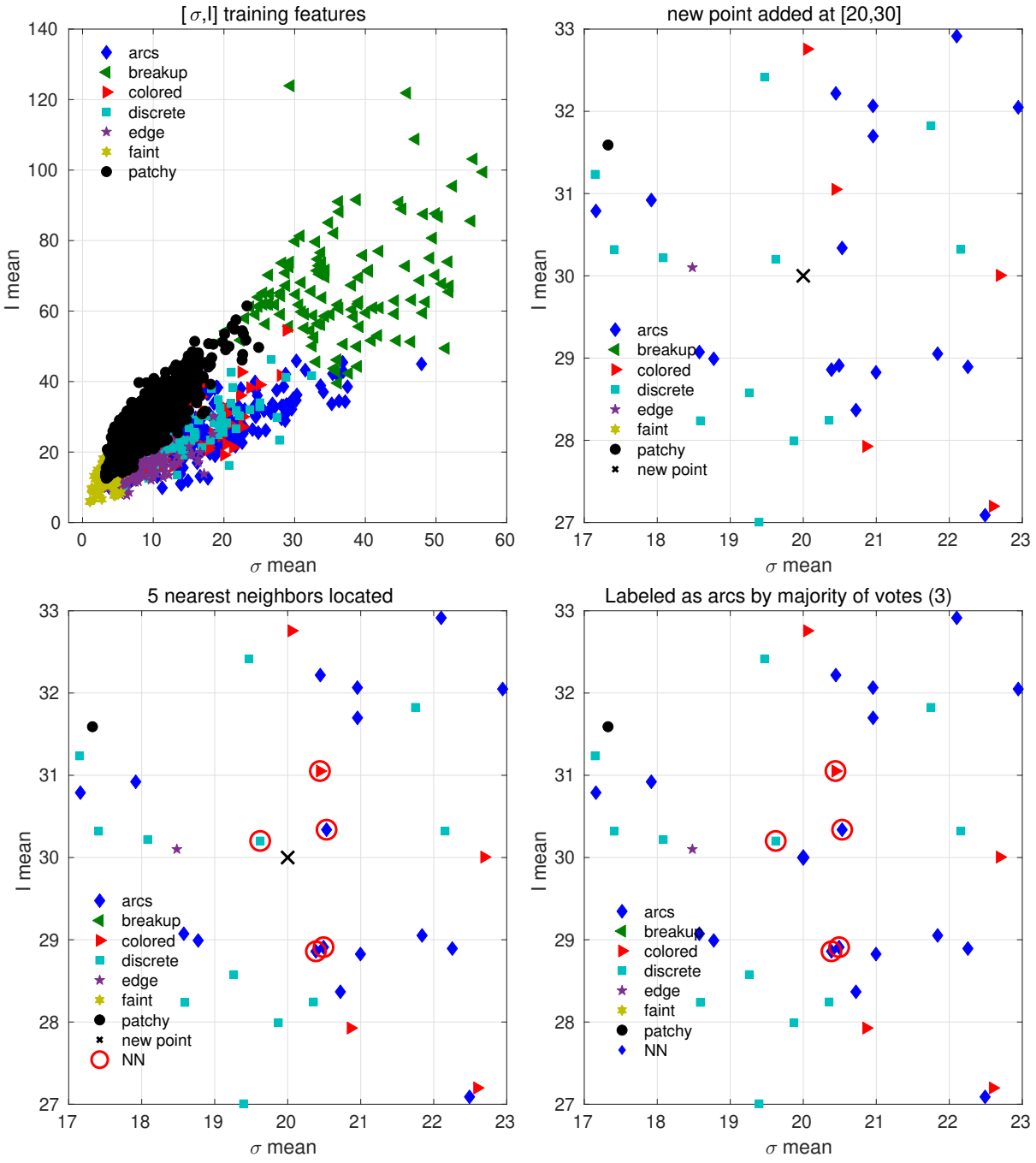


Figure 2.13: An illustration of a 5-Nearest Neighbor classification within a mixed region. In the top left, the training data plotted as a function of $\bar{\sigma}$ and $\bar{I}$. Top right, a zoomed in image of the training data plot with a new sample, representing an auroral image with an unknown auroral form, added at [20, 30]. Bottom left, the 5-Nearest Neighbors are located. Bottom right, a count of the labels is performed (3 *arcs*, 1 *colored* and 1 *discrete*) and the new point is labeled as *arcs* (the majority). Note that if there is a tie between two or more labels, then the nearest neighbor among the tied groups is used.

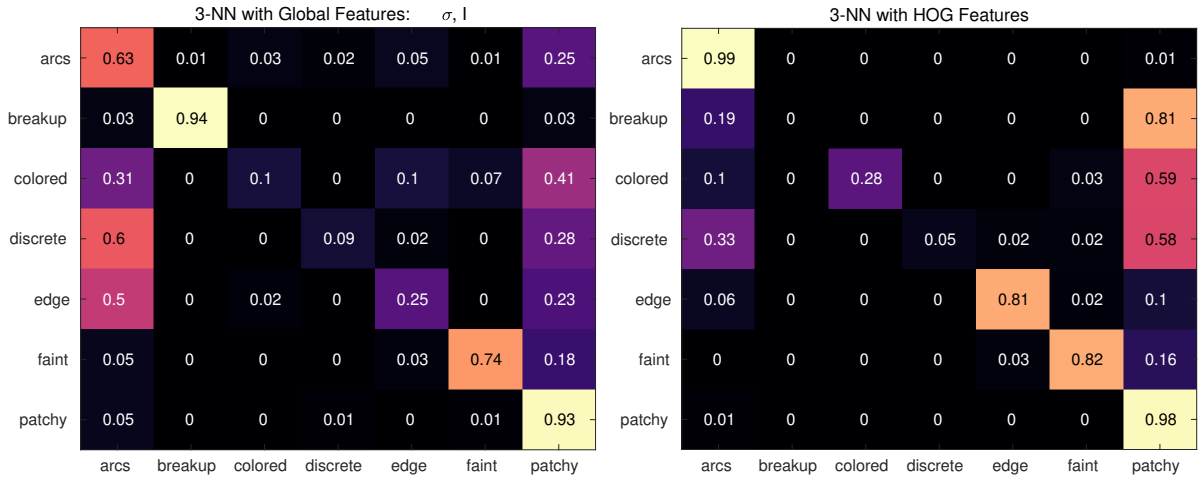


Figure 2.14: To the left, the 3-NN confusion matrix using the $[\bar{\sigma}, \bar{I}]$ features. To the right, the 3-NN confusion matrix using the HOG features. Overall, the 3-NN classifier using HOG features perform better than the 3-NN classifier using the $[\bar{\sigma}, \bar{I}]$ features, with *breakup* being the only clear exception (as expected by studying the feature vector spaces in Figure 2.12).

Figure 2.15: To the left, the 5-NN confusion matrix with the $[\bar{\sigma}, \bar{I}]$ features. To the right, the 5-NN confusion matrix using the HOG features. Again, the HOG feature classification outperforms the $[\bar{\sigma}, \bar{I},]$ feature classification. Overall, the 5-NN classifier is outperformed by the 3-NN classifier.

when the model incorrectly predicts the negative class. F1-score is defined as:

$$\text{F1} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \tag{2.6}$$

and acts as a weighted average of precision and recall.

Table 2.3: The Table summarizes the precision, recall and F1-score for the K-NN classifier with K = 3 (3-NN) and K = 5 (5-NN) using the HOG features. Note that these results are also reported in Paper [I].

Algorithm	Precision	Recall	F1-score
3-NN	0.84 ± 0.0	0.56 ± 0.0	0.58 ± 0.0
5-NN	0.66 ± 0.0	0.54 ± 0.0	0.53 ± 0.0

The results presented in Figures 2.14 and 2.15 and in Table 2.3 can be summarized as follows: It can be concluded, by comparing the confusion matrices, that the K-NN classifier performed better using the HOG feature vectors than the global $[\bar{\sigma}, \bar{I}]$ vectors. Furthermore, the 3-NN classifier generally performs better than the 5-NN classifier. Nevertheless, even when using the HOG features and K = 3, the algorithm was not able to classify reliably *breakup*, *colored* or *discrete* aurora and had a low *edge* aurora accuracy (< 90 %). The overall conclusion with the available data set is therefore; the K-NN algorithm does not provide a satisfactory auroral image classification accuracy (> 90 %), neither using the $[\bar{\sigma}, \bar{I}]$ features nor the HOG feature vectors.

Decision Tree – Linear Support Vector Machine (SVM)

It is interesting, although intuitive, to observe that different feature vectors provide varying class-wise discriminative power in the K-NN classification. For example, *breakup* was well classified using the simple $[\bar{\sigma}, \bar{I}]$ feature vector while *arcs* and *patchy* aurora was better separated in the HOG feature vector space. This observation can be exploited to create a decision tree, where each branch represent a binary logical operator (the image is either classified as -1 or +1) that discerns one auroral class from the rest using a suitable feature vector. An example of such a decision tree is presented in Figure 2.16. The Support Vector Machine (SVM) classifier (Cortes and Vapnik, 1995) is a well-known and robust binary classifier that is suitable for this task.

The SVM classifier has previously been used in (Rao et al., 2014) for classifying color images into three classes; aurora, no aurora and cloudy. In Paper [I], the linear SVM classifier was one of the evaluated classification methods. The SVM classifier in Paper [I] was used in a one-against-all classification scheme for the multi-class auroral image classification problem. Although the SVM classifier is a binary classifier that can not be used directly in the multi-class case, the one-against-all technique handles this problem by considering M two-class problems, where M is the number of classes. For details, see e.g. (Theodoridis and Koutroumbas, 2009c). Note however

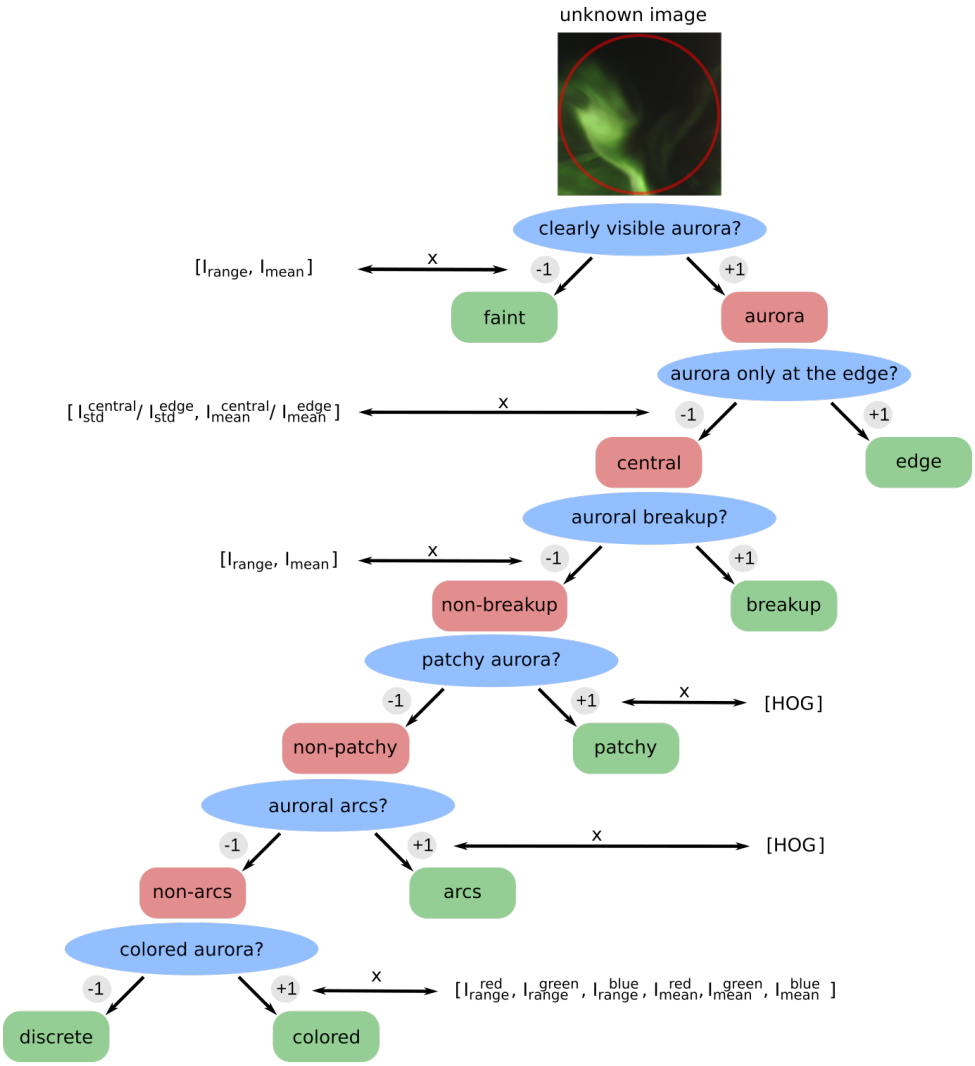


Figure 2.16: The diagram depicts the logical structure of the SVM decision tree used for classification of an unknown input image. Firstly, a trained SVM classifies the image, represented by the feature vector $\boldsymbol{x} = [I_{range}, \bar{I}]$, as faint (–1) or aurora (+1), where (I_{range}) is the maximum-minimum pixel value difference of a heavily (7×7) median filtered version of the input image and ($\bar{I}$) is the mean pixel value. Secondly, if the image is classified as aurora, the image is divided into one central region and four edge regions, the central region is defined by the red circle in the auroral image while the four edge regions are defined by the regions left out by the circle. The feature vector at this branch consists of the standard deviation and mean intensity ratio between the central region and the edge region with the highest mean pixel value. Another linear SVM is then used to classify the features with central auroral emission (–1) and emission predominantly at the edge (+1). Thirdly, if the image is classified as aurora with central emission, the feature vector $\boldsymbol{x} = [I_{range}, \bar{I}]$ is once again used to classify breakup images. At the 4th and 5th decision branch, the HOG feature vectors are used to classify patchy aurora and auroral arcs. Lastly, the intensity range and mean pixel values in the three (RGB) image color channels are used as feature vectors to discern colored aurora and discrete aurora.

that this approach is different to the more transparent and adaptive decision tree approach presented in this section. Both approaches are compared and evaluated towards the end of this section.

At all branches of the decision tree, an auroral image (A_i) is represented by a suitable feature vector ($\boldsymbol{x}_i$) with dimension (p). In addition, the decision tree is designed in such a way that all feature vectors in the labeled training data set ($X = \{\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_N\}$) belong to either class ($\omega_i = \omega_1$) or ($\omega_i = \omega_2$) at each branch. If the feature vectors are linearly separable by some hyperplane with dimension ($p - 1$), defined by the wight vector ($\boldsymbol{w}$) and the threshold (w_0), then the task becomes finding the hyperplane;

$$g(\boldsymbol{x}) = \boldsymbol{w}^T\boldsymbol{x} + w_0 = 0 \tag{2.7}$$

that has the highest discriminative power on the training data. This can be achieved by using the support vectors (the nearest points to the hyperplane) to maximize the margins (z) (Theodoridis and Koutroumbas, 2009c), defined as:

$$z = \frac{|g(\boldsymbol{x})|}{||\boldsymbol{w}||} \tag{2.8}$$

The hyperplane can be defined such that the support vectors belonging to class (ω_1) are equal to 1 and the (ω_2) support vectors are equal to –1. With this convention, the margin becomes:

$$z = \frac{1}{||\boldsymbol{w}||} + \frac{1}{||\boldsymbol{w}||} = \frac{2}{||\boldsymbol{w}||} \tag{2.9}$$

And the decision line is such that:

$$\begin{aligned} \boldsymbol{w}^T\boldsymbol{x} + w_0 \geq 1, \forall \boldsymbol{x} \in \omega_1 \\ \boldsymbol{w}^T\boldsymbol{x} + w_0 \leq -1, \forall \boldsymbol{x} \in \omega_2 \end{aligned} \tag{2.10}$$

Thus, maximizing (z) is equivalent to searching for the parameters ($\boldsymbol{w}, w_0$) that minimize the cost function; $\mathcal{L}(\boldsymbol{w}, w_0)$:

$$\mathcal{L}(\boldsymbol{w}, w_0) = \frac{1}{2}||\boldsymbol{w}||^2 \tag{2.11}$$

subject to;

$$y_i(\boldsymbol{w}^T\boldsymbol{x}_i + w_0) \geq 1, i = 1, 2, ..., N \tag{2.12}$$

where ($y_i = +1$) for (ω_1) samples and ($y_i = -1$) for (ω_2) samples. Minimizing the cost function $\mathcal{L}(\boldsymbol{w}, w_0)$ is a nonlinear optimization problem that is beyond the scope of this thesis to cover, see for example (Theodoridis and Koutroumbas, 2009c) for a description of this procedure. It should however be noted that any local minimum reached by the cost function is also global and unique, thus the optimal hyperplane classifier of a SVM is unique. Figure 2.17 illustrates how the support vectors can be used to define the support vector machine, i.e. the optimal decision hyperplane

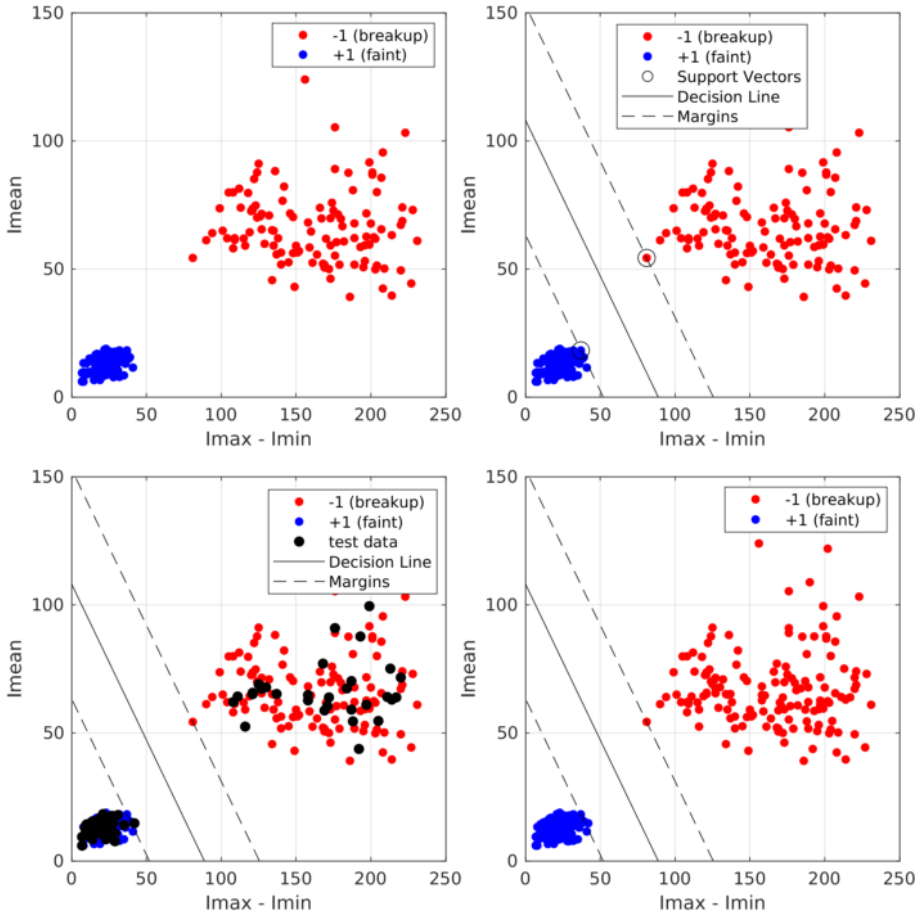


Figure 2.17: The Figure illustrates how labeled images are used to train and test a linear SVM classifier. Top left, the training images labeled as *breakup* and *faint* are plotted. Top right, the line $g(\boldsymbol{x})$ that minimizes Equation 2.11 is plotted along with the associated support vectors and margins. The support vectors lie on either the +1 or the –1 hyperplane margin border; $\boldsymbol{w}^T\boldsymbol{x} + w_0 = \pm 1$. Bottom left, the trained SVM classifier is used to classify images in the testing data. Bottom right, all labels are correctly classified although one *faint* data point fell within the margin.

(in this case a line), that discerns faint and breakup images, represented by the ($[I_{range}, \bar{I}]$) feature vectors.

In general, it can not be assumed that the labeled feature vectors are linearly separable. Nevertheless, with some adjustments, the SVM technique is still applicable. In this thesis, this problem is handled by introducing a slack variable (ξ) into the cost function (Equation 2.11):

$$\mathcal{L}(\boldsymbol{w}, w_0) = \frac{1}{2}||\boldsymbol{w}||^2 + F_\xi \sum_{i=1}^{N} \xi_i \tag{2.13}$$

where ($\xi_i = 0$) for all feature vectors that are correctly classified and outside the margins. The features that are correctly classified but inside the margins have a value ($0 < \xi_i \leq 1$) while the feature vectors that are misclassified have values

($\xi_i > 1$). In words: we insist on the linear separator although the data is not linearly separable by introducing a slack term that penalizes the misclassified feature vectors and the vectors that fall within the margins.

The constant (F_ξ) controls the influence of the slack variables on the decision hyperplane which now is determined by the competing terms in the minimization of the cost function (Equation 2.13). The hyperplane $g(\boldsymbol{x})$ that minimizes Equation 2.13 with ($F_\xi = 1$) for *breakup* classification (the third branch in Figure 2.16) is presented in Figure 2.18.

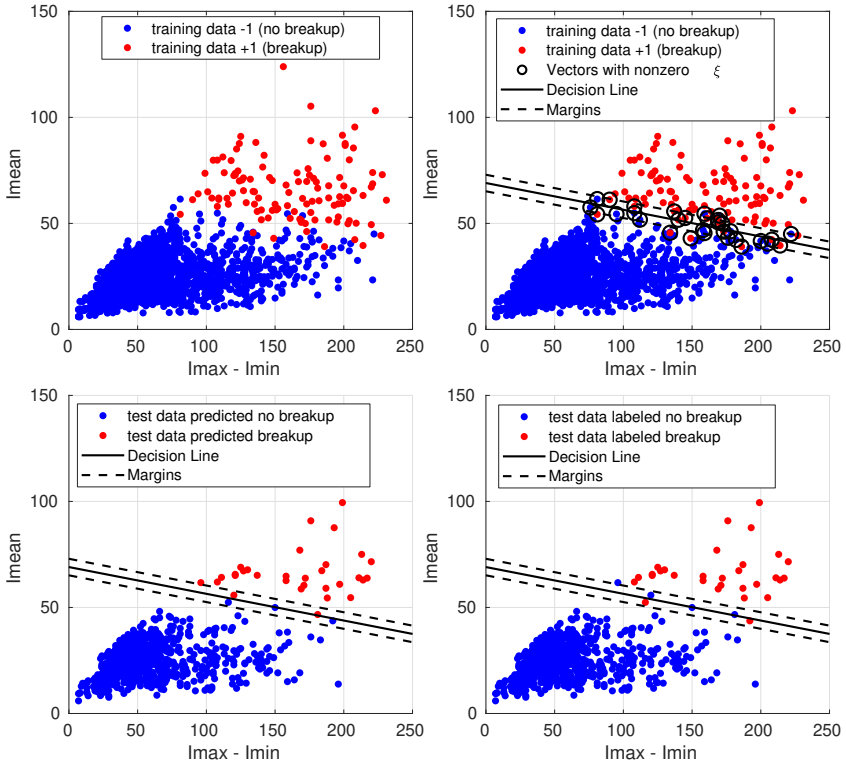


Figure 2.18: The Figure illustrates how a linear SVM can be trained and tested with labeled feature vectors that are not linearly separable. Top left, the feature vectors labeled as *breakup* are plotted in red while the other feature vectors are plotted in blue, it is clear that the classes are mixed and not linearly separable in the $[I_{range}, \bar{I}]$ feature vector space. Top right, minimizing Equation 2.13 yields the decision hyperplane $g(\boldsymbol{x})$ (in this case a line) represented by the solid line. The feature vectors with nonzero (ξ_i) are encircled and the ($\boldsymbol{w}^T\boldsymbol{x} + w_0 = \pm 1$) margins are plotted as dashed lines. Bottom left, the testing data is classified by the decision hyperplane $g(\boldsymbol{x})$. Bottom right, the true labels of the feature vectors are revealed, it is clear that a few samples were misclassified by the linear decision line. However, all the misclassified feature vectors are located near the decision line and are therefore likely to represent somewhat ambiguously labeled images.

The overall performance of the linear SVM decision tree classification on the testing set is presented as a confusion matrix in Figure 2.19. It is clear that the general performance of the SVM classifier is better than the K-NN performance. However, the SVM classifier has a problem with detecting the *colored* and *discrete* auroral

forms. In addition, there is a tendency of error propagation in the decision tree where the classifier struggles to classify the auroral classes at the lower part of the decision tree; even the *arcs* class, that was classified with 96 % percent accuracy by the 3-NN classifier using the HOG feature vectors, only reaches an accuracy of 0.77 % using the same feature vectors but with the SVM decision tree classifier.

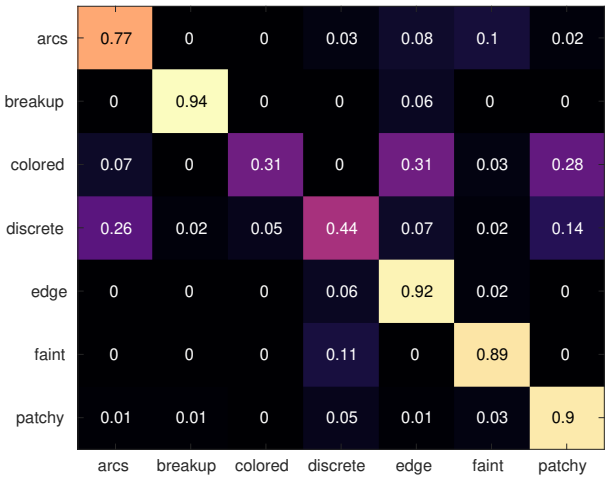


Figure 2.19: The linear SVM decision tree confusion matrix. Overall, the SVM decision tree classifier outperforms the K-NN classifiers. However, the SVM decision tree classifier has a problem detecting *colored* and *discrete* aurora and struggles with error propagation through the decision tree.

The SVM precision, recall and F1 scores are presented in Table 2.4 along with the resulting scores from the 3-NN and 5-NN classifiers with HOG feature vectors. In addition, the linear SVM one-against-all classification precision, recall and F1 scores from Paper [I] are added to Table 2.4 for comparison. The SVM decision tree and the SVM one-against-all performance is comparable with a similar F1-score.

Table 2.4: The Table summarizes the precision, recall and F1-score of the classifiers; 3-NN, 5-NN, SVM decision tree and SVM one-against-all. The 3-NN, 5-NN and SVM one-against-all scores are calculated using the HOG feature vectors while the SVM decision tree feature vectors differs according to the binary classification branch, see Figure 2.16 for details. Note that the 3-NN, 5-NN and SVM one-against-all scores are also reported in Paper [I].

Algorithm	Precision	Recall	F1-score
3-NN	0.84 ± 0.0	0.56 ± 0.0	0.58 ± 0.0
5-NN	0.66 ± 0.0	0.54 ± 0.0	0.53 ± 0.0
SVM decision tree	0.71 ± 0.0	0.74 ± 0.0	0.72 ± 0.0
SVM one-against-all	0.78 ± 0.0	0.70 ± 0.0	0.72 ± 0.0

The main drawback of the SVM one-against-all approach is that only one feature vector representation can be used; here the results using the HOG feature vector is reported. On the other hand, the main drawback of the SVM decision tree approach is the error propagation through the tree. However, the great asset of the

decision tree approach is its transparency and adaptive power. The slack variable constant (F_ξ), here set to 1 at all branches, could be adjusted and more advanced feature vectors could be used to increase the performance of the branches with low classification accuracy. In addition, although only linear SVM classifiers were considered and discussed here, an enhanced performance is expected if non-linear SVM classifiers also were evaluated, see for example Theodoridis and Koutroumbas (2009d) for details.

Neural Networks

Artificial neural networks are inspired by biological neural networks. Yet, even the largest artificial neural networks with ∼1000 million parameters are tiny compared to the human brain with roughly 100 billion neurons and $\sim 10^{15}$ connections between the neurons (Deweerdt, 2019). Nevertheless, similar to the human brain, artificial networks are highly versatile and can be specialized, using training data, to solve numerous tasks, such as speech recognition, computer-aided medical diagnosis and character recognition (Theodoridis and Koutroumbas, 2009a). For some tasks, such as playing chess (Silver et al., 2017) and some image classification tasks (Buetti-Dinh et al., 2019), the deep neural networks even achieve superhuman performances, despite its relatively simple architecture.

Figure 2.20 is a diagram of a fully connected neural network, where each node in a layer (L) is connected to all nodes in layer ($L-1$). A node represents a placeholder

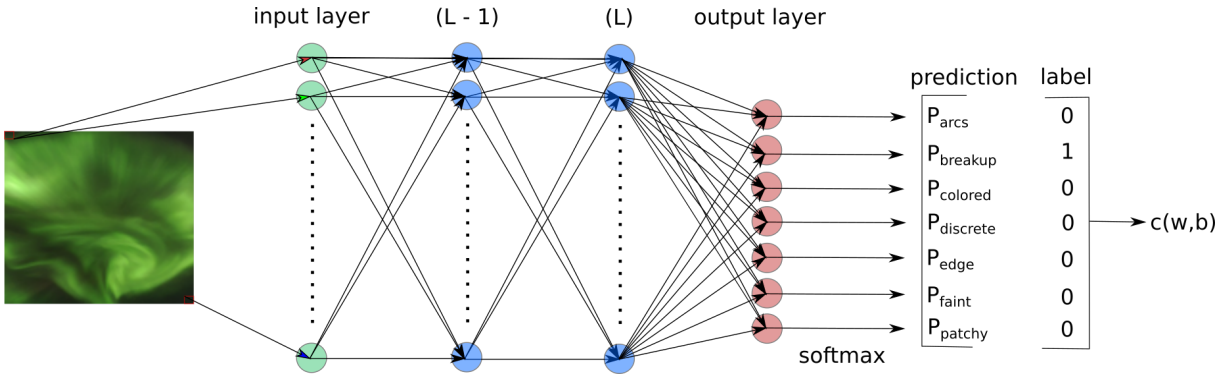


Figure 2.20: A fully connected neural network with two hidden layers (L-1) and (L). The number of neurons and the activation value of each neuron in the input layer is determined by the image pixel dimension (128×128×3) and the pixel values. The arrows between the layers represent the connections between the neurons. The output layer provides the scores of each class. The scores are normalized to probabilities using the softmax function, Equation 2.16. The predicted class probability is thereafter compared to the true label vector of the image and the performance of the network is quantified using the cost function (c). The cost function is minimized by optimizing the weights and biases of the network (the connections), subject to N training images and labels. Note that the network architecture can be specialized for the task at hand by optimizing the network hyper-parameters (the number of hidden layers and the number of neurons in each layer).

for a floating number. In the input layer, these nodes are the vector representation

of the pixel values in the image. The node values in the layers after the input layer are determined by a set of weights (w) and biases (b) that need to be optimized to perform the task at hand. The connections between neuron (j) in layer (L) to all neurons in layer ($L-1$), see Figure 2.20, can be expressed as:

$$z_j^L = \left(\sum_k w_{jk}^L a_k^{L-1} \right) + b_j^L \tag{2.14}$$

where (a_k^{L-1}) is the value (the activation) of neuron (k) in layer ($L-1$). The activation of neuron (a_j^L) is then determined from the connections (z_j^L) by a non-linear activation function $\sigma(z_j^L)$, allowing the networks to model non-linear relationships. A commonly used activation function is the rectified linear unit function (ReLU):

$$\begin{aligned} \mathsf{ReLU}(x) &= \mathsf{max}(0, x) \\ a_j^L = \sigma(z_j^L) &= \mathsf{ReLU}(z_j^L) \end{aligned} \tag{2.15}$$

Note that the number of neurons in the output layer in Figure 2.20 has the same dimension as the number of target classes (7 in the auroral image classification example), and can therefore be used to compare the network output to the true label input. The output values of these neurons, the class scores (s), are normalized to probabilities that sum to 1 using the softmax function (Theodoridis and Koutroumbas, 2009d). For class (o), out of (O) classes, with the associated output score (s_o), the softmax probability is given by:

$$P_o = \frac{e^{s_o}}{\sum_{o=1}^{O} e^{s_o}} \tag{2.16}$$

The predicted label probability vector of image number (n) with the associated true label vector (label^n) are used to evaluate the network performance. Ideally, the target class predicted probability should be 1 while all other predicted class probabilities should be 0. Thus, the goal is to find the weights (w) and biases (b) that minimize the difference between the predicted class probability vector and the true label vector, see Figure 2.20, subject to (N) training images and labels.

For each image matrix (A^n), the performance is quantified by the cost function (c), often evaluated by the cross-entropy loss function. While the optimization is achieved by minimizing the summed cost function ($\mathcal{L}$):

$$\mathcal{L} = \frac{1}{N} \sum_{n=1}^{N} c(w, b; A^n, \mathsf{label}^n) \tag{2.17}$$

Note that (c) is a function of all the weights and biases that define the network. At first, these weights and biases are randomly initialized. The weights and biases are thereafter optimized by minimizing ($\mathcal{L}$). At this point in time, it is common to use the Adam stochastic gradient decent optimizer and the backpropagation algorithm to update the weights and biases to minimize ($\mathcal{L}$) (Kingma and Ba, 2014; Theodoridis and Koutroumbas, 2009d). It is however beyond the scope of this thesis to discuss the optimization task.

The network architecture in Figure 2.20 could be implemented directly and used to classify the auroral images. However, fully connected neural networks are sub-optimal for the auroral image classification tasks due to the large input data size (128×128×3) and therefore the excessive amount of connections needed to perform the classification task adequately. Consequently, a fully connected neural network would be time consuming to train and prone to overfitting. Overfitting in this context is when the network is successfully optimized using the training data, but does not perform well on data beyond the domain of the training data.

A Convolutional Neural Network (CNN) is a more suitable architecture that is commonly used for image classification tasks. CNNs are regularized versions of the fully connected neural network with fewer parameters relative to the amount of hidden layers (the network depth).

A CNN typically consists of tree types of layers; convolutional, pooling and fully connected layers.

- **Convolutional layers:** The inclusion of the convolutional layers to the neural network is motivated by the ability of the convolution operation to extract image features with high class-wise discriminative power. The convolution filters are not pre-defined, as in the HOG gradient feature extraction technique, but are optimized to solve the classification task at hand. In the convolutional layers, the image volumes play a similar role as the neurons in the fully connected network and the convolution filters act as the connection weights. Therefore, in order to maintain a notation analogous to Equation 2.15, we express the activation of a convolved image (j) of layer L as (a_j^L) and the associated convolution filter as (w_j^L). With this notation, the connection between (a_j^L) and (a_k^{L-1}) can be expressed as:

$$\begin{aligned} a_j^L &= \sigma\left(w_j^L * a_k^{L-1} + b_j^L\right) \\ a_j^L &= \sigma(z_j^L) \\ a_j^L &= \mathsf{ReLU}(z_j^L) \end{aligned} \tag{2.18}$$

  where ($*$) denote the convolution operation and (k) is the data depth index in layer ($L-1$). The first convolutional layer, from the input image (a_1^0) to the convolved images (a_K^1), is illustrated as K convolution operations with K different filters in Figure 2.21.

- **Pooling layers:** The pooling layers are used to reduce the size of the data being propagated through the network. This is motivated by a desire to speed up computation and increase the robustness of the CNN. A pooling layer have no unknown parameters (no optimization needed) but is defined by a predefined pooling filter size (f_p) and a stride (s_p). Pooling filters with size ($f_p = 2$) and stride ($s_p = 2$) are often used. The data reduction is achieved by extracting a representative value within a sliding pooling window ($f_p \times f_p$). The average value (average pooling) or the maximum value (max pooling) are commonly used as the representative value within the pooling window. The output image matrix is populated by sliding the pooling window across the entire input

image with a step size of (s_p) pixels. The pooling is performed over all image channels. Thus, the pooling reduces the input image volume of size ($n \times n \times n_c$) to an output image volume of size ($n_p \times n_p \times n_c$), where ($n_p$) is determined (if no padding is included) by:

$$n_p = \left\lfloor \left(\frac{n - f_p}{s_p} + 1 \right) \right\rfloor \tag{2.19}$$

For example, if a pooling window of size ($f_p = 2$) and a stride of ($s_p = 2$) is used (pooling without overlap), then an input image of size ($128 \times 128 \times 3$) is reduced by a factor of 4 to an output image of size ($64 \times 64 \times 3$).

- **Fully connected layers:** The convolutional and pooling layers are used to extract the auroral features and reduce the feature data size. However, these features needs to be combined and evaluated in order to perform a prediction of the image label. This task is accomplished by the fully connected layers. The fully connected layers are designed similarly to the fully connected layers in the fully connected neural network (Figure 2.20), but the input is now the convolved feature maps and not the full resolution auroral image. The output of the fully connected layers is the label score vector that can be transformed to predicted label probabilities using the softmax function, as described in Equation 2.16.

Different combinations of the convolution, pooling and fully connected layers result in different CNN architectures. AlexNet is often considered as the breakthrough of deep CNNs after winning the 2012 ImageNet image classification contest with 1.2 million training images divided into 1000 target labels (Krizhevsky et al., 2012). Figure 2.22 is a diagram of the AlexNet architecture with 7 output scores (corresponding to the 7 auroral target labels in Figure 2.8).

In the context of deep neural networks, the term deep refers to the depth of the classifier, i.e. consisting of multiple hidden layers, often with millions of parameters (wights and biases) that needs to be optimized. As with the fully connected network in Figure 2.20, the optimization is achieved by minimizing the cost function ($\mathcal{L}$), subject to (N) training images and labels. Note however that the CNN optimization is now a combination of the weights and biases that define the convolutional activations (Equation 2.18) and the weights and biases that define the fully connected activations (Equation 2.15).

After the breakthrough of AlexNet, several deep CNN architectures have been designed and employed for various image classification task such as; character recognition (Wu and Chen, 2015), brain tumor classification (Amin et al., 2018) and facial recognition (Balaban, 2015). However, CNNs have not been widely used for auroral image classification. Classification of dayside aurora using CNNs have been studied in Han et al. (2017); Niu et al. (2018); Yang and Zhou (2020). While CNN classification of nighttime aurora has only been studied (and published) in Clausen and Nickisch (2018), prior to Paper [I].

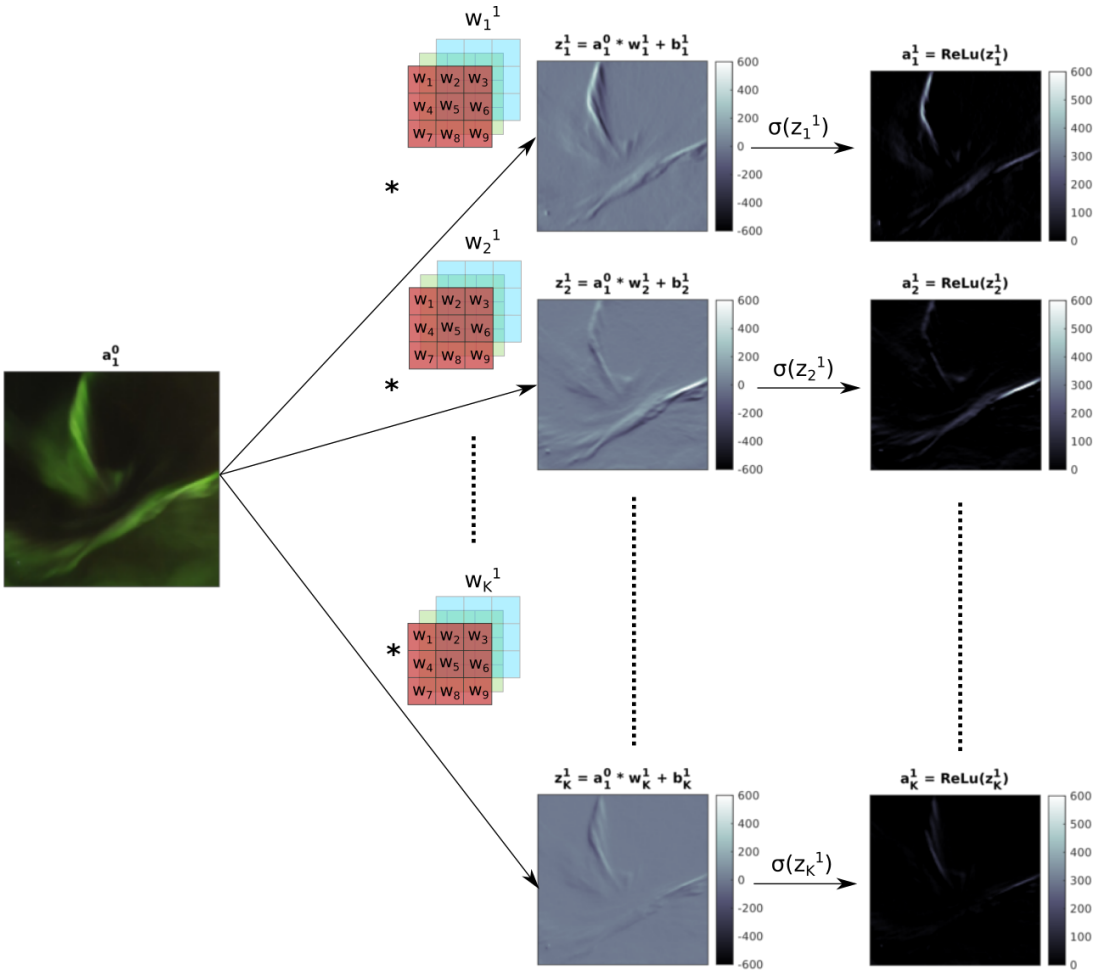


Figure 2.21: An illustration of the convolution procedure from one layer (in this example the input layer) to the next layer. The input image (a_1^0) is convolved with (K) (RGB) filters. The size (f) and number of filters (K) are hyper-parameters that are specified (pre-defined) for the given CNN architecture. The convolutional layer extracts auroral features with high discriminative power by optimizing the weights in each filter (in this example (3×3×3) = 27 weights) and the associated bias. For illustration purposes, the biases are all set to zero ($b_1 = b_2 = b_K = 0$) and a gradient (Prewitt) filter is used along the x-axis at the top (k = 1), along the the y-axis in the second row (k = 2) and along the diagonal at the bottom (k = K). The ReLU function is used as the activation function (σ).

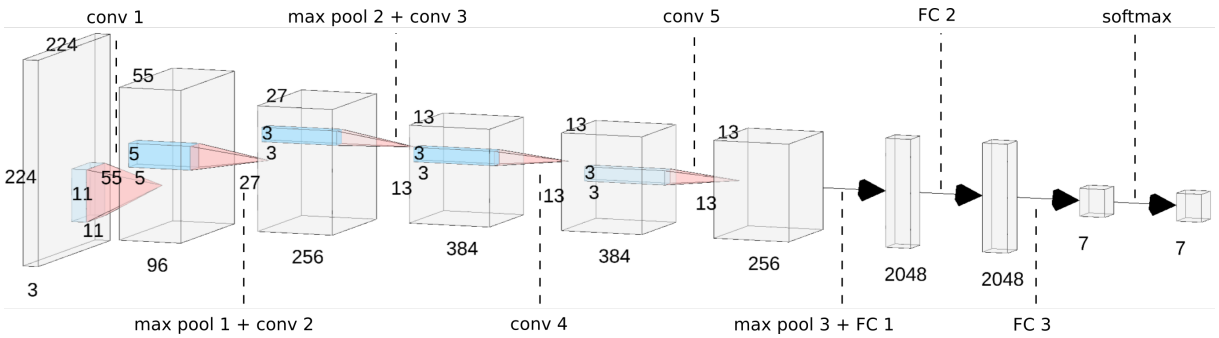


Figure 2.22: A diagram of the AlexNet architecture with 7 target classes. Alexnet consist of 5 convolutional (conv) layers, 3 max pooling (max pool) layers and 3 fully connected (FC) layers. The convolution filter size and the data dimension are indicated in each data volume. The input image needs to be resized to ($224 \times 224 \times 3$) pixels before being feed to the network. The convolutional networks (usually) propagate the features through the network by a stepwise reduction in the image (n) size and an increased data volume depth (n_c). At the end of the convolutional/pooling layers, the data volume is flattened into a 1–D vector. The combined 1–D feature vector is the input to the fully connected layers. The number of nodes in the fully connected layers are indicated by the number below the FC columns. The final output of AlexNet is the target class scores from the fully connected layers. The class scores are normalized to predicted class probabilities using the softmax function.

In Clausen and Nickisch (2018), a pre-trained CNN was used to extract features from black-and-white THEMIS images with only the last layer of the CNN being optimized using auroral image training data. The output features were classified into the 6 classes; *clear/no aurora*, *cloudy*, *moon*, *arc*, *diffuse*, and *discrete*. An average class-wise accuracy of 82 % was achieved, but with an enhanced confusion between the auroral subclasses (*arc*, *discrete* and *diffuse*).

Paper [I] was motivated by the lack of studies where nighttime auroral images were classified using CNNs. We were further inclined to conduct the study given the high performance of CNNs for other image classification tasks in general and the promising results from Clausen and Nickisch (2018) in particular. The objective of Paper [I] was to study the auroral image classification performance of different CNN architectures that were trained from scratch, in contrast to the pre-trained CNN approach in Clausen and Nickisch (2018). Furthermore, Clausen and Nickisch (2018) experienced enhanced confusion between auroral subclasses (*arc*, *discrete* and *diffuse*). Paper [I] therefore focused on classifying auroral subclasses into even more subclasses while aiming at achieving an error rate ($<$10 %), which is considered sufficient for operational purposes (Syrjasuo and Partamies, 2011).

The error rate ($<$10 %) goal was achieved by the Residual Learning Network (ResNet) architecture (He et al., 2016), but not by the AlexNet architecture, as presented in Paper [I]. The main differences between the ResNet and the AlexNet architecture is the network depth and the addition of skip connections between the convolutional layers in the ResNet architecture. Ideally, we want to design a very deep CNN

that can extract complex image features. However, the weights and biases of very deep CNNs are difficult to optimize using gradient decent and backpropagation, ultimately leading to a degradation of the CNN performance with an increasing the number of layers.

The ResNet CNN architecture solves this problem by adding skip connections between the convolutional layers. It is beyond the scope of this thesis to study the design of deep CNNs architectures, however a brief description of the basic concept of the skip connections used in ResNet-50 is included. The skip connections can be mathematically expressed in terms of the connections between the activations of the convolutional layers in Equation 2.18. Starting with the connection between the activations in layer ($L-1$) to layer (L) over the entire data volume:

$$\begin{aligned} a^L &= \sigma\left(w^L * a^{L-1} + b^L\right) \\ a^L &= \sigma(z^L) \end{aligned} \tag{2.20}$$

Similarly, the connection between the activations in layer (L) to layer ($L+1$) can be expressed as

$$\begin{aligned} a^{L+1} &= \sigma\left(w^{L+1} * a^L + b^{L+1}\right) \\ a^{L+1} &= \sigma(z^{L+1}) \end{aligned} \tag{2.21}$$

Now, the problem with the optimization of very deep CNNs is that information between layers, for example from layer (a^{L-1}) to layer (a^{L+1}), needs to propagate through layer (a^L). In He et al. (2016), it was demonstrated that the diminishing of the information throughput in CNNs can be largely avoided by adding the residual activation to the activation function. Thus, in the ResNet architecture, the activation function of layer ($L+1$) can be expressed as:

$$\begin{aligned} a^{L+1} &= \sigma\left(\left(w^{L+1} * a^L + b^{L+1}\right) + a^{L-1}\right) \\ a^{L+1} &= \sigma(z^{L+1} + a^{L-1}) \end{aligned} \tag{2.22}$$

The addition of (a^{L-1}) to the (a^{L+1}) activation function acts as an identity mapping from (a^{L-1}) to (a^{L+1}), which can be adjusted during the optimization of the weights (w^{L+1}) and biases (b^{L+1}). Explicitly including the identity mapping eases the information throughput in the ResNet architecture, allowing training of deeper CNNs without fast performance degradation.

In Paper [I], two ResNet architectures were evaluated; ResNet-18 with 18 convolutional layers and ResNet-50 with 50 convolutional layers (i.e. much deeper networks than the 5 convolutional layers in AlexNet). ResNet-50 achieved the highest performance with an average class-wise accuracy of 91 %. The resulting ResNet-50 performance on the testing data is presented as a confusion matrix in Figure 2.23. In addition, the precision, recall and F1 scores are summarized in Table 2.5 for the classifiers; AlexNet, Resnet-18 and ResNet-50.

Overall, the convolutional neural networks outperforms the K-NN and the SVM classifiers. The enhanced performance of the CNNs is due to the superior feature extraction capability that is optimized for the auroral image classification task. For

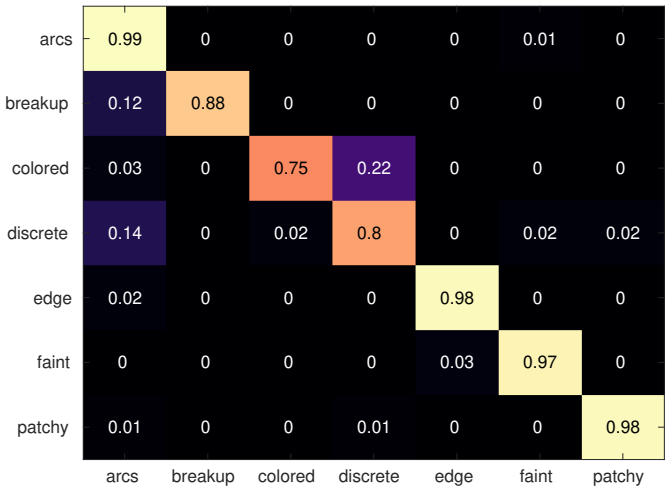


Figure 2.23: The ResNet-50 classification confusion matrix. The ResNet-50 classifier clearly outperforms both the K-NN and the SVM classifers, see Figures 2.14, 2.15 and 2.19 for comparison. The ResNet-50 classifier is able to detect and classify most images correctly, but with some problems discerning *colored* and *discrete* images. In addition, $\sim$12 % of the testing images labeled as *breakup* are classified as *arcs*.

the K-NN and SVM classifiers, the features were extracted by fixed and manually defined techniques. However, defining feature vectors with high discriminative power is a challenging task, as seen in the $[\bar{\sigma}, \bar{I}]$ and HOG feature vector spaces in Figure 2.12. For comparison, the dimension reduced (t-SNE) feature vector space output of the ResNet-50 architecture is presented in Figure 2.24. It is clear that the auroral images are much better separated in the final ResNet-50 feature vector space, than in the $[\bar{\sigma}, \bar{I}]$ and HOG feature vector spaces.

To conclude; automatic auroral image classification is a suitable task for convolutional neural networks. In particular, the ResNet-50 architecture achieved an error rate ($<$10 %), which is considered to be sufficient for operational purposes (Syrjasuo and Partamies, 2011).

Table 2.5: The Table summarizes the precision, recall and F1-score of the CNN classifiers. In addition, the results from the K-NN and SVM classifiers are added for comparison. The uncertainty in the CNN scores are calculated from the average performance of 10 runs with random parameter initialization of the weights and biases.

Algorithm	Precision	Recall	F1-score
3-NN	0.84 ± 0.0	0.56 ± 0.0	0.58 ± 0.0
5-NN	0.66 ± 0.0	0.54 ± 0.0	0.53 ± 0.0
SVM decision tree	0.71 ± 0.0	0.74 ± 0.0	0.72 ± 0.0
SVM one-against-all	0.78 ± 0.0	0.70 ± 0.0	0.72 ± 0.0
AlexNet	0.88 ± 0.03	0.88 ± 0.03	0.87 ± 0.03
ResNet-18	0.92 ± 0.03	0.87 ± 0.05	0.89 ± 0.04
ResNet-50	0.92 ± 0.03	0.89 ± 0.04	0.90 ± 0.03

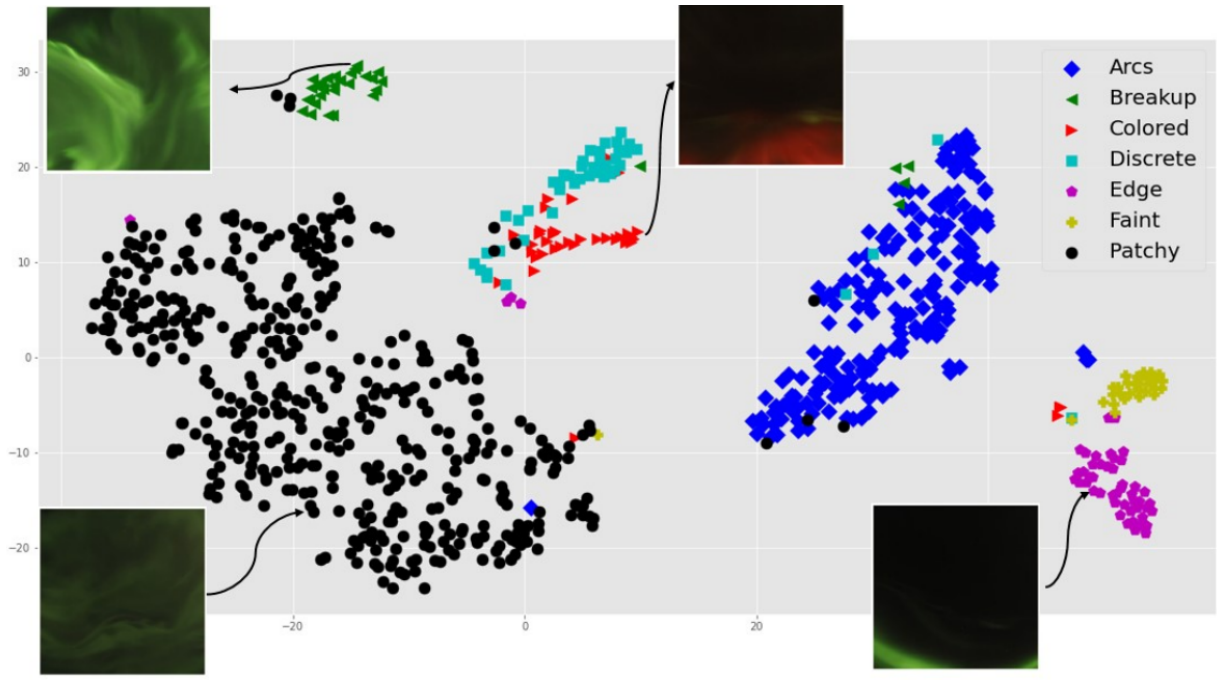


Figure 2.24: The auroral test images in the dimension reduced (t-SNE) ResNet-50 feature vector space. It is clear that most classes are well separated and are expected to obtain a high classification accuracy. However, the *colored* and *discrete* images are somewhat mixed, as indicated by the confusion matrix in Figure 2.23. Nevertheless, the ResNet-50 architecture achieved an average class-wise accuracy of 91 %, which is comparable to the 95 % label agreement between the two experts. The Figure is reprinted from Kvammen et al. (2020), under the Creative Commons Attribution 4.0 License.

Chapter 3

Multi-Viewpoint Analysis of Auroral Images

When once your point of view is changed, the very thing which was so damning becomes a clue to the truth.

– Sir Arthur Conan Doyle, *The Complete Sherlock Holmes*

The aurora has a three-dimensional (3–D) intensity distribution, yet isolated cameras represent the emission as a 2–D projection, diminishing the volumetric information. The spatial information is however needed to calculate auroral quantities such as: excitation rates, electron precipitation energies, optical flow velocities and ionospheric electric fields (Whiter et al., 2013).

Consequently, assumptions of the 3–D distribution are typically made when numerical input is required. For example, an auroral height of 110 km is commonly assumed, often without considering the uncertainty of this estimate. Still, even for a thin–stable auroral arc, the peak height can vary up to 40 km along its length (Sangalli et al., 2011a), making height assumptions a significant uncertainty factor in auroral studies.

Dubious spatial assumptions can be avoided if the aurora is observed simultaneously from separated locations by using multi-viewpoint analysis techniques. Three such methods are described in this chapter; triangulation in Section 3.2, shell projection in Section 3.3 and tomography-like reconstruction in Section 3.4, where the overall objective is to retrieve 3–D auroral information from the 2–D images.

Both images of the naturally occurring Aurora Borealis, discussed in Section 1.1, and images of artificially induced aurora, discussed in Section 1.2, are used to illustrate the techniques. In particular, all-sky images of the Lumikot aurora are used to exemplify the stereoscopic triangulation and shell projection methods, supporting the results and analysis in Paper [III]. While wide field-of-view (54°–90°) images of artificial aurora are used to exemplify the multi-viewpoint triangulation and the tomography-like reconstruction, describing the technique that was applied in Paper [IV]. However, the presented methods can, in principle, be used for any distinct auroral feature.

3.1 3–D Projection and Image Pre-Processing

In order to obtain the volumetric information from the image projections, without relying on a priori assumptions, it is required that the imaging stations observe the object of interest with overlapping fields-of-view and that the imaging stations are well-separated (with baselines $>\sim$10 km). In addition, since the aurora is a highly dynamic phenomena, it is also critical that the images are taken near-synchronously (separated in time by $<\sim$0.1 seconds). The appropriate baseline separation and the necessary timing precision is however dependent on the studied auroral phenomena and the camera system.

Moreover, information about the camera's position, orientation and the pointing direction of each pixel (i.e. the field-of-view) is needed in order to map features within an image into the 3–D scene, this is illustrated in Figure 3.1.

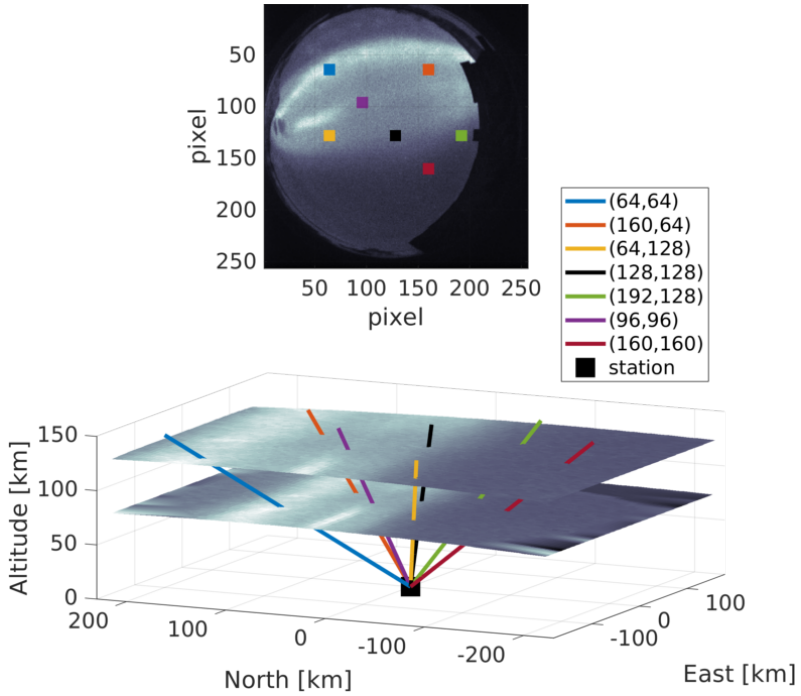


Figure 3.1: At the top, an auroral image acquired by the MIRACLE all-sky camera located in Kilpisjärvi, Finland, with a few marked pixels (enlarged in the figure for clarity). At the bottom, the corresponding pixel line-of-sights in a 3–D Euclidean coordinate system with the camera station at the origin. Each pixel value is assumed to be the line-of-sight integrated column emission rate (i.e. image formation without error). Thus, it is impossible from a single image (without a priori assumptions) to determine if the peak emission height should be located at for example 90 km or 130 km, as illustrated by the projected images. Note also that there is a non-linear relationship between pixel distances in the all-sky images and the corresponding distances in the projected planes, making it difficult to infer the horizontal size of auroral features without 3–D projection.

The position of a scientific camera is normally provided by the host observatory while the field-of-view can be estimated by geometric calibration using stellar po-

sitions. In Papers [III] and [IV], the orientation and field-of-view was estimated by identifying stars in the image and matching them with the corresponding stars in the Yale Bright Star Catalog (Hoffleit and Jaschek, 1991). The sky position of the cataloged stars has been accurately measured, providing precise line-of-sight estimates for the pixels containing stars, taking the atmospheric refraction at the time of observation into account. The field-of-view of the entire image can thereafter be calculated by fitting an appropriate camera model to the pointing direction of the star pixels, for example by using the Auroral Image Data Analysis (AIDA) toolbox (Gustavsson, 2015). The AIDA star calibration tool typically provides line-of-sight estimates with 1σ accuracies of ± 0.2 pixel, which is sufficient for most auroral research purposes (Gustavsson et al., 2008a).

In the following sections, it is assumed that the cameras are perfectly calibrated and that the images are formed without error. Thus, optical aberration, atmospheric absorption, detector imperfections and other effects are not discussed in this thesis. For a through description of the image formation and an error analysis of wide field-of-view camera systems, see for example Gustavsson (2000). Note however that optical- and mechanical vignetting is corrected by flat-field-correction and that noise, to some extent, is reduced by 2–D median filtering.

In addition, the auroral images are background subtracted in order to enhance and isolate the auroral feature of interest. An example of the background subtraction procedure for simultaneous stereo observations of Lumikko[1] aurora, acquired by the MIRACLE imaging stations in Kilpisjärvi and Abisko, is presented in Figure 3.2. A similar procedure was used for isolating the artificially induced auroral patch. However, a thorough description of the pre-processing procedures is omitted since the appropriate method depends on background conditions, the camera system and the feature of interest, thus making it difficult to specify a general pre-processing procedure for multi-viewpoint analysis.

3.2 Triangulation

Triangulation of auroral displays has been used to estimate the height of the emission since the first scientific stereo-imaging observations near Alta, Norway (Størmer, 1916; Vegard and Krogness, 1920). Størmer (1916) established that the height of the peak auroral brightness is typically located at 100–110 km by analyzing thousands of image pairs. Still, triangulation of auroral features continues to provide useful results as the camera technology and image processing techniques advances (Kozelov et al., 2019; Sangalli et al., 2011a).

The main advantage of triangulation is that the method can be used to estimate the 3–D coordinate of any point that is recognizable in simultaneous multi-station images, without a priory assumptions. The disadvantage however is that triangulation do not provide distribution estimates of the volume emission rates.

The triangulation problem for the stereo-imaging case, with ($n = 2$) imaging sta-

[1] Singular: Lumikko. Plural: Lumikot, from the Finnish word for weasel

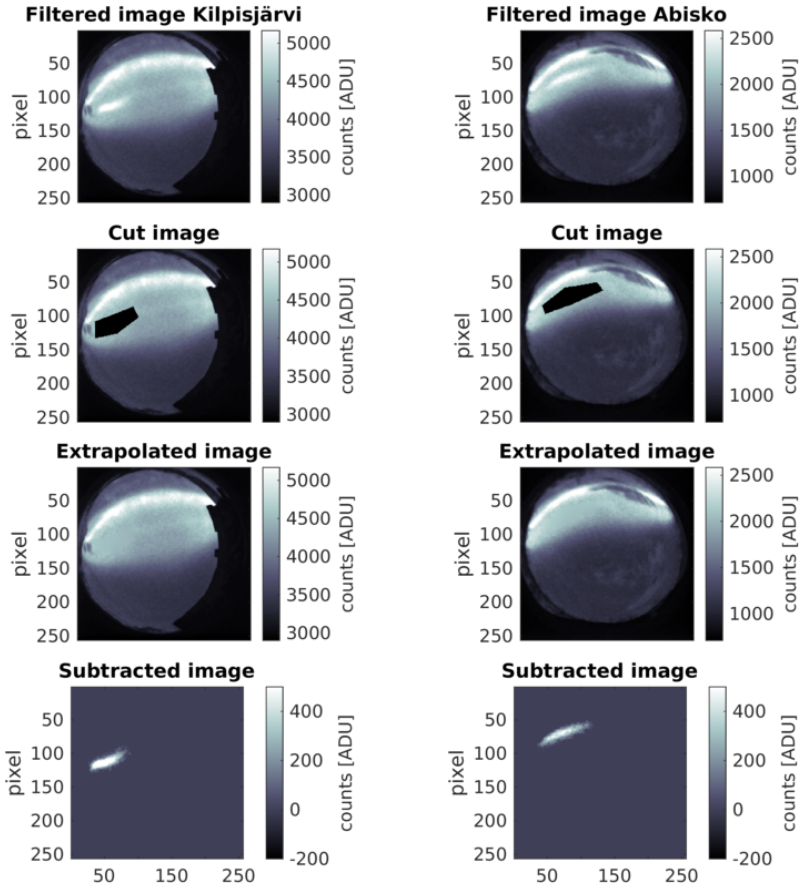


Figure 3.2: The Lumikko, discussed in detail in Paper [III], is recognized in the images as a patch of aurora, located equatorward of the auroral arc, that is rapidly traversing parallel to the arc. In order to isolate the feature of interest (the Lumikko), a background subtraction procedure was performed: 1. The image is filtered with a 2–D median filter. 2. A region enclosing the feature of interest is manually defined and cut out. 3. The cut image is in-painted, where the interpolated pixel values are estimated by solving a system of partial differential equations (D'Errico, 2012). 4. The filtered image is subtracted by the extrapolated image, isolating the feature of interest.

tions, and for the multi-station case, with ($n > 2$), is formulated differently in this section and therefore treated separately.

Stereoscopic Triangulation

The first step of stereoscopic triangulation is to define the lines-of-sight ($\boldsymbol{L}_1$) and ($\boldsymbol{L}_2$), from the camera positions ($\boldsymbol{r}_1$) and ($\boldsymbol{r}_2$), towards an identified common point ($\hat{\boldsymbol{P}}$) that represent the object of interest in the 3–D scene.

Various techniques can be used for determining the lines-of-sight, for example; the midpoint of a fitted model, the location of feature edges or matching points along epipolar lines. However, for the Lumikko triangulation case, the position of the brightest pixel of a heavily 2–D median-filtered image version is used. This point

is illustrated in Figure 3.3 by the yellow square in the Kilpisjärvi image (camera 1) and by the red square in the Abisko image (camera 2). The pointing direction of these pixels (the star-calibrated lines-of-sight) defines the skew lines ($\boldsymbol{L}_1$) and ($\boldsymbol{L}_2$), and can be expressed in a 3–D Euclidean space as:

$$\begin{aligned} \boldsymbol{L}_1 &= \boldsymbol{r}_1 + \lambda_1 \hat{\boldsymbol{e}}_1 \\ \boldsymbol{L}_2 &= \boldsymbol{r}_2 + \lambda_2 \hat{\boldsymbol{e}}_2 \end{aligned} \tag{3.1}$$

with separation:

$$\boldsymbol{L}_D = \boldsymbol{L}_2 - \boldsymbol{L}_1 \tag{3.2}$$

where ($\hat{e}_1$) and ($\hat{e}_2$) are the line-of-sight unit vectors, with (unscaled) pointing directions as indicated by the red and yellow arrows in Figure 3.3, while (λ_1) and (λ_2) are the scaling factors, determining the lengths along the lines-of-sight. The stereoscopic triangulation point ($\boldsymbol{T}_S$) is defined as the midpoint of the shortest possible line ($\tilde{\boldsymbol{L}}_D$) that separates ($\boldsymbol{L}_1$) and ($\boldsymbol{L}_2$).

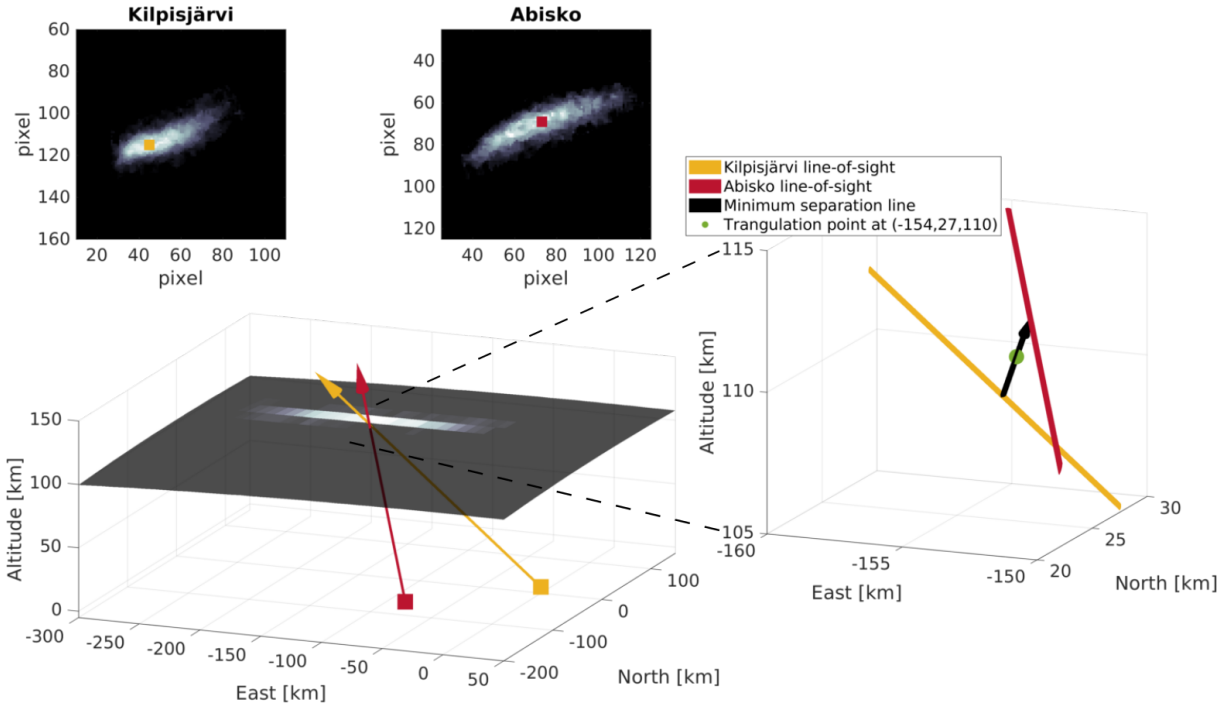


Figure 3.3: Triangulation of the Lumikko in the 5577 Å emission line. Top left, a representative Lumikko pixel is located in the Kilpisjärvi and Abisko images, determined by the position of the brightest pixel value of a (5×5) median filtered image version (not shown). Bottom left, the corresponding lines-of-sight from Kilpisjärvi (yellow) and Abisko (red) are drawn in the 3–D scene. To the right, a zoomed in illustration of the skew lines in proximity to the triangulation point (green dot), defined as the midpoint of the shortest possible line (black arrow) that separates the lines-of-sight.

Note that two lines generally do not intersect in 3–D (unlike the 2–D case) and that there exist no unique solution for the parallel case. Thus, assuming that the lines

($\boldsymbol{L}_1$) and ($\boldsymbol{L}_2$) are non-parallel, the shortest possible distance can be expressed as a minimization problem:

$$\left|\tilde{\boldsymbol{L}}_D\right| = \sqrt{\left((\boldsymbol{r}_2 + \tilde{\lambda}_2\hat{\boldsymbol{e}}_2) - (\boldsymbol{r}_1 + \tilde{\lambda}_1\hat{\boldsymbol{e}}_1)\right)^2} \tag{3.3}$$

where ($\tilde{\lambda}_1$) and ($\tilde{\lambda}_2$) are determined by;

$$\tilde{\lambda}_1, \tilde{\lambda}_2 = \arg\min_{\lambda_1,\lambda_2} \sqrt{((\boldsymbol{r}_2 + \lambda_2\hat{\boldsymbol{e}}_2) - (\boldsymbol{r}_1 + \lambda_1\hat{\boldsymbol{e}}_1))^2} \tag{3.4}$$

However, the minimization problem can be solved analytically, since the separation line ($\boldsymbol{L}_D$) is perpendicular to the skew lines when the distance is minimal, i.e: ($\hat{\boldsymbol{e}}_1 \perp \tilde{\boldsymbol{L}}_D$) and ($\hat{\boldsymbol{e}}_2 \perp \tilde{\boldsymbol{L}}_D$):

$$\hat{\boldsymbol{e}}_1 \cdot ((\boldsymbol{r}_2 + \tilde{\lambda}_2\hat{\boldsymbol{e}}_2) - (\boldsymbol{r}_1 + \tilde{\lambda}_1\hat{\boldsymbol{e}}_1)) = 0 \tag{3.5}$$

$$\hat{\boldsymbol{e}}_2 \cdot ((\boldsymbol{r}_2 + \tilde{\lambda}_2\hat{\boldsymbol{e}}_2) - (\boldsymbol{r}_1 + \tilde{\lambda}_1\hat{\boldsymbol{e}}_1)) = 0 \tag{3.6}$$

Solving these two linear equations for ($\tilde{\lambda}_1$) and ($\tilde{\lambda}_2$) yields ($\tilde{\boldsymbol{L}}_1$) and ($\tilde{\boldsymbol{L}}_2$), the points along the lines-of-sight that minimizes the separation distance. Thus, the triangulation point can be calculated in terms of ($\tilde{\boldsymbol{L}}_1$) and ($\tilde{\boldsymbol{L}}_2$):

$$\begin{aligned} \boldsymbol{T}_S &= \tilde{\boldsymbol{L}}_1 + \frac{\tilde{\boldsymbol{L}}_D}{2} \quad \text{or equivalently;} \quad \boldsymbol{T}_S = \tilde{\boldsymbol{L}}_2 - \frac{\tilde{\boldsymbol{L}}_D}{2} \\ \boldsymbol{T}_S &= \frac{(\boldsymbol{r}_1 + \tilde{\lambda}_1\hat{\boldsymbol{e}}_1) + (\boldsymbol{r}_2 + \tilde{\lambda}_2\hat{\boldsymbol{e}}_2)}{2} \end{aligned} \tag{3.7}$$

The triangulation point solution for the Lumikko example is presented in Table 3.1.

Table 3.1: The Lumikko coordinates with respect to the Kilpisjärvi imaging at the origin. The uncertainty is estimated by the length of the separation line ($\tilde{\boldsymbol{L}}_D$) components.

Technique	Stereoscopic triangulation
East [km]	–154 $\pm$ 1
North [km]	27 $\pm$ 1
Height [km]	110 $\pm$ 2

Multi-Viewpoint Triangulation

A more robust triangulation estimate can be obtained if the feature of interest is observed by more than two sites. For the multi-station imaging case with ($n > 2$) considered lines-of-sight, defining the set of 3–D lines;

$$\begin{aligned} \boldsymbol{L}_1 &= \boldsymbol{r}_1 + \lambda_1\hat{\boldsymbol{e}}_1 \\ \boldsymbol{L}_2 &= \boldsymbol{r}_2 + \lambda_2\hat{\boldsymbol{e}}_2 \\ &\vdots \\ \boldsymbol{L}_n &= \boldsymbol{r}_n + \lambda_n\hat{\boldsymbol{e}}_n \end{aligned} \tag{3.8}$$

there exist;

$$n_{\boldsymbol{L}_D} = \binom{n}{2} = \frac{n(n-1)}{2} \tag{3.9}$$

separation lines, and therefore no unique midpoint solution. The triangulation problem must therefore be reformulated. The triangulation point ($\boldsymbol{T}_M$) for the multi-station case can be defined as the estimated point ($\hat{\boldsymbol{P}}$) that minimizes the mean Euclidean distance to the (n) considered lines-of-sight:

$$\left|\tilde{\boldsymbol{L}}_D\right| = \frac{1}{n}\sqrt{\sum_{i=1}^{n}\left(\boldsymbol{T}_M - (\boldsymbol{r}_i + \tilde{\lambda}_i\hat{\boldsymbol{e}}_i)\right)^2} \tag{3.10}$$

where ($\boldsymbol{T}_M$), ($\tilde{\lambda}_1$), ($\tilde{\lambda}_2$) ... ($\tilde{\lambda}_n$) are determined by;

$$\boldsymbol{T}_M, \tilde{\lambda}_1, \tilde{\lambda}_2, ...\tilde{\lambda}_n = \arg\min_{\hat{\boldsymbol{P}},\lambda_1,\lambda_2,...\lambda_n} \frac{1}{n}\sqrt{\sum_{i=1}^{n}\left(\hat{\boldsymbol{P}} - (\boldsymbol{r}_i + \lambda_i\hat{\boldsymbol{e}}_i)\right)^2} \tag{3.11}$$

Equation 3.11 defines an over-determined system that can not be optimized analytically. However, an estimate can be obtained numerically by searching for the parameters ($\hat{\boldsymbol{P}}, \lambda_1, \lambda_2, ...\lambda_n$) that minimizes the mean Euclidean distance error function.

A suitable start guess ($\hat{\boldsymbol{P}}_0$) for the estimated triangulation point can be obtained by stereoscopic triangulation, preferably using the station-pair with the largest baseline. Figure 3.4 presents the best-fit results from a numerical search, applied on the images of artificial aurora observed by the ALIS stations near Abisko (–65.5, 58.3, –0.7), Kiruna (0.4, 1.7, –0.0), Silkimoutka (53.3, 21.6, –0.3) and Tjautjas (14.6, –56.6, –0.2), where the coordinates are given in kilometers in the Eastward, Northward and altitude direction with respect to the location of the Swedish Institute for Space Physics (IRF) in Kiruna, Sweden, at the origin.

The triangulation point solution for the artificial aurora example is presented in Table 3.2.

Table 3.2: The artificial aurora coordinates with respect to IRF at the origin. The uncertainties are estimated by the length of the mean Euclidean distance components ($\tilde{\boldsymbol{L}}_D$).

Technique	multi-station triangulation
East [km]	–52.6 $\pm$ 0.3
North [km]	151.8 $\pm$ 0.2
Height [km]	247.2 $\pm$ 0.2

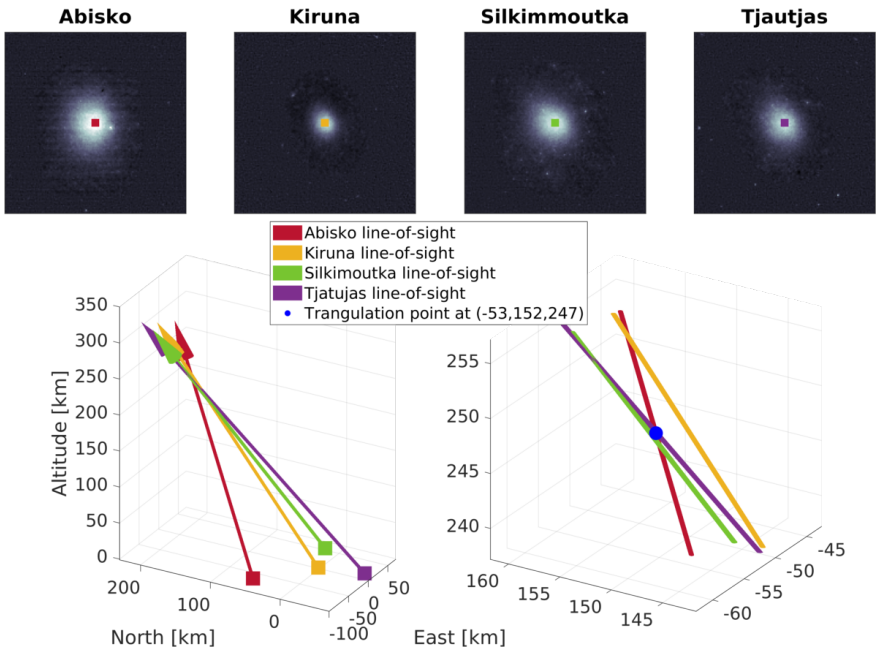


Figure 3.4: Triangulation of artificial aurora in the 6300 Å emission line. At the top, the simultaneous multi-station images from Abisko, Kiruna, Silkimoutka and Tjautjas. As for the Lumikko triangulation, the line-of-sight from each station is determined by the pointing direction of the brightest pixel value, marked in each image. Bottom left, the corresponding 3–D lines are drawn. Bottom right, the lines-of-sight in the vicinity of the triangulation point (blue dot) are illustrated. Note that the aurora patch appear smaller in the Kiruna image since the Kiruna field-of-view is 90°, while the other cameras have a 54° opening angle.

3.3 Shell Projection

The fundamental idea of the shell projection technique is that a feature, observed simultaneously from multiple sites with overlapping fields-of-view, can be mapped onto a common height-shell (in the 3–D scene). The mapped multi-station images are highly correlated if the projection height is close to the center of the true emission-altitude distribution. While poorly correlated projections indicate that the estimated height is far from the true center, this effect is depicted in Figure 3.5. Thus, the auroral height can be estimated by searching for the altitude that maximizes the correlation.

The shell projection technique was used to estimate the Lumikot height in Paper [III]. Previously, the method has been used to estimate the altitude of artificial aurora in Ashrafi et al. (2005) and the altitude of pulsating aurora in e.g. Partamies et al. (2017) and Partamies et al. (2019). Furthermore, the method was extensively analyzed for synthetic stereoscopic images of auroral arcs in Whiter et al. (2013).

The main advantage of the shell-projection technique is that the method can be automated and used to estimate the height of any auroral feature without assumptions. The main shortcoming is that the auroral emission has an extended altitude distribution and can not be projected onto an infinitesimal height-shell without er-

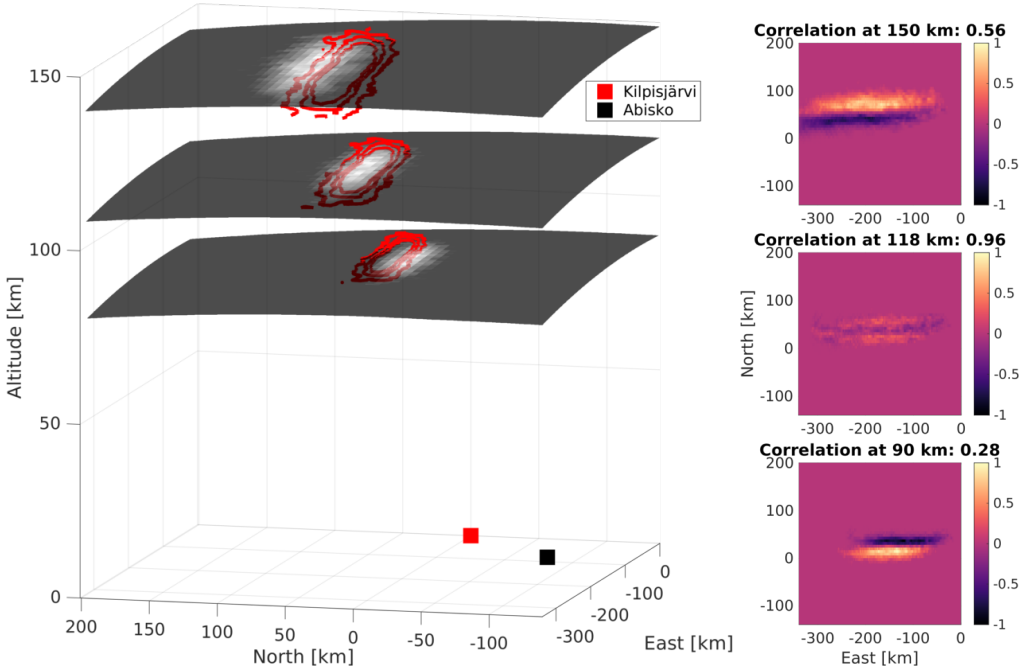


Figure 3.5: To the left, the Lumikko images from Kilpisjärvi (red intensity contours) and Abisko (black-and-white image), projected onto three height shells at; 90 km, 118 km and 150 km. The camera locations are indicated by the red box (Kilpisjärvi) and black box (Abisko). The Lumikko projections are poorly correlated at 90 km and 150 km (the mapped feature in the image and the contour plot do not overlap), but is well matched at 118 km. To the right, the residual of the Abisko projection subtracted by the Kilpisjärvi projection, a correlation value of one indicate fully overlapping images whereas a correlation of zero represent completely un-matched projections. Note that the altitude axis is not to scale with the xy-plane.

ror and loss of information. Whiter et al. (2013) reported an altitude bias of about –2 km, as compared to the peak-height of the synthetic arcs, almost independently of the evaluated arc widths and the location of the arc in the image-pairs.

Note that a horizontal height-layer has been assumed in the previous work, and that the technique therefore was named the horizontal plane method in Whiter et al. (2013). However, the analogous technique presented in this section is named the shell-projection method since the curvature of the Earth is taken into account.

First, the stereoscopic case is considered, with images from two ($n = 2$) separated stations; (A_1) and (A_2). The objective is to find the height (z) that maximizes the correlation (C). The correlation is evaluated on images projected onto a common height shell at altitude (z); (A_1^z) and (A_2^z), over all pixels (i, j):

$$C_{12}(z) = \frac{\sum_{i,j} \left(\left[A_1^z(i,j) - \bar{A}_1^z \right] \left[A_2^z(i,j) - \bar{A}_2^z \right] \right)}{\sqrt{\left(\sum_{i,j} \left[A_1^z(i,j) - \bar{A}_1^z \right]^2 \right) \left(\sum_{i,j} \left[A_2^z(i,j) - \bar{A}_2^z \right]^2 \right)}} \tag{3.12}$$

where ($\bar{A}_1^z$) and ($\bar{A}_1^z$) are the mean pixel values. ($C_{12} = 1$) indicate that the projected images are perfectly correlated, while ($C_{12} = 0$) represents completely un-matched projections.

Thus, the best-fit height ($\tilde{z}$) of the auroral feature, observed by a stereoscopic imaging system, can be estimated by solving the optimization problem:

$$\tilde{z}_2 = \arg\max_z \left(C_{12}(z)\right) \tag{3.13}$$

Equation 3.13 can be generalized for the multi-viewpoint ($2 \geq n$) case, where all (unordered) combinations of the projected image pairs are considered:

$$n_{\text{pairs}} = \binom{n}{2} = \frac{n(n-1)}{2} \tag{3.14}$$

The generalized correlation value optimization problem can be expressed as:

$$\begin{aligned}
\tilde{z}_n &= \arg\max_z \left(\frac{1}{n_{\text{pairs}}} \sum_{\substack{l,m=1 \\ l \neq m}}^{n} C_{lm}(z) \right) \\
&= \arg\max_z \left(\frac{1}{n_{\text{pairs}}} \sum_{\substack{l,m=1 \\ l \neq m}}^{n} \frac{\sum_{i,j} \left(\left[A_l^z(i,j) - \bar{A}_l^z \right] \left[A_m^z(i,j) - \bar{A}_m^z \right] \right)}{\sqrt{\left(\sum_{i,j} \left[A_l^z(i,j) - \bar{A}_l^z \right]^2 \right) \left(\sum_{i,j} \left[A_m^z(i,j) - \bar{A}_m^z \right]^2 \right)}} \right)
\end{aligned} \tag{3.15}$$

where (l) and (m) are the index of two separated imaging stations. Note however that if the feature of interest is observed by more than two stations, it is recommended to consider using the 3–D reconstruction technique, discussed in Section 3.4.

Equation 3.13 is used to estimate the best-fit height for the stereoscopic Lumikko case, by scanning over a typical auroral altitude range (between 90 km and 150 km) with 1 km resolution. In addition to the correlation value, the normalized summed residual value (R_{12}) is evaluated:

$$R_{12}(z) = \frac{\sum_{i,j} \sqrt{\left[A_1^z(i,j) - A_2^z(i,j) \right]^2}}{\sum_{i,j} \sqrt{\left[A_1^z(i,j) + A_2^z(i,j) \right]^2}} \tag{3.16}$$

The normalized residual value is used to quantify the relative error of the projected images at height (z), thus providing an estimate of the uncertainty of the best-fit height. The correlation and the relative error results are presented in Figure 3.6. Note that Equation 3.16 was used in the optimization task in Paper [III]. Both methods (maximizing the correlation or minimizing the normalized summed residuals) converge towards the same best-fit height, as seen in Figure 3.6.

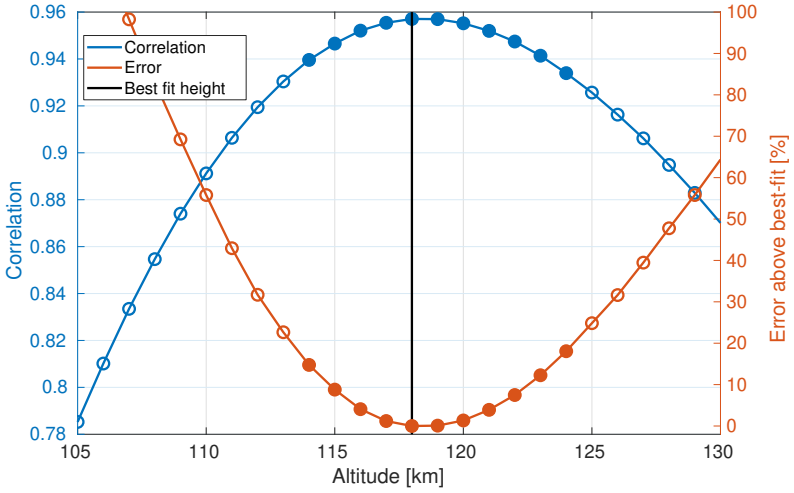


Figure 3.6: The correlation value results are indicated by the blue circles, obtained using Equation 3.12. The best-fit Lumikko height (118 km) is depicted as the black vertical line. The red circles indicate the residual error value, in per cent above the best-fit value. The projection results within 20 % of the best-fit error value, marked by filled circles, are used as an uncertainty estimate, corresponding to the altitude uncertainty interval; [114, 124].

In addition to the best-fit height, it is often useful to estimate the auroral feature coordinates and distribution in the xy-plane (i.e. the East-North-plane). This can be achieved by fitting a parameterized model to the feature of interest. The Lumikko aurora, projected onto the xy-plane, has a 2–D distribution that can be approximated by an asymmetric ($\sigma_x \neq \sigma_y$) 2–D Gaussian function:

$$f_{2D}(x,y) = I_0 \cdot \exp\left(-\left(\frac{((x-x_0)\cos(\phi) - (y-y_0)\sin(\phi))^2}{2\sigma_x^2} + \frac{((x-x_0)\sin(\phi) + (y-y_0)\cos(\phi))^2}{2\sigma_y^2} \right) \right) \tag{3.17}$$

where (I_0) is the midpoint intensity and (x_0, y_0) are the midpoint coordinates in the Eastward and Northward direction. (σ_x) and (σ_y) are the distribution widths and (ϕ) is the rotation angle. The 2–D distribution approximation is illustrated in Figure 3.7. The mean spatial parameters of the fitted Gaussians are used to quantify the coordinates and distribution widths in the East-North-plane. The overall results are presented in Table 3.3.

The resulting coordinates estimated by the shell-projection technique is different from the coordinates calculated by the stereoscopic triangulation, as seen in Table 3.3. The reason for the disagreement is that the methods calculate different representative coordinates. The triangulation technique estimate the 3–D coordinate of the brightest point within the Lumikko feature, while the shell projection method estimate a point similar to the centroid of the 3–D structure. In general, these points are not identical and the resulting coordinates are therefore not directly comparable.

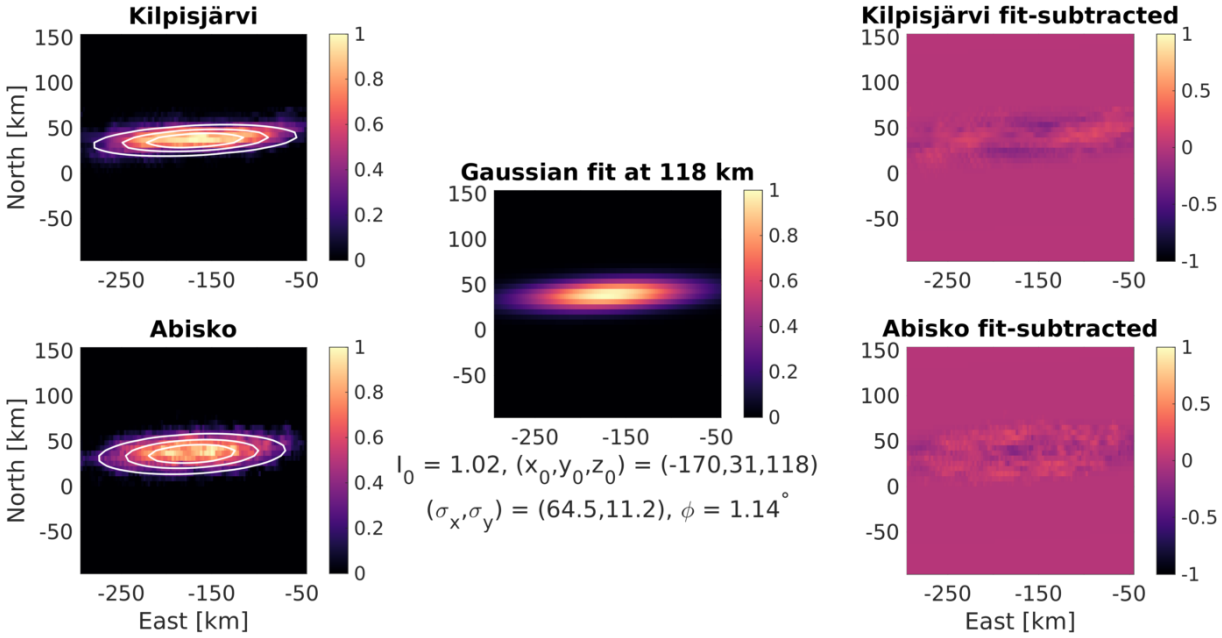


Figure 3.7: The 2–D Gaussian fit to the Lumikko projections in the xy-plane (i.e. the East-North-plane). To the left, the xy-projections of the Lumikko, from the best-fit height, observed at Kilpisjärvi and Abisko. The white contour lines indicate the best-fit of an asymmetric 2–D Gaussian. The Lumikko coordinates and distributions estimates in the xy-plane is determined by the mean values of the fitted Gaussians. In the middle, mean Gaussian parameter values are used to create a representative Lumikko distribution. To the right, the residual pixel values of the projected images subtracted by the representative 2–D Gaussian. The Gaussian distribution is a reasonable approximation, with no significant structures omitted by the fit, making the parameter results meaningful.

Table 3.3: The Lumikko coordinates with respect to the Kilpisjärvi station at the origin. The best-fit coordinate and distribution values are presented with error intervals, corresponding to the z-coordinate uncertainty interval; [114, 124]. The results from the triangulation method are added for comparison.

Technique	Stereoscopic triangulation	Shell projection
East [km]	-154 ± 1	–170 [–176, –166]
North [km]	27 ± 1	31 [28, 35]
Height [km]	110 ± 2	118 [114, 124]
Width West-East [km]		64.5 [62.2, 67.7]
Width South-North [km]		11.2 [10.8, 11.8]
Rotation [°]		1.14 [1.09, 1.18]

3.4 3–D Reconstruction

The full 3–D intensity structure of artificial aurora is estimated by tomography-like reconstruction in this section. The overall objective of the 3–D reconstruction is to estimate the emission distribution within a modeling volume, illustrated in Figure 3.8, so that the projections of the modeled aurora best fit the observations.

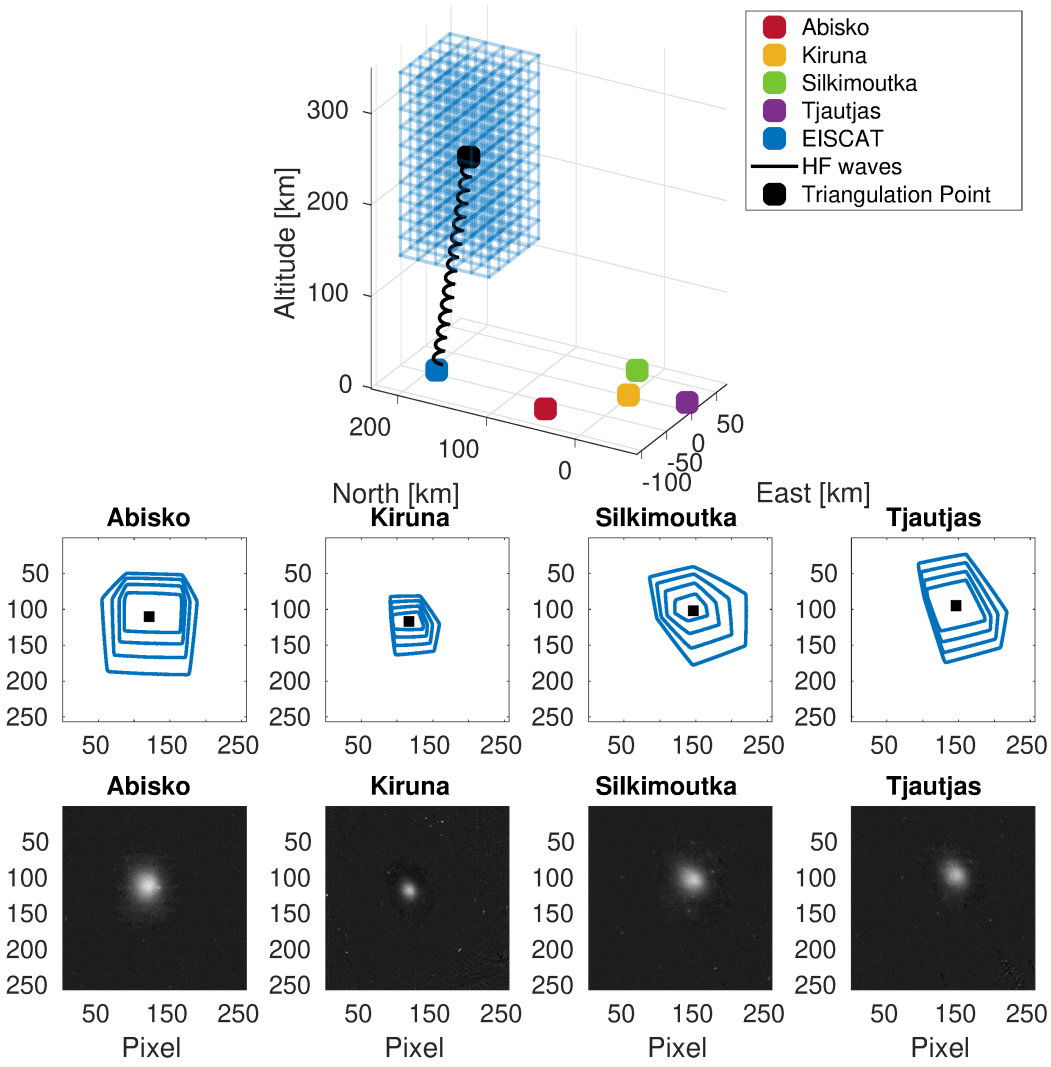


Figure 3.8: At the top, a diagram of the 3–D reconstruction problem geometry for the artificial aura example. The modeling volume is roughly defined by the volume ranging from ±50 km East, ±50 km North and ±100 km altitude of the triangulation point (black box). The HF radio waves, transmitted from the EISCAT Heater (blue box), is illustrated by the black line. In the middle, the isocontour projections of the modeling volume onto the the ALIS camera stations in Abisko, Kiruna, Silkimoutka and Tjautjas, the identified triangulation point is marked by the black square. At the bottom, the observed artificial aurora images in the 6300 Å line. The objective is to construct a 3–D aurora model with projections that match the observations.

In standard tomographic problems, such as medical scanning of brain tumors with Positron-Emission Tomography (PET) (Yen et al., 1982) or identifying bone fractures with Computed Tomography (CT) scanning (Shaffer et al., 1980), the object of interest is observed from as many positions as required until an adequate reconstruction is obtained. In reconstruction of auroral displays however, the object of interest is observed from a very limited number of positions (camera stations) and all camera stations are generally located at altitudes below the object of interest. Thus, the aurora is observed from a suboptimal set of viewing-directions, a set-up that makes the observations close to being linearly dependent, as described in Gustavsson et al. (2001).

The aurora reconstruction problem is therefore ill-posed and can not be adequately solved analytically, using e.g. the 3–D version of the inverse Radon transform (Radon, 1986), nor by conventional tomographic methods, using for example Algebraic Reconstruction Techniques (ART) (Gordon et al., 1970), without significant use of a priori information. Ill-posed, in this context, refers to a reconstruction problem that is sensitive to noise and either have no exact solution or an infinite number of solutions (Gustavsson, 2000). The ill-posed problem must therefore be regularized. A parameterized aurora model is used in this section to regularize the reconstruction problem and thus transforming the 3–D reconstruction task into a well-posed problem, in a restricted sense.

A parameterized 3–D reconstruction technique was used in paper [IV] to estimate the excitations rates of artificial aurora in the green (5577 Å), red (6300 Å) and infrared (8446 Å) lines. This method was first used to estimate the artificial aurora volume distribution in the red line in Gustavsson et al. (2001) and later in both the green and red lines in Gustavsson et al. (2008a). Pedersen et al. (2010) also used a parameterized aurora model to estimate the 5577 Å volume emission rates during very high power HF transmission at HAARP, causing artificial ionization and a descending wave–plasma resonance altitude. In addition, techniques for reconstructing the 3–D distribution of “naturally” occurring aurora have been studied, for example, by black aurora modeling in Gustavsson et al. (2008b) and by reconstruction of auroral arcs in e.g.; Frey et al. (1996), Frey et al. (1998), Aso et al. (1998) and Partamies et al. (2012).

The main advantage of the parameterized reconstruction technique is that the full 3–D intensity structure can be robustly estimated, using only a few unknown parameters. Moreover, the best-fit parameter values provide results that can be used to monitor the temporal evolution of the auroral structure and can further be compared to other events. The disadvantage, however, is that the emission distribution is heavily regularized, constraining the solution space. Thus, great care must be made when parameterizing the problem. There is a risk that the best-fit parameter results are meaningless if the aurora model do not adequately reconstruct the ‘true” auroral distribution. The projections of the aurora model must therefore be compared to the multi-viewpoint observations and evaluated before being presented as statistically significant 3–D emission rate reconstructions.

In this section, it is assumed that the artificial aurora can be (meaningfully) represented by an asymmetric ($\sigma_{xy} \neq \sigma_z$) 3–D Gaussian distribution:

$$f_{3D}(x,y,z) = I_0 \cdot \exp\left(-\left(\frac{(x-x_0)^2}{2\sigma_{xy}^2} + \frac{((y-y_0)+(z-z_0)\sin(\theta))^2}{2\sigma_{xy}^2} + \frac{(z-z_0)^2}{2\sigma_z^2} \right) \right) \tag{3.18}$$

where (I_0) is the maximum emission rate, located at the center of the emission volume (x_0, y_0, z_0). (σ_{xy}) and (σ_z) are the widths in the horizontal plane and along the altitude axis, while (θ) is magnetic field angle from zenith, $\sim$12 degrees South at the EISCAT Heating facility. This model was selected after inspecting the shape of the artificial aurora in the images, studying the results from previous work, e.g. Gustavsson et al. (2001) and Gustavsson and Eliasson (2008), and inference based on the shape of the transmitted HF beam pattern, as presented in paper [IV].

The aurora model can be constructed within the modeling volume by voxel elements (3–D cubes) where each voxel has a uniform intensity distribution, as suggested by Gordon (1974) and illustrated in Figure 3.8 by the depicted 20 km cubes. However, the voxel approximation results in a discontinuous intensity transition from one voxel to the next. This approach is therefore suboptimal for reconstructing smoothly varying objects (such as the 6300 Å artificial aurora). The voxels are thus exchanged in favor of overlapping three-dimensional spheroid elements with a symmetric ($\cos^2$) base function distribution, as suggested in Rydesäter and Gustavsson (2000) and further disused therein.

In the presented work, the spheroid centers are uniformly spaced within the modeling volume on a 3–D grid with 1 km resolution. The spheroids elements are selected to have a 2 km extent in each direction (0.5 km overlap to each neighbor), resulting in a fully spatially filling model space. The projected contribution from each spheroid element to a specified camera location and field-of-view can be accurately calculated by the dot-projection algorithm, proposed in Rydesäter and Gustavsson (2000). Model images of the constructed aurora distribution are thus formed by calculating and adding the projected intensity contribution from all ($51 \times 51 \times 101$) spheroid elements.

Figure 3.9 presents a 3–D aurora model, constructed by spheroid elements with intensities determined by the Gaussian distribution function. As a start guess, the triangulation point result ($\boldsymbol{T}_M$) from Section 3.2 was used as the center of the aurora model (x_0, y_0, z_0) and the intensity at the center (I_0) was set to $3{\times}10^6$ m^{-3}s^{-1}, similar to the 6300 Å emission rate results from Gustavsson et al. (2008a). Finally, a 20 km width was used both in the horizontal plane and along the magnetic field direction, reported as typical scale sizes in Gustavsson et al. (2001).

The high residual values in Figure 3.9 indicate that the constructed 3–D model is an inadequate reconstruction of the “true” aurora emission distribution. Once again, in order to obtain a better estimate, an optimization problem is defined. The goal is to optimize the Gaussian parameters such that the difference between the mod-

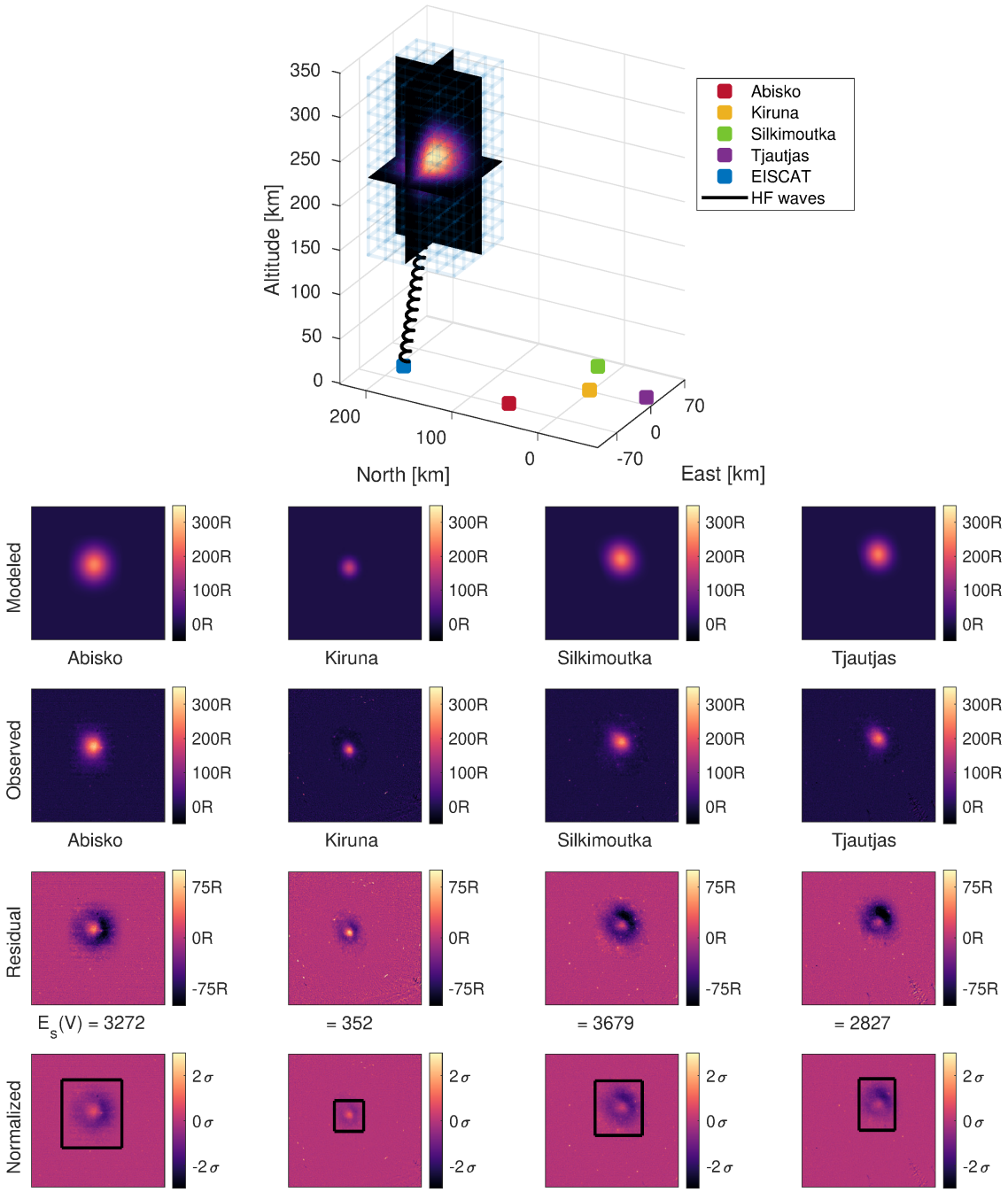


Figure 3.9: At the top, a 3–D aurora model constructed by the 3–D spheroid elements with intensities quantified by the Gaussian distribution function with guessed initial parameters. The projections of the 3–D aurora model to the location and field-of-view of the imaging stations are presented in the first row, calculated by the dot-projection algorithm. The intensity is presented in Rayleigh (R) units. The observed images are shown in the second row. The third row displays the residual emission, the observed images subtracted by the associated model projection. At the bottom, the residuals normalized by the standard deviation value (σ). The summed absolute normalized residual values, $E_s(\boldsymbol{V})$, over the indicated pixel region is provided as an error measure. The black-and-white images are artificially colored to better illustrate the residuals and highlight the quality of the fit.

eled images and the observed images is minimal. In this section, a weighted sum of the squared image residuals, $E(\boldsymbol{V})$, is used to evaluate the fit:

$$E(\boldsymbol{V}) = \frac{\sum_{s}^{n} \sum_{i,j} \left(A_s(i,j) - \hat{A}_s(i,j,I(\boldsymbol{V})) \right)^2}{\sum_{s}^{n} \sum_{i,j} \sigma_s^2(i,j)} \tag{3.19}$$

where ($\boldsymbol{V}$) denote the parameter vector for the 3–D Gaussian emission distribution $I(\boldsymbol{V})$. As before, (n) is the number of imaging stations and (s) is the station index. $A_s(i,j)$ denote the observed image at station (s) with pixel indices (i,j), while $\hat{A}_s(i,j,I(\boldsymbol{V}))$ represent the corresponding image projection of the 3–D model. Finally, $\sigma_s(i,j)$ is the pixel-dependent standard deviation, calculated from the noise characteristic of the CCD, known to be normally distributed $\mathcal{N}(\mu, I_c + 100)$, where ($I_c$) is the expected pixel intensity (Rydesäter et al., 2001). Note that the presented error function is different from the error function in Paper [IV], where it was assumed that the pixel uncertainty was equal over the entire CCD and the weights $\sigma_s(i,j)$ were therefore omitted.

In order to ease the analysis, a matrix representation of the error function is formulated. First, the observed data ($A_s(i,j)$) is reshaped into a 1–D vector (D) with size (n_D), containing all the pixel values from all stations. Secondly, the corresponding modeled data ($\hat{A}_s(i,j,I(\boldsymbol{V}))$) is reshaped into a 1–D vector $\hat{G}(\boldsymbol{V})$ with similar size. Finally, with this formulation, the sum of the squared residuals can be calculated by matrix multiplication (Theodoridis and Koutroumbas, 2009c):

$$\sum_{i}^{n_D} (D_i - G_i(\boldsymbol{V}))^2 = \left(D - \hat{G}(\boldsymbol{V}) \right)^T \left(D - \hat{G}(\boldsymbol{V}) \right) \tag{3.20}$$

Using this expression, the error function is re-written as:

$$E(\boldsymbol{V}) = (D_{obs} - \hat{G}(\boldsymbol{V}))^T \Sigma_D^{-1} (D_{obs} - \hat{G}(\boldsymbol{V})) \tag{3.21}$$

where (Σ_d) is the data covariance matrix, in this case a diagonal matrix of size ($n_D \times n_D$), containing the pixel intensity variance ($\sigma_s^2(i,j)$) information. The data covariance matrix is diagonal since the pixel intensities are only statistically correlated within the point-spread function, which in this case has a full-width half maximum ($<$1) pixel Gustavsson (2000). Thus, it is assumed that there is no pixel-to-pixel co-variance (perfect point-spread function) and the expressions in Equations 3.19 and 3.21 are therefore analogous.

The best-fit parameters are determined by minimizing $E(\boldsymbol{V})$, an optimization task that is often referred to as weighted least-squares fitting:

$$\tilde{\boldsymbol{V}} = \arg\min_{\boldsymbol{V}} \left(E(\boldsymbol{V}) \right) \tag{3.22}$$

where ($\tilde{\boldsymbol{V}}$) denotes the best-fit parameters. The best-fit aurora model is presented in Figure 3.10. The Gaussian parameters are optimized by the Nelder-Mead simplex search algorithm (Lagarias et al., 1998; Nelder and Mead, 1965) and reaches a

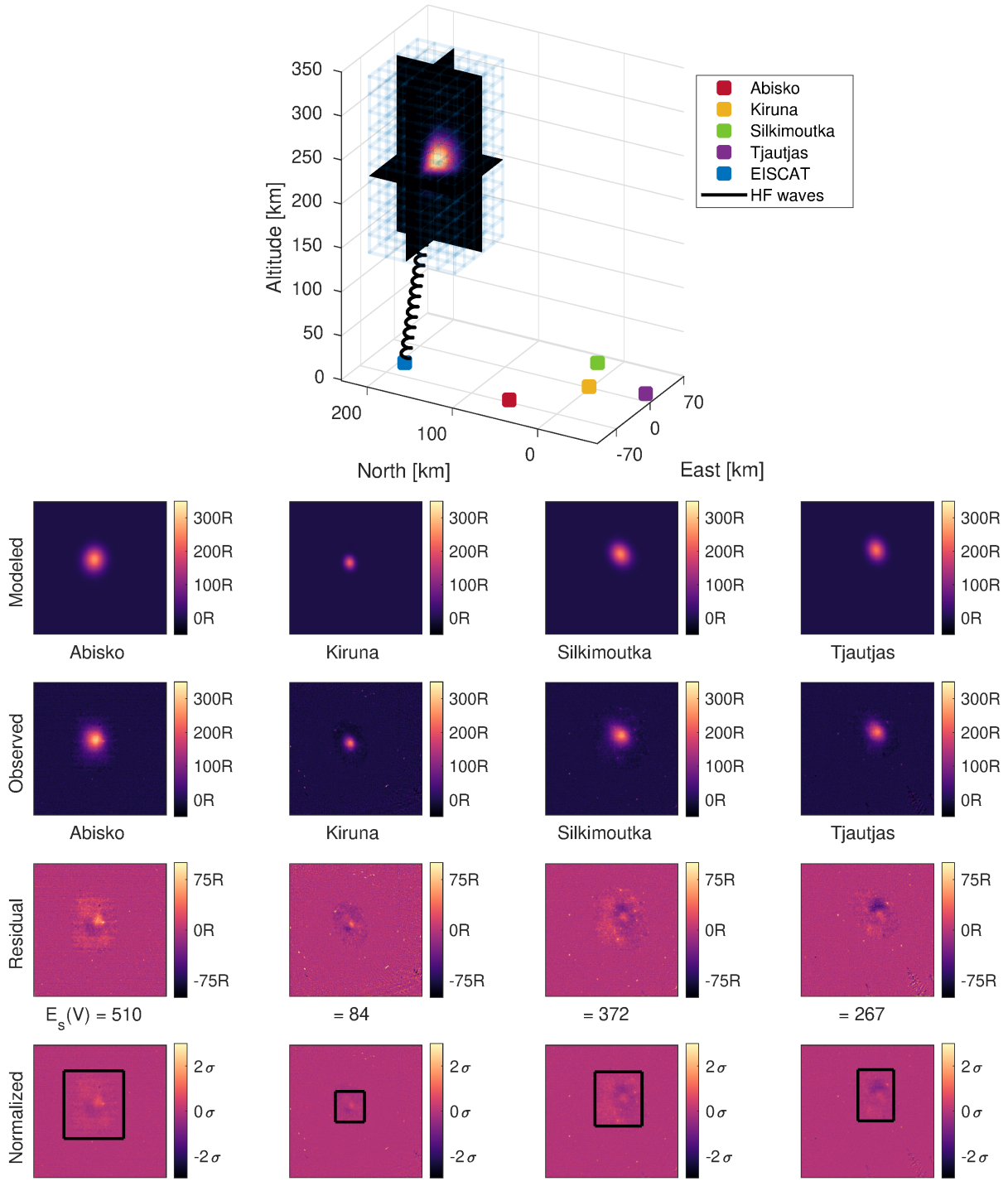


Figure 3.10: An illustration of the best-fit aurora model. Although some residuals remain, there are no large clusters of pixels with high (σ) values. Thus, the fit is meaningful and the optimized parameters can be used as a statistically significant estimate of the "true" auroral volume distribution.

minimum after 279 iterations (∼2 min of calculation time on a laptop with a 3GHz i5 dual core processor).

The best-fit aurora model has projections that adequately fit the observed images. However, the uncertainty of the optimized parameters remains to be determined. The uncertainty can be estimated by calculating the error function response to small changes in the best-fit parameters, $E(\tilde{\boldsymbol{V}} + \Delta\boldsymbol{V})$. This task can be mathematically expressed in terms of the discretized Jacobian matrix (Aster et al., 2013a):

$$J = \frac{\Delta E(\tilde{\boldsymbol{V}})}{\Delta \boldsymbol{V}} \tag{3.23}$$

(J) is a ($n_D \times n_P$) dimensional matrix, where (n_P) is the number of partial derivatives to evaluate, six for the Gaussian model. A change of (ΔI_0 = $10^5 \mathrm{m}^{-3}\mathrm{s}^{-1}$), ($\Delta x_0$ = Δy_0 = Δz_0 = 1 km) and ($\Delta\sigma_{xy}$ = $\Delta\sigma_z$ = 0.5 km) is selected. The effect of the parameter change to a modeled image is presented in Figure 3.11. Here it is worth noting that increasing or decreasing the Δ-values made no significant change in the parameter uncertainty results, tested within reasonable bounds for Δ-scaling factors of 0.5, 2, 4 and 10.

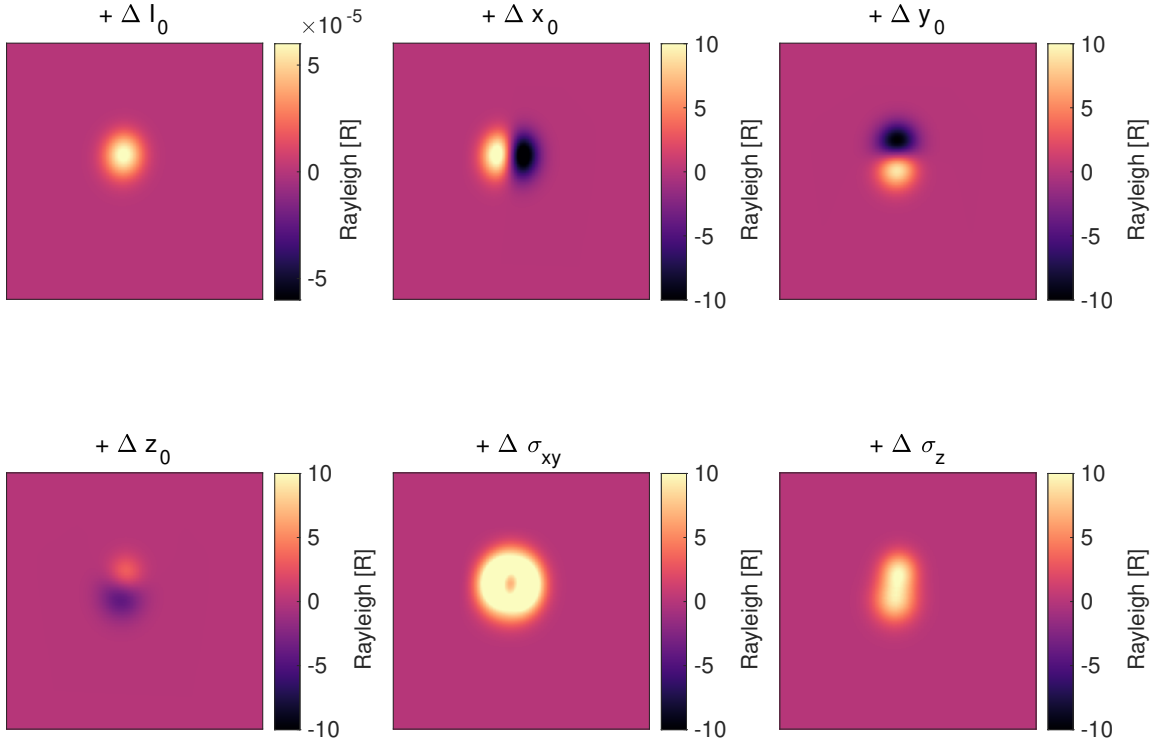


Figure 3.11: The effect of the change (ΔI_0 = $10^5 \mathrm{m}^{-3}\mathrm{s}^{-1}$), ($\Delta x_0$ = Δy_0 = Δz_0 = 1 km) and $\Delta\sigma_{xy}$ = $\Delta\sigma_z$ = 0.5 km) to a modeled image, projected to the location and field-of-view in Abisko. The (Δy_0) and (Δz_0) derivatives appears to be highly correlated for the Abisko imaging station. Also, note that the effect of the (ΔI_0) change is much smaller than for the other derivatives.

The parameter covariance matrix ($\Sigma_{\tilde{\boldsymbol{V}}}$) can now be determined by the data variance matrix (Σ_d) and the Jacobian derivatives (J), see Aster et al. (2013a) for details:

$$\Sigma_{\tilde{\boldsymbol{V}}} = (J^T \Sigma_d^{-1} J)^{-1} \tag{3.24}$$

$\Sigma_{\tilde{\boldsymbol{V}}}$ is a ($n_P \times n_P$) dimensional matrix, in this case (6 $\times$ 6), where the diagonal entries are the parameter variances ($\sigma^2_{\tilde{\boldsymbol{V}}}$), assuming that the estimated parameters themselves behave like random variables. The best-fit parameters and the corresponding uncertainties are presented in Table 3.4.

Table 3.4: The artificial aurora parameters. The East-North coordinates are presented with respect to the location of the Swedish Institute for Space Physics in Kiruna, Sweden, at the origin. The best-fit 3–D reconstruction parameters are presented with $1\pm\sigma_{\tilde{\boldsymbol{V}}}$ uncertainties. The results from the multi-viewpoint triangulation method are added for comparison.

Technique	multi-station triangulation	3–D reconstruction
East [km]	–52.6 $\pm$ 0.3	–49.3 $\pm$ 0.1
North [km]	151.8 $\pm$ 0.2	150.3 $\pm$ 0.3
Height [km]	247.2 $\pm$ 0.2	243.9 $\pm$ 0.5
Horizontal width [km]		13.2 $\pm$ 0.1
Altitude width [km]		16.5 $\pm$ 0.3
Intensity [$10^6 m^{-3} s^{-1}$]		4.0 $\pm$ 0.1

The resulting coordinates from the 3–D reconstruction is slightly different from the multi-viewpoint triangulation results. The reason for this difference is likely that the auroral emission distribution can not be fully represented by a 3–D Gaussian, thus the coordinate of maximum emission might be off-set from the estimated Gaussian centroid.

It should be noted that a better fit can be obtained if more shape parameters are added to the model function, e.g. by allowing the model to have different widths below and above the height center or by making the widths independent along the x– and y-axis.

However, using an excessively complex model should be avoided in order to prevent high parameter correlation (and thus high uncertainty) and model aliasing. Model aliasing, in this case, refers to a situation where several parameter combinations result in similar projections, complicating the parameter estimation and the interpretation of the resulting model.

Model aliasing effects can be studied by performing several parameter optimization runs with different initializations (start guesses) and comparing the results. 10 optimization runs were tested for the artificial aurora example and all runs converged towards the same parameter solution (although some initializations required a larger number of iteration steps). It can therefore be concluded that the selected set of parameters are not prone to aliasing effects.

Moreover, the parameter correlation can be determined by analyzing the parameter covariance matrix. The correlation of two parameters is the scaled covari-

ance (Aster et al., 2013b), calculated using $\Sigma_{\tilde{\boldsymbol{V}}}(i,j)$ for parameters with indices (i) and (j) as:

$$C_{ij} = \frac{\Sigma_{\tilde{\boldsymbol{V}}}(i,j)}{\sqrt{\Sigma_{\tilde{\boldsymbol{V}}}(i,i)\Sigma_{\tilde{\boldsymbol{V}}}(j,j)}} \tag{3.25}$$

The resulting correlation values for all parameter-pairs are presented by the correlation matrix in Figure 3.12. The (y_0) and (z_0) parameters are highly correlated, as expected from studying the Δy_0 and Δz_0 derivatives in Figure 3.11.

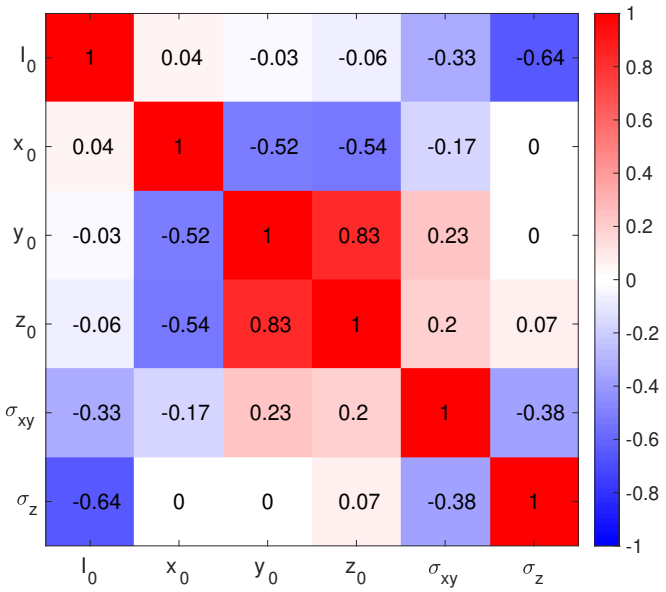


Figure 3.12: The $\sigma_{\tilde{\boldsymbol{V}}}$ Pearson correlation values. The (y_0) and (z_0) parameters are highly correlated due to the unfortunate position of the camera stations, all located South and below of the studied auroral emission.

The reason for the high (y_0)–(z_0) correlation is that all imaging stations are located South and below of the object of interest, as seen in the top panel of Figure 3.10, making the emission distribution somewhat ambiguous in the North-altitude plane. This uncertainty could be avoided by adding observations from another camera, seeing the object of interest in a different plane, preferably located Westward, Eastward or directly below the object of interest.

Chapter 4

Conclusion and Outlook

Analysis of big data sets is a rapidly progressing field of research with a largely untapped potential in auroral science. This thesis has presented auroral image processing techniques that are suitable for big data sets, requiring a minimal amount of user interaction. The processing techniques can be divided into two branches: *machine learning classification methods*, evaluated and discussed in Chapter 2, and *multi-viewpoint analysis techniques*, studied in Chapter 3.

4.1 Concluding Remarks

Auroral Image Classification with Machine Learning Methods

Three machine learning classifiers were trained, tested and compared: K-Nearest Neighbor (K-NN), Support Vector Machine (SVM) and Convolutional Neural Network (CNN). The main conclusion of the presented work is that the CNN is a particularly suitable classifier for auroral image data and outperforms the K-NN and SVM classifiers.

The enhanced performance of the CNN is due to the superior feature extraction capability. The CNN optimizes the feature extraction process for the available training data, while the K-NN and SVM classifiers require pre-defined feature vectors with high discriminative power, which are challenging to obtain from auroral images.

Six different CNN architectures were evaluated in Paper [I]. All networks achieved precisions above 80 % and the ResNet–50 architecture stood out with the highest performance, obtaining an overall precision of 92 % and an average class-wise accuracy of 91 %. Thus, the ResNet–50 classifier is adequate for operational purposes, with error rates below 10 % (Syrjasuo and Partamies, 2011).

The networks were trained and tested on image data without clearly ambiguous auroral forms. Classification of ambiguous images is problematic since the “correct” label is a matter of subjective interpretation and possible bias factors. Paper [II] discusses the implications of biases in the training and testing data for machine learning classification and further provides suggestions on how to manually label image data with high integrity.

Multi-Viewpoint Analysis of Auroral Images

The overall objective of multi-viewpoint analysis techniques is to obtain 3–D auroral information from the 2–D image projections. Three such methods have been evaluated and described: triangulation, shell projection and 3–D reconstruction.

The appropriate technique for a particular auroral study depends on the investigated feature, the camera system geometry and background conditions, but the general conclusions are:

- The triangulation method can be largely automated and is suitable for estimating the 3–D position of the peak auroral intensity. This information can further be used as a priori input to more advanced multi-viewpoint analysis techniques.
- The shell projection method can be fully automated and used without a priori information. The method is suitable for for estimating the mean height and the horizontally projected emission distribution of auroral displays.
- The 3–D reconstruction method estimates the full volumetric auroral distribution and can be fully automated. The method does however require heavy regularization.

The shell projection method was used in Paper [III] to estimate the height (115–120 km) and the horizontally projected midpoint coordinate of the newly reported auroral feature: Lumikot aurora. A velocity estimate ($\approx$ 12 km/s Eastward or Westward) was inferred by calculating the coordinate displacement in consecutive images. The processes that cause the bidirectional movement of Lumikot aurora, parallel to the growth-phase arc, is unknown and may be significant for understanding the mechanisms leading to the substorm onset.

The volume excitation rates of artificial aurora were estimated by the 3–D reconstruction method in Paper [IV]. The reconstruction results are significant since the altitude distribution of the excitation rates does not agree with theoretical predictions from previous work (Gustavsson and Eliasson, 2008). An anisotropic electron acceleration, caused by the radio wave–plasma interaction, may explain the discrepancy.

4.2 Outlook

The Auroral Morphology

Auroral image classification has great potential for statistically relating auroral forms to magnetospheric–ionospheric coupling processes and conditions. However, this field of research is constrained by subjective interpretations and biases. In the presented work, approximately two-thirds of the auroral images were labeled as having an ambiguous form, and therefore not used in the training and testing of the classifiers.

After Akasofu (1964) characterized the dynamic morphology of auroral substorms, numerous auroral features (e.g. auroral streamers, auroral fingers, Sun-aligned arcs, black aurora and flickering aurora) have been reported in addition to, or as sub-classes of, the initial classes. However, it is not always clear (without an extensive literature search) how these classes relate to each other, to magnetospheric–ionospheric conditions and to the auroral substorm in general. The communication of auroral research is further complicated since auroral forms often are referred to by different terms. For example, Sun-aligned arcs also go by the names: high-latitude polar arcs, theta auroras, teardrops, and horse collar auroras (Elphinstone et al., 1996). Overall, these problems make auroral science prone to confusion, ambiguity, misinterpretation and further makes it difficult to introduce new scientists to the field.

Formulating an updated auroral morphology would therefore greatly benefit the auroral research community in general, but also the field of auroral image classification in particular, and should be considered in future work. The morphological diagram could be branched, including definitions of the auroral classes and subclasses, illustrations of typical displays and short descriptions of the relationship to processes and conditions in the magnetosphere–ionosphere (when applicable). Inspiration could be taken from morphological classification of snowflakes (Nakaya diagram), stars (Hertzsprung–Russell diagram) or galaxies (Hubble sequence).

Suggestions for Future Endeavors in Auroral Classification

The aurora is a dynamic phenomena and the auroral classes (e.g. *arcs*, *breakup* and *patchy*) have temporal characteristics. Thus, including the time dimension will likely enhance the classification performance. In practice, this can be achieved using a CNN, where 3–D volumes (short videos) with stacked consecutive images (e.g. 3–10 frames in each volume) are labeled by the most prominent auroral feature in the sample video, and used instead of single image frames.

In addition, the classification performance on mixed data sets, including images from several cameras and locations, should be studied in order to detect possible camera-specific bias effects. It is crucial that biases are minimized in order to obtain reliable statistical results from the classified auroral images. The camera-specific bias effects can be reduced by image pre-processing and by adding labeled images from different cameras to the training data.

A pre-processing procedure, standardizing the image input to the classifier, was suggested in Section 2.1. However, only careful examination of the classification performance on mixed data sets can determine if the suggested pre-processing procedure is adequate.

3–D Auroral Reconstruction in Conjunction with EISCAT3D

At present, there is a new incoherent scatter radar under construction in Northern Fenno-Scandinavia: EISCAT3D, and its “first light” observations are expected in the next few years (Kero et al., 2019). When completed, it is anticipated to be

the world's most advanced incoherent scatter radar, providing volumetric measurements of the ionospheric plasma parameters with high spatial and temporal resolution (McCrea et al., 2015).

For auroral science, the EISCAT3D radar will make it feasible to study the 3–D structural evolution of auroral forms with sub-second resolution based on both volumetric plasma measurements and 3–D auroral reconstructions. Moreover, the energy and spatial distribution of the precipitating electrons can be estimated by combining the plasma measurements and the optical observations, e.g. using the Generalized-Aurora Computed Tomography (G-ACT) method, proposed in Tanaka et al. (2011), or by using the 3–D inversion technique presented in Simon Wedlund et al. (2013).

Finally, studies of artificial aurora with EISCAT3D in conjunction with optical observations will provide new and exciting methods for examining the radio wave–plasma interaction. In particular, the rise and fall of plasma cavities can be investigated in detail using aperture-synthesis radar imaging techniques (Stamm et al., 2021), and analysis of the plasma cavity formation in relation to 3–D excitation rate estimates will likely give new insights to the electron energization process.

References

Akasofu, S.-I.: The development of the auroral substorm, Planetary and Space Science, 12, 273–282, 1964.

Akasofu, S.-I.: Auroral morphology: A historical account and major auroral features during auroral substorms, Auroral phenomenology and magnetospheric processes: earth and other planets, 197, 29–38, 2012.

Akasofu, S.-I.: The relationship between the magnetosphere and magnetospheric/auroral substorms, Annales Geophysicae, 31, 387–394, doi: 10.5194/angeo-31-387-2013, URL https://angeo.copernicus.org/articles/31/387/2013/, 2013.

Akasofu, S.-I.: Auroral substorms: search for processes causing the expansion phase in terms of the electric current approach, Space Science Reviews, 212, 341–381, 2017.

Amin, J., Sharif, M., Yasmin, M., and Fernandes, S. L.: Big data analysis for brain tumor detection: Deep convolutional neural networks, Future Generation Computer Systems, 87, 290–297, 2018.

Angelopoulos, V., McFadden, J. P., Larson, D., Carlson, C. W., Mende, S. B., Frey, H., Phan, T., Sibeck, D. G., Glassmeier, K.-H., Auster, U., et al.: Tail reconnection triggering substorm onset, Science, 321, 931–935, 2008.

Antiochos, S., Mikić, Z., Titov, V., Lionello, R., and Linker, J.: A model for the sources of the slow solar wind, The Astrophysical Journal, 731, 112, 2011.

Ashrafi, M., Kosch, M., and Kaila, K.: Height triangulation of artificial optical emissions in the F-layer, in: Proceedings of the 31st Annual European Meeting on Atmospheric Studies by Optical Methods and 1st International Riometer Workshop, pp. 8–16, 2005.

Aso, T., Ejiri, M., Urashima, A., Miyaoka, H., Steen, Å., Brändström, U., and Gustavsson, B.: First results of auroral tomography from ALIS-Japan multi-station observations in March, 1995, Earth, planets and space, 50, 81–86, 1998.

Aster, R. C., Borchers, B., and Thurber, C. H.: Chapter Nine - Nonlinear Regression, in: Parameter Estimation and Inverse Problems (Second Edition), edited by Aster, R. C., Borchers, B., and Thurber, C. H., pp. 217–238, Academic Press, Boston, second edition edn., doi: https://doi.org/10.1016/

B978-0-12-385048-5.00009-4, URL https://www.sciencedirect.com/science/article/pii/B9780123850485000094, 2013a.

Aster, R. C., Borchers, B., and Thurber, C. H.: Appendix B - Review of Probability and Statistics, in: Parameter Estimation and Inverse Problems (Second Edition), edited by Aster, R. C., Borchers, B., and Thurber, C. H., pp. 315–337, Academic Press, Boston, second edition edn., doi: https://doi.org/10.1016/B978-0-12-385048-5.00014-8, URL https://www.sciencedirect.com/science/article/pii/B9780123850485000148, 2013b.

Baker, D. e., Fritz, T., McPherron, R., Fairfield, D., Kamide, Y., and Baumjohann, W.: Magnetotail energy storage and release during the CDAW 6 substorm analysis intervals, Journal of Geophysical Research: Space Physics, 90, 1205–1216, 1985.

Balaban, S.: Deep learning and face recognition: the state of the art, in: Biometric and Surveillance Technology for Human and Activity Identification XII, vol. 9457, p. 94570B, International Society for Optics and Photonics, 2015.

Baumjohann, W. and Treumann, R.: Basic Space Plasma Physics (Revised Edition), World Scientific Publishing Company, 2012.

Bazilchuk, Z. S.: Angular dependence of wide altitude ion line enhancements (WAILEs) during ionospheric heating at the EISCAT Tromsø Facility, Master's thesis, UiT Norges arktiske universitet, 2019.

Bernhardt, P. A., Siefring, C. L., Briczinski, S. J., McCarrick, M., and Michell, R. G.: Large ionospheric disturbances produced by the HAARP HF facility, Radio Science, 51, 1081–1093, 2016.

Biondi, A., Sipler, D., and Hake, R.: Optical (λ6300) detection of radio frequency heating of electrons in the F region, Journal of Geophysical Research, 75, 6421–6424, 1970.

Bissacco, A., Cummins, M., Netzer, Y., and Neven, H.: PhotoOCR: Reading Text in Uncontrolled Conditions, in: Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2013.

Blagoveshchenskaya, N.: Perturbing the high-latitude upper ionosphere (F region) with powerful HF radio waves: A 25-year collaboration with EISCAT, URSI Radio Science Bulletin, 2020, 40–55, 2020.

Brändström, B., Leyser, T., Steen, Å., Rietveld, M., Gustavsson, B., Aso, T., and Ejiri, M.: Unambiguous evidence of HF pump-enhanced airglow at auroral latitudes, Geophysical research letters, 26, 3561–3564, 1999.

Brändström, U.: The auroral large imaging system: design, operation and scientific results, IRF Institutet för rymdfysik, 2003.

Brekke, A.: Physics of the upper polar atmosphere, Springer Science & Business Media, 2012.

Buetti-Dinh, A., Galli, V., Bellenberg, S., Ilie, O., Herold, M., Christel, S., Boretska, M., Pivkin, I. V., Wilmes, P., Sand, W., et al.: Deep neural networks outperform human expert's capacity in characterizing bioleaching bacterial biofilm composition, Biotechnology Reports, 22, e00 321, 2019.

Carlson, H. C., Gordon, W. E., and Showen, R. L.: High frequency induced enhancements of the incoherent scatter spectrum at Arecibo, Journal of Geophysical Research, 77, 1242–1250, 1972.

Chen, F. F. et al.: Introduction to plasma physics and controlled fusion, vol. 1, Springer, 1984.

Clausen, L. B. and Nickisch, H.: Automatic classification of auroral images from the Oslo Auroral THEMIS (OATH) data set using machine learning, Journal of Geophysical Research: Space Physics, 123, 5640–5647, 2018.

Colpitts, C. A., Hakimi, S., Cattell, C. A., Dombeck, J., and Maas, M.: Simultaneous ground and satellite observations of discrete auroral arcs, substorm aurora, and Alfvénic aurora with FAST and THEMIS GBO, Journal of Geophysical Research: Space Physics, 118, 6998–7010, 2013.

Cortes, C. and Vapnik, V.: Support-vector networks, Machine learning, 20, 273–297, 1995.

Coster, A., Djuth, F., Jost, R., and Gordon, W.: The temporal evolution of 3-m striations in the modified ionosphere, Journal of Geophysical Research: Space Physics, 90, 2807–2818, 1985.

Dahlgren, H., Ivchenko, N., Sullivan, J., Lanchester, B. S., Marklund, G., and Whiter, D.: Morphology and dynamics of aurora at fine scale: first results from the ASK instrument, Annales Geophysicae, 26, 1041–1048, doi: 10.5194/angeo-26-1041-2008, URL https://angeo.copernicus.org/articles/26/1041/2008/, 2008.

Dalal, N. and Triggs, B.: Histograms of oriented gradients for human detection, in: 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05), vol. 1, pp. 886–893, Ieee, 2005.

Davis, T. N.: Observed characteristics of auroral forms, Space Science Reviews, 22, 77–113, 1978.

D'Errico, J.: Interpolate NaN elements, URL https://se.mathworks.com/matlabcentral/fileexchange/4551-inpaint_nans, 2012.

Deweerdt, S.: Deep connections, Nature, 571, S6–S8, 2019.

Djuth, F., Isham, B., Rietveld, M., Hagfors, T., and La Hoz, C.: First 100 ms of HF modification at Tromsø, Norway, Journal of Geophysical Research: Space Physics, 109, 2004.

Donovan, E., Liu, W., Liang, J., Spanswick, E., Voronkov, I., Connors, M., Syrjäsuo, M., Baker, G., Jackel, B., Trondsen, T., et al.: Simultaneous THEMIS in situ and auroral observations of a small substorm, Geophysical research letters, 35, 2008.

Dungey, J. W.: Interplanetary magnetic field and the auroral zones, Physical Review Letters, 6, 47, 1961.

Ebrahimzadeh, R. and Jampour, M.: Efficient handwritten digit recognition based on histogram of oriented gradients and SVM, International Journal of Computer Applications, 104, 2014.

Elphinstone, R., Murphree, J., and Cogger, L.: What is a global auroral substorm?, Reviews of Geophysics, 34, 169–232, 1996.

Fang, X., Randall, C. E., Lummerzheim, D., Solomon, S. C., Mills, M. J., Marsh, D. R., Jackman, C. H., Wang, W., and Lu, G.: Electron impact ionization: A new parameterization for 100 eV to 1 MeV electrons, Journal of Geophysical Research: Space Physics, 113, 2008.

Farley Jr, D.: Artificial heating of the electrons in the F region of the ionosphere, Journal of Geophysical Research, 68, 401–413, 1963.

Feldstein, Y.: A quarter of a century with the auroral oval, Eos, Transactions American Geophysical Union, 67, 761–767, 1986.

Feldstein, Y., Vorobjev, V., Zverev, V., and Förster, M.: Investigations of the auroral luminosity distribution and the dynamics of discrete auroral forms in a historical retrospective, History of geo-and space sciences, 5, 81–134, 2014.

Frey, H., Frey, S., Lanchester, B., and Kosch, M.: Optical tomography of the aurora and EISCAT, Annales Geophysicae, 16, 1332–1342, 1998.

Frey, H. U., Han, D., Kataoka, R., Lessard, M. R., Milan, S. E., Nishimura, Y., Strangeway, R. J., and Zou, Y.: Dayside aurora, Space Science Reviews, 215, 1–32, 2019.

Frey, S., Frey, H., Carr, D., Bauer, O. H., and Haerendel, G.: Auroral emission profiles extracted from three-dimensionally reconstructed arcs, Journal of Geophysical Research: Space Physics, 101, 21 731–21 741, 1996.

Gallardo-Lacourt, B., Liang, J., Nishimura, Y., and Donovan, E.: On the origin of STEVE: Particle precipitation or ionospheric skyglow?, Geophysical Research Letters, 45, 7968–7973, 2018.

Gallardo-Lacourt, B., Frey, H., and Martinis, C.: Proton Aurora and Optical Emissions in the Subauroral Region, Space Science Reviews, 217, 1–36, 2021.

Ganushkina, N. Y., Liemohn, M., and Dubyagin, S.: Current systems in the Earth's magnetosphere, Reviews of Geophysics, 56, 309–332, 2018.

Gilmore, F. R., Laher, R. R., and Espy, P. J.: Franck–Condon factors, r-centroids, electronic transition moments, and Einstein coefficients for many nitrogen and oxygen band systems, Journal of physical and chemical reference data, 21, 1005–1107, 1992.

Goenka, C., Semeter, J., Noto, J., Baumgardner, J., Riccobono, J., Migliozzi, M., Dahlgren, H., Marshall, R., Kapali, S., Hirsch, M., et al.: LiCHI–Liquid Crystal Hyperspectral Imager for simultaneous multispectral imaging in aeronomy, Optics express, 23, 17 772–17 782, 2015.

Gondarenko, N., Ossakow, S., and Milikh, G.: Generation and evolution of density irregularities due to self-focusing in ionospheric modifications, Journal of Geophysical Research: Space Physics, 110, 2005.

Gordon, R.: A tutorial on ART (algebraic reconstruction techniques), IEEE Transactions on Nuclear Science, 21, 78–93, 1974.

Gordon, R., Bender, R., and Herman, G. T.: Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography, Journal of theoretical Biology, 29, 471–481, 1970.

Gosling, J. T.: The solar wind, in: Encyclopedia of the solar system, pp. 261–279, Elsevier, 2014.

Grach, S. M., Kosch, M. J., Yashnov, V. A., Sergeev, E. N., Atroshenko, M. A., and Kotov, P. V.: On the location and structure of the artificial 630-nm airglow patch over Sura facility, Annales Geophysicae, 25, 689–700, doi: 10.5194/angeo-25-689-2007, URL `https://angeo.copernicus.org/articles/25/689/2007/`, 2007.

Gurevich, A. V.: Nonlinear effects in the ionosphere, Physics-Uspekhi, 50, 1091, 2007.

Gustavsson, B.: Three dimensional imaging of aurora and airglow, 2000.

Gustavsson, B.: AIDA-tools, URL `https://github.com/space-physics/AIDA-tools`, 2015.

Gustavsson, B. and Eliasson, B.: HF radio wave acceleration of ionospheric electrons: Analysis of HF-induced optical enhancements, Journal of Geophysical Research: Space Physics, 113, 2008.

Gustavsson, B., Sergienko, T., Rietveld, M., Honary, F., Steen, A., Brändström, B., Leyser, T., Aruliah, A., Aso, T., Ejiri, M., et al.: First tomographic estimate of volume distribution of HF-pump enhanced airglow emission, Journal of Geophysical Research: Space Physics, 106, 29 105–29 123, 2001.

Gustavsson, B., Sergienko, T., Kosch, M. J., Rietveld, M. T., Brändström, B. U. E., Leyser, T. B., Isham, B., Gallop, P., Aso, T., Ejiri, M., Grydeland, T., Steen, Å., La-Hoz, C., Kaila, K., Jussila, J., and Holma, H.: The electron energy distribution during

HF pumping, a picture painted with all colors, Annales Geophysicae, 23, 1747–1754, doi: 10.5194/angeo-23-1747-2005, URL https://angeo.copernicus.org/articles/23/1747/2005/, 2005.

Gustavsson, B., Leyser, T., Kosch, M., Rietveld, M., Steen, Å., Brändström, B. U. E., and Aso, T.: Electron gyroharmonic effects in ionization and electron acceleration during high-frequency pumping in the ionosphere, Physical review letters, 97, 195 002, 2006.

Gustavsson, B., Kosch, M., Wong, A., Pedersen, T., Heinselman, C., Mutiso, C., Bristow, B., Hughes, J., and Wang, W.: First estimates of volume distribution of HF-pump enhanced emissions at 6300 and 5577 Å: a comparison between observations and theory, Annales Geophysicae, 26, 3999–4012, doi: 10.5194/angeo-26-3999-2008, URL https://angeo.copernicus.org/articles/26/3999/2008/, 2008a.

Gustavsson, B., Kosch, M. J., Senior, A., Kavanagh, A. J., Brändström, B., and Blixt, E.: Combined EISCAT radar and optical multispectral and tomographic observations of black aurora, Journal of Geophysical Research: Space Physics, 113, 2008b.

Han, B., Chu, F., Gao, X., and Yan, Y.: A multi-size kernels CNN with eye movement guided task-specific initialization for aurora image classification, in: CCF Chinese Conference on Computer Vision, pp. 533–544, Springer, 2017.

He, K., Zhang, X., Ren, S., and Sun, J.: Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778, 2016.

Hoffleit, D. and Jaschek, C.: The Bright star catalogue, New Haven, 1991.

Isham, B., Rietveld, M., Hagfors, T., La Hoz, C., Mishin, E., Kofman, W., Leyser, T., and Van Eyken, A.: Aspect angle dependence of HF enhanced incoherent backscatter, Advances in Space Research, 24, 1003–1006, 1999.

Jackman, C. M., Arridge, C. S., André, N., Bagenal, F., Birn, J., Freeman, M. P., Jia, X., Kidder, A., Milan, S. E., Radioti, A., et al.: Large-scale structure and dynamics of the magnetotails of Mercury, Earth, Jupiter and Saturn, Space Science Reviews, 182, 85–154, 2014.

Karlsson, T., Andersson, L., Gillies, D., Lynch, K., Marghitu, O., Partamies, N., Sivadas, N., and Wu, J.: Quiet, discrete auroral arcs—Observations, Space Science Reviews, 216, 1–50, 2020.

Kelley, M. C., Vickrey, J. F., Carlson, C., and Torbert, R.: On the origin and spatial extent of high-latitude F region irregularities, Journal of Geophysical Research: Space Physics, 87, 4469–4475, 1982.

Kendall, E., Marshall, R., Parris, R. T., Bhatt, A., Coster, A., Pedersen, T., Bernhardt, P., and Selcher, C.: Decameter structure in heater-induced airglow at the High frequency Active Auroral Research Program facility, Journal of Geophysical Research: Space Physics, 115, 2010.

Kero, J., Kastinen, D., Vierinen, J., Grydeland, T., Heinselman, C. J., Markkanen, J., and Tjulin, A.: EISCAT 3D: the next generation international atmosphere and geospace research radar, Tech. rep., European Space Agency, 2019.

Kingma, D. P. and Ba, J.: Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980, 2014.

Kozelov, B., Pilgaev, S., Borovkov, L., and Yurov, V.: Multi-scale auroral observations in Apatity: winter 2010–2011, Geoscientific Instrumentation, Methods and Data Systems, 1, 1–6, 2012.

Kozelov, B., Roldugin, A., Pilgaev, S., and Grigoriev, V.: Triangulation of auroral structures in Barentsburg, the first data of the season 2018-2019, Physics of Auroral Phenomena, 42, 152–155, 2019.

Krizhevsky, A., Sutskever, I., and Hinton, G. E.: Imagenet classification with deep convolutional neural networks, Advances in neural information processing systems, 25, 1097–1105, 2012.

Kudeki, E. and Milla, M. A.: Incoherent scatter spectral theories—Part I: A general framework and results for small magnetic aspect angles, IEEE Transactions on Geoscience and Remote Sensing, 49, 315–328, 2010.

Kvammen, A., Gustavsson, B., Sergienko, T., Brändström, U., Rietveld, M., Rexer, T., and Vierinen, J.: The 3-D Distribution of Artificial Aurora Induced by HF Radio Waves in the Ionosphere, Journal of Geophysical Research: Space Physics, 124, 2992–3006, doi: https://doi.org/10.1029/2018JA025988, 2019.

Kvammen, A., Wickstrøm, K., McKay, D., and Partamies, N.: Auroral image classification with deep neural networks, Journal of Geophysical Research: Space Physics, 125, e2020JA027 808, doi: https://doi.org/10.1029/2020JA027808, 2020.

Lagarias, J. C., Reeds, J. A., Wright, M. H., and Wright, P. E.: Convergence properties of the Nelder–Mead simplex method in low dimensions, SIAM Journal on optimization, 9, 112–147, 1998.

Laundal, K. M. and Richmond, A. D.: Magnetic coordinate systems, Space Science Reviews, 206, 27–59, 2017.

LeLevier, R.: Modification of the Ionosphere by Radiowaves, Tech. rep., RAND CORP SANTA MONICA CA, 1969.

Lessard, M.: A review of pulsating aurora, Auroral phenomenology and magnetospheric processes: Earth and other planets, 197, 55–68, 2012.

Leyser, T.: Stimulated electromagnetic emissions by high-frequency electromagnetic pumping of the ionospheric plasma, Space Science Reviews, 98, 223–328, 2001.

Li, W., Thorne, R., Bortnik, J., Nishimura, Y., and Angelopoulos, V.: Modulation of whistler mode chorus waves: 1. Role of compressional Pc4–5 pulsations, Journal of Geophysical Research: Space Physics, 116, 2011.

Libbrecht, K. G.: Physical dynamics of ice crystal growth, Annual Review of Materials Research, 47, 271–295, 2017.

Luhmann, J. and Solomon, S.: The Sun-Earth Connection, Encyclopedia of the Solar System„ pp. 213–226, 2007.

Lui, A.: A synthesis of magnetospheric substorm models, Journal of Geophysical Research: Space Physics, 96, 1849–1856, 1991.

Lui, A.: Reduction of the cross-tail current during near-Earth dipolarization with multisatellite observations, Journal of Geophysical Research: Space Physics, 116, 2011.

Lyons, L., Nishimura, Y., Kim, H.-J., Donovan, E., Angelopoulos, V., Sofko, G., Nicolls, M., Heinselman, C., Ruohoniemi, J., and Nishitani, N.: Possible connection of polar cap flows to pre-and post-substorm onset PBIs and streamers, Journal of Geophysical Research: Space Physics, 116, 2011.

MacDonald, E. A., Donovan, E., Nishimura, Y., Case, N. A., Gillies, D. M., Gallardo-Lacourt, B., Archer, W. E., Spanswick, E. L., Bourassa, N., Connors, M., et al.: New science in plain sight: Citizen scientists lead to the discovery of optical structure in the upper atmosphere, Science advances, 4, eaaq0030, 2018.

Mao, L., Xie, M., Huang, Y., and Zhang, Y.: Preceding vehicle detection using Histograms of Oriented Gradients, in: 2010 International Conference on Communications, Circuits and Systems (ICCCAS), pp. 354–358, IEEE, 2010.

Mauer, C. and Wueller, D.: Measuring the spectral response with a set of interference filters, in: Digital Photography V, vol. 7250, p. 72500S, International Society for Optics and Photonics, 2009.

McCrea, I., Aikio, A., Alfonsi, L., Belova, E., Buchert, S., Clilverd, M., Engler, N., Gustavsson, B., Heinselman, C., Kero, J., et al.: The science case for the EISCAT_3D radar, Progress in Earth and Planetary Science, 2, 1–63, 2015.

McKay, D. and Kvammen, A.: Auroral classification ergonomics and the implications for machine learning, Geoscientific Instrumentation, Methods and Data Systems, 9, 267–273, doi: 10.5194/gi-9-267-2020, 2020.

McKay, D., Paavilainen, T., Gustavsson, B., Kvammen, A., and Partamies, N.: Lumikot: Fast auroral transients during the growth phase of substorms, Geophysical Research Letters, 46, 7214–7221, doi: https://doi.org/10.1029/2019GL082985, 2019.

McPherron, R. L.: Growth phase of magnetospheric substorms, Journal of Geophysical Research, 75, 5592–5599, 1970.

Mende, S., Frey, H., Angelopoulos, V., and Nishimura, Y.: Substorm triggering by poleward boundary intensification and related equatorward propagation, Journal of Geophysical Research: Space Physics, 116, 2011.

Mendillo, M., Baumgardner, J., and Wroten, J.: SAR arcs we have seen: Evidence for variability in stable auroral red arcs, Journal of Geophysical Research: Space Physics, 121, 245–262, 2016.

Milan, S. E., Clausen, L. B. N., Coxon, J. C., Carter, J. A., Walach, M.-T., Laundal, K., Østgaard, N., Tenfjord, P., Reistad, J., Snekvik, K., et al.: Overview of solar wind–magnetosphere–ionosphere–atmosphere coupling and the generation of magnetospheric currents, Space Science Reviews, 206, 547–573, 2017.

Milikh, G., Gurevich, A., Zybin, K., and Secan, J.: Perturbations of GPS signals by the ionospheric irregularities generated due to HF-heating at triple of electron gyrofrequency, Geophysical research letters, 35, 2008.

Moss, K. and Stauning, P.: Sophus Peter Tromholt: an outstanding pioneer in auroral research, History of Geo-and Space Sciences, 3, 53–72, 2012.

Najmi, A., Eliasson, B., Shao, X., Milikh, G., and Papadopoulos, K.: Simulations of ionospheric turbulence produced by HF heating near the upper hybrid layer, Radio Science, 51, 704–717, 2016.

Najmi, A., Eliasson, B., Shao, X., Milikh, G., Sharma, A., and Papadopoulos, K.: Vlasov simulations of electron acceleration by radio frequency heating near the upper hybrid layer, Physics of Plasmas, 24, 102 904, 2017.

Nelder, J. A. and Mead, R.: A simplex method for function minimization, The computer journal, 7, 308–313, 1965.

Nishimura, Y., Lyons, L., Zou, S., Angelopoulos, V., and Mende, S.: Substorm triggering by new plasma intrusion: THEMIS all-sky imager observations, Journal of Geophysical Research: Space Physics, 115, 2010.

Niu, C., Zhang, J., Wang, Q., and Liang, J.: Weakly supervised semantic segmentation for joint key local structure localization and classification of aurora image, IEEE Transactions on Geoscience and Remote Sensing, 56, 7133–7146, 2018.

Partamies, N., Sangalli, L., Donovan, E., Connors, M., and Charrois, D.: Tomography-like Approach for Analysing Colour Auroral Images, Geophysica, 48, 81–90, 2012.

Partamies, N., Juusola, L., Whiter, D., and Kauristie, K.: Substorm evolution of auroral structures, Journal of Geophysical Research: Space Physics, 120, 5958–5972, 2015.

Partamies, N., Whiter, D., Kadokura, A., Kauristie, K., Nesse Tyssøy, H., Massetti, S., Stauning, P., and Raita, T.: Occurrence and average behavior of pulsating aurora, Journal of Geophysical Research: Space Physics, 122, 5606–5618, 2017.

Partamies, N., Bolmgren, K., Heino, E., Ivchenko, N., Borovsky, J. E., and Sundberg, H.: Patch size evolution during pulsating aurora, Journal of Geophysical Research: Space Physics, 124, 4725–4738, 2019.

Pedersen, T., Gustavsson, B., Mishin, E., MacKenzie, E., Carlson, H., Starks, M., and Mills, T.: Optical ring formation and ionization production in high-power HF heating experiments at HAARP, Geophysical research letters, 36, 2009.

Pedersen, T., Gustavsson, B., Mishin, E., Kendall, E., Mills, T., Carlson, H., and Snyder, A.: Creation of artificial ionospheric layers using high-power HF waves, Geophysical Research Letters, 37, 2010.

Perreault, P. and Akasofu, S.: A study of geomagnetic storms, Geophysical Journal International, 54, 547–573, 1978.

Phan, T., Kistler, L., Klecker, B., Haerendel, G., Paschmann, G., Sonnerup, B. Ö., Baumjohann, W., Bavassano-Cattaneo, M., Carlson, C., DiLellis, A., et al.: Extended magnetic reconnection at the Earth's magnetopause from detection of bi-directional jets, Nature, 404, 848–850, 2000.

Picone, J., Hedin, A., Drob, D. P., and Aikin, A.: NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues, Journal of Geophysical Research: Space Physics, 107, SIA–15, 2002.

Radon, J.: On the determination of functions from their integral values along certain manifolds, IEEE transactions on medical imaging, 5, 170–176, 1986.

Rao, J., Partamies, N., Amariutei, O., Syrjäsuo, M., and van de Sande, K. E.: Automatic auroral detection in color all-sky camera images, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7, 4717–4725, 2014.

Reidy, J. A., Fear, R., Whiter, D., Lanchester, B., Kavanagh, A. J., Milan, S., Carter, J., Paxton, L., and Zhang, Y.: Interhemispheric survey of polar cap aurora, Journal of Geophysical Research: Space Physics, 123, 7283–7306, 2018.

Rexer, T.: Current systems associated with Non-Conjugate Aurora, Master's thesis, The University of Bergen, 2015.

Rexer, T., Gustavsson, B., Leyser, T., Rietveld, M., Yeoman, T., and Grydeland, T.: First Observations of Recurring HF-Enhanced Topside Ion Line Spectra Near the Fourth Gyroharmonic, Journal of Geophysical Research: Space Physics, 123, 8649–8663, 2018.

Rexer, T., Leyser, T., Gustavsson, B., and Rietveld, M.: Conditions for topside ion line enhancements, Journal of Geophysical Research: Space Physics, p. e2021JA029379, 2021.

Rhodes, B.: Skyfield v1.10, doi: http://rhodesmill.org/skyfield/, URL `http://rhodesmill.org/skyfield/`, 2019.

Rietveld, M., Isham, B., Kohl, H., La Hoz, C., and Hagfors, T.: Measurements of HF-enhanced plasma and ion lines at EISCAT with high-altitude resolution, Journal of Geophysical Research: Space Physics, 105, 7429–7439, 2000.

Rietveld, M. T., Senior, A., Markkanen, J., and Westman, A.: New capabilities of the upgraded EISCAT high-power HF facility, Radio Science, 51, 1533–1546, 2016.

Rostoker, G., Akasofu, S., Baumjohann, W., Kamide, Y., and McPherron, R.: The roles of direct input of energy from the solar wind and unloading of stored magnetotail energy in driving magnetospheric substorms, Space science reviews, 46, 93–111, 1988.

Rydesäter, P. and Gustavsson, B.: Investigation of smooth basis functions and an approximated projection algorithm for faster tomography, International journal of imaging systems and technology, 11, 347–354, 2000.

Rydesäter, P., Gustavsson, B., Brändström, U., and Steen, Å.: Lossy compression of scientific images of aurora, in: Proc. of 28th Annual European Meeting on Atmospheric Studies by Optical Methods, Sodankylä Geophysical Observatory, 2001.

Sangalli, L., Gustavsson, B., Partamies, N., and Kauristie, K.: Estimating the peak auroral emission altitude from all-sky images, Opt. Pura Apl., 44, 593–598, 2011a.

Sangalli, L., Partamies, N., Syrjäsuo, M., Enell, C.-F., Kauristie, K., and Mäkinen, S.: Performance study of the new EMCCD-based all-sky cameras for auroral imaging, International Journal of Remote Sensing, 32, 2987–3003, 2011b.

Senior, A., Rietveld, M., Yeoman, T., and Kosch, M.: The dependence of F-region electron heating on HF radio pump power: Measurements at EISCAT Tromsø, Journal of Geophysical Research: Space Physics, 117, 2012.

Senior, A., Rietveld, M., Häggström, I., and Kosch, M.: Radio-induced incoherent scatter ion line enhancements with wide altitude extents in the high-latitude ionosphere, Geophysical research letters, 40, 1669–1674, 2013.

Sergeev, V., Angelopoulos, V., Kubyshkina, M., Donovan, E., Zhou, X.-Z., Runov, A., Singer, H., McFadden, J., and Nakamura, R.: Substorm growth and expansion onset as observed with ideal ground-spacecraft THEMIS coverage, Journal of Geophysical Research: Space Physics, 116, 2011.

Shaffer, K., Haughton, V., and Wilson, C. R.: High resolution computed tomography of the temporal bone., Radiology, 134, 409–414, 1980.

Shiokawa, K., Otsuka, Y., and Connors, M.: Statistical study of auroral/resonant-scattering 427.8-nm emission observed at subauroral latitudes over 14 years, Journal of Geophysical Research: Space Physics, 124, 9293–9301, 2019.

Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., et al.: Mastering chess and shogi by self-play with a general reinforcement learning algorithm, arXiv preprint arXiv:1712.01815, 2017.

Simon Wedlund, C., Lamy, H., Gustavsson, B., Sergienko, T., and Brändström, U.: Estimating energy spectra of electron precipitation above auroral arcs from ground-based observations with radar and optics, Journal of Geophysical Research: Space Physics, 118, 3672–3691, 2013.

Stamm, J., Vierinen, J., Urco, J. M., Gustavsson, B., and Chau, J. L.: Radar imaging with EISCAT 3D, Annales Geophysicae, 39, 119–134, doi: 10.5194/angeo-39-119-2021, URL https://angeo.copernicus.org/articles/39/119/2021/, 2021.

Størmer, C.: Altitudes of aurorae, Nature, 97, 5–5, 1916.

Størmer, C.: The distribution in space of the sunlit aurora rays, Nature, 123, 82–83, 1929.

Streltsov, A., Berthelier, J.-J., Chernyshov, A., Frolov, V., Honary, F., Kosch, M., McCoy, R., Mishin, E., and Rietveld, M.: Past, present and future of active radio frequency experiments in space, Space Science Reviews, 214, 1–122, 2018.

Syrjäsuo, M. and Donovan, E.: Diurnal auroral occurrence statistics obtained via machine vision, in: Annales Geophysicae, vol. 22, pp. 1103–1113, Copernicus GmbH, 2004.

Syrjasuo, M. and Partamies, N.: Numeric image features for detection of aurora, IEEE Geoscience and Remote Sensing Letters, 9, 176–179, 2011.

Tanaka, Y.-M., Aso, T., Gustavsson, B., Tanabe, K., Ogawa, Y., Kadokura, A., Miyaoka, H., Sergienko, T., Brändström, U., and Sandahl, I.: Feasibility study on generalized-aurora computed tomography, in: Annales Geophysicae, vol. 29, pp. 551–562, Copernicus GmbH, 2011.

Tenfjord, P. and Østgaard, N.: Energy transfer and flow in the solar wind-magnetosphere-ionosphere system: A new coupling function, Journal of Geophysical Research: Space Physics, 118, 5659–5672, 2013.

Tenfjord, P., Østgaard, N., Snekvik, K., Laundal, K. M., Reistad, J. P., Haaland, S., and Milan, S.: How the IMF By induces a By component in the closed magnetosphere and how it leads to asymmetric currents and convection patterns in the two hemispheres, Journal of Geophysical Research: Space Physics, 120, 9368–9384, 2015.

Theodoridis, S. and Koutroumbas, K.: Chapter 1 - Introduction, in: Pattern Recognition (Fourth Edition), edited by Theodoridis, S. and Koutroumbas, K., pp. 1 – 12, Academic Press, Boston, fourth edition edn., doi: https://doi.org/10.1016/B978-1-59749-272-0.50004-9, URL http://www.sciencedirect.com/science/article/pii/B9781597492720500049, 2009a.

Theodoridis, S. and Koutroumbas, K.: Chapter 2 - Classifiers Based on Bayes Decision Theory, in: Pattern Recognition (Fourth Edition), edited by Theodoridis, S. and Koutroumbas, K., pp. 13 – 89, Academic Press, Boston, fourth edition edn., doi: https://doi.org/10.1016/B978-1-59749-272-0.50004-9, URL http://www.sciencedirect.com/science/article/pii/B9781597492720500049, 2009b.

Theodoridis, S. and Koutroumbas, K.: Chapter 3 - Linear Classifiers, in: Pattern Recognition (Fourth Edition), edited by Theodoridis, S. and Koutroumbas, K., pp. 91 – 150, Academic Press, Boston, fourth edition edn., doi: https://

doi.org/10.1016/B978-1-59749-272-0.50004-9, URL http://www.sciencedirect.com/science/article/pii/B9781597492720500049, 2009c.

Theodoridis, S. and Koutroumbas, K.: Chapter 4 - Nonlinear Classifiers, in: Pattern Recognition (Fourth Edition), edited by Theodoridis, S. and Koutroumbas, K., pp. 151 – 260, Academic Press, Boston, fourth edition edn., doi: https://doi.org/10.1016/B978-1-59749-272-0.50004-9, URL http://www.sciencedirect.com/science/article/pii/B9781597492720500049, 2009d.

Utlaut, W.: An ionospheric modification experiment using very high power, high frequency transmission, Journal of Geophysical Research, 75, 6402–6405, 1970.

Van der Maaten, L. and Hinton, G.: Visualizing data using t-SNE., Journal of machine learning research, 9, 2008.

Vegard, L. and Krogness, O. A.: The Position in Space of the Aurora Polaris: From Observations Made at the Haldde-observatory, 1913-14, vol. 1, AW Brøggers boktrykkeri, 1920.

Wang, Q., Liang, J., Hu, Z.-J., Hu, H.-H., Zhao, H., Hu, H.-Q., Gao, X., and Yang, H.: Spatial texture based automatic classification of dayside aurora in all-sky images, Journal of Atmospheric and Solar-Terrestrial Physics, 72, 498–508, 2010.

Whiter, D., Lanchester, B., Gustavsson, B., Ivchenko, N., and Dahlgren, H.: Using multispectral optical observations to identify the acceleration mechanism responsible for flickering aurora, Journal of Geophysical Research: Space Physics, 115, 2010.

Whiter, D., Gustavsson, B., Partamies, N., and Sangalli, L.: A new automatic method for estimating the peak auroral emission height from all-sky camera images, Geoscientific Instrumentation, Methods and Data Systems, 2, 131–144, 2013.

Wu, M. and Chen, L.: Image recognition based on deep learning, in: 2015 Chinese Automation Congress (CAC), pp. 542–546, IEEE, 2015.

Yang, Q. and Zhou, P.: Representation and Classification of Auroral Images Based on Convolutional Neural Networks, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 523–534, 2020.

Yen, C.-K., Yano, Y., Budinger, T., Friedland, R., Derenzo, S., Huesman, R., and O'Brien, H.: Brain tumor evaluation using Rb-82 and positron emission tomography., Journal of Nuclear Medicine: Official Publication, Society of Nuclear Medicine, 23, 532–537, 1982.

Zhang, Q.-H., Lockwood, M., Foster, J., Zhang, S.-R., Zhang, B.-C., McCrea, I., Moen, J., Lester, M., and Ruohoniemi, J. M.: Direct observations of the full Dungey convection cycle in the polar ionosphere for southward interplanetary magnetic field conditions, Journal of Geophysical Research: Space Physics, 120, 4519–4530, 2015.

Zou, S., Moldwin, M., Lyons, L., Nishimura, Y., Hirahara, M., Sakanoi, T., Asamura, K., Nicolls, M., Miyashita, Y., Mende, S., et al.: Identification of substorm onset location and preonset sequence using Reimei, THEMIS GBO, PFISR, and Geotail, Journal of Geophysical Research: Space Physics, 115, 2010.

PAPER I:
Auroral Image Classification With Deep Neural Networks

Kvammen, A., Wickstrøm, K., McKay, D., and Partamies, N.: Auroral image classification with deep neural networks, Journal of Geophysical Research: Space Physics, 125, e2020JA027 808, doi: https://doi.org/10.1029/2020JA027808, 2020

JGR Space Physics

RESEARCH ARTICLE
10.1029/2020JA027808

Auroral Image Classification With Deep Neural Networks

Andreas Kvammen[1], Kristoffer Wickstrøm[1], Derek McKay[2,3], and Noora Partamies[4,5]

[1]Department of Physics and Technology, UiT-The Arctic University of Norway, Tromsø, Norway, [2]NORCE Norwegian Research Centre AS, Tromsø, Norway, [3]Finnish Centre for Astronomy with ESO, FINCA, University of Turku, Turku, Finland, [4]Department of Arctic Geophysics, The University Centre in Svalbard, Longyearbyen, Norway, [5]Birkeland Centre for Space Science, Bergen, Norway

Key Points:
- Auroral images were labeled into seven classes and used to train and test machine learning classifiers
- The deep neural networks outperformed the k-nearest neighbor and support vector machine classifiers
- The ResNet-50 deep neural network had the highest performance and achieved 92% precision

Correspondence to:
A. Kvammen,
andreas.kvammen@uit.no

Citation:
Kvammen, A., Wickstrøm, K., McKay, D., & Partamies, N. (2020). Auroral image classification with deep neural networks. *Journal of Geophysical Research: Space Physics*, *125*, e2020JA027808. https://doi.org/10.1029/2020JA027808

Received 16 JAN 2020
Accepted 23 SEP 2020
Accepted article online 5 OCT 2020

Abstract Results from a study of automatic aurora classification using machine learning techniques are presented. The aurora is the manifestation of physical phenomena in the ionosphere-magnetosphere environment. Automatic classification of *millions* of auroral images from the Arctic and Antarctic is therefore an attractive tool for developing auroral statistics and for supporting scientists to study auroral images in an objective, organized, and repeatable manner. Although previous studies have presented tools for detecting aurora, there has been a lack of tools for classifying aurora into subclasses with a high precision (>90%). This work considers seven auroral subclasses: *breakup*, *colored*, *arcs*, *discrete*, *patchy*, *edge*, and *faint*. Six different deep neural network architectures have been tested along with the well-known classification algorithms: k-nearest neighbor (KNN) and a support vector machine (SVM). A set of clean nighttime color auroral images, without clearly ambiguous auroral forms, moonlight, twilight, clouds, and so forth, were used for training and testing the classifiers. The deep neural networks generally outperformed the KNN and SVM methods, and the ResNet-50 architecture achieved the highest performance with an average classification precision of 92%.

1. Introduction

Spectacular auroral displays can be seen on the night sky at high latitudes if the solar wind, magnetospheric, and ionospheric conditions are opportune. The auroral excitation processes are activated by energetic electrons and protons from the magnetosphere. The charged particles follow the magnetic field from the plasma sheet down to their magnetic footprint in the ionosphere where the energy is dissipated by ionization, excitation, and heating of thermospheric particles. The auroral displays, with typical emission intensity peaks at altitudes between 100 and 130 km, are therefore an indicator of dynamical processes that occur much further out into the Earth's magnetotail. Understanding how different ionospheric and magnetospheric conditions are manifested in the shape, color, intensity, and time evolution of the aurora is not well understood, even over 100 years after the first big auroral imaging campaigns in Bossekop, Norway (Störmer, 1913).

Since the first auroral imaging campaigns, millions of auroral images have been taken in the Arctic and Antarctic regions, and auroral scientists now have access to more data than what is possible to search and analyze by visual inspections. During recent decades, it has been demonstrated that machine learning methods are a valuable and highly applicable tool for automatically classifying large image data sets, for instance, by letter, brain tumor, and facial recognition. Machine learning techniques are, however, not widely used within the auroral research community. Automatic classification of millions of images captured every year will make it easier for scientists to study the images that are of interest in an organized, objective, and repeatable manner. This will further make statistical studies easier to conduct. For example, probability distributions of different auroral structures will facilitate studies of the temporal evolution of the aurora under different geomagnetic condition. In addition, statistical studies of the ionosphere-magnetosphere environment can be conducted by investigating the occurrence of auroral classes which are related to physical processes in the magnetosphere, such as the occurrence and evolution of *patchy aurora* that can be statistically studied in relation to the pitch angle scattering due to wave-particle interactions in the ionosphere-magnetosphere system.

Machine classification of aurora is a difficult task due to the transparent nature of the emission and thus the soft boundaries of the observed forms. Manual classification of auroral images is also a challenging task

since no two auroral events are alike and the aurora is a very dynamic phenomena which changes rapidly and manifests differently depending on geomagnetic conditions. An additional complication is that there is no clear consensus about how many nighttime auroral classes exist and what they are. Nevertheless, some automatic classification and feature extraction techniques have been developed and tested over the last two decades, for instance, using k-nearest neighbor (KNN) (Syrjäsuo & Donovan, 2002, 2004), Fourier descriptors (Syrjäsuo et al., 2007), support vector machines (SVMs) (Rao et al., 2014), and polynomial fitting (Partamies et al., 2015). However, none of the developed tools have achieved broader usage, possibly due to the techniques not being general enough, low classification accuracy, or few and/or unwanted auroral labels. It should also be noted that machine learning techniques have been extensively used for dayside aurora classification in the cusp region using hidden Marcov models (Yang et al., 2012) and to find key features detected by the cycle-consistent generative adversarial network (CycleGAN) (Yang et al., 2019). However, the developed dayside tools cannot be used on nightside aurora images without additional training due to the differences in auroral morphology. In a recent study, Clausen and Nickisch (2018) presented automatic nighttime auroral classification results by employing a pretrained deep neural network (DNN) and 5,824 labeled images. Clausen and Nickisch (2018) achieved 81.7 ± 0.1% classification accuracy, distributed into six labels: *arc*, *discrete*, *diffuse*, *cloudy*, *moon*, and *clear/no aurora*. Furthermore, when clustering together the auroral classes (*arc*, *discrete*, and *diffuse*) and the nonauroral classes (*cloudy*, *moon*, and *clear/no aurora*), an accuracy of 95.60 ± 0.03% was achieved for detecting *aurora* versus *no aurora* conditions.

Following the previous work, the main purpose of this study was to classify color (RGB) nighttime auroral images (manually preselected to contain clear skies) with higher precision and into more labels than those in the previous studies. Classification error rates of <10% are considered sufficient for operational purposes (Syrjasuo & Partamies, 2011); this study therefore aimed at an average classification precision of >90%. An additional objective was to define auroral labels which represent an exclusive production mechanism or characterize a physical property of the aurora when possible. Finally, several neural networks and machine learning techniques were tested and compared in this study. The pretrained DNN considered in Clausen and Nickisch (2018) is among the evaluated classifiers. In total, 14,030 auroral images were labeled. Out of these, 3,854 auroral images did not contain clearly ambiguous auroral forms and were therefore selected for network training and testing. The methodology is described in section 2. A comparison of the performance of the different algorithms and an analysis of misidentified images is presented in section 3, with a more general discussion in section 4.

2. Methodology

A description of the data processing, the auroral labels, and the applied machine learning techniques are presented in this section. A detailed description of auroral classification ergonomics and potential biases in the data set is presented in McKay and Kvammen (2020).

2.1. Data Acquisition and Preprocessing

The images used in this study were acquired by the all-sky camera located near Kiruna, Sweden, at 425 m above mean sea level with geographic (latitude and longitude) location: 67.84°N, 20.42°E and CGM location: 64.69°, 102.64°, and operated by the Swedish Institute for Space Physics. The camera is a Nikon D700 equipped with a Nikon Nikkor 8 mm 1:2.8 lens giving almost 180° field of view. The color sensitivity of the detector, from Mauer (2009), is depicted in Figure 1 along with the characteristic 427.8, 557.7, and 630.0 nm auroral lines. The camera was installed in 2009 with the available data set used in this project extending from 2010 to 2019. The exposure time is 6 s, with images taken automatically on each minute. To ease data transfer rates and processing, JPEG images (720 × 479 pixels) were used, rather than the full-resolution images.

Images containing clouds and moon where excluded, since earlier studies by Rao et al. (2014) and Clausen and Nickisch (2018) showed high accuracy for labeling clouds and moon while being less successful in labeling auroral subclasses. A preprocessing stage was carried out where keograms were manually inspected and areas of potential auroral emission were selected, thus rejecting the bulk of overcast sky conditions. Celestial positions were calculated using the Skyfield software (Rhodes, 2019) and images where the Sun was at an apparent elevation greater than −15°, or where the moon was above the horizon, were automatically rejected.

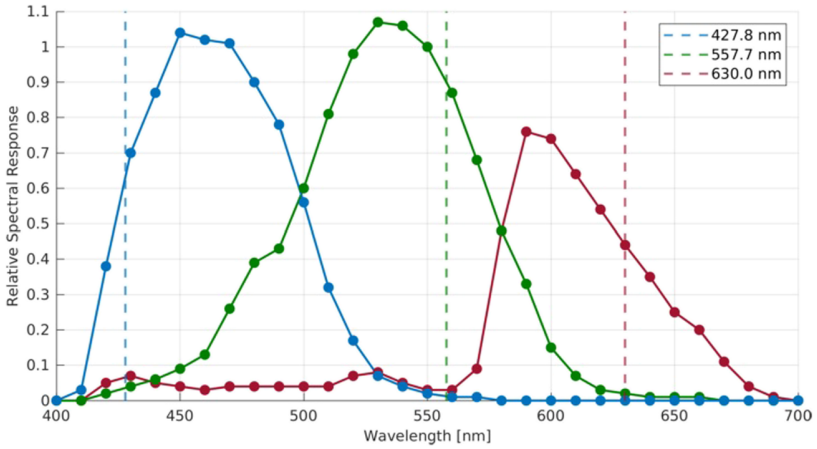


Figure 1. The figure shows the color sensitivity of the blue, green, and red channels of the detector. It is clear that the camera sensitivity is not uniform with respect to the wavelength. The relative response is approximately 0.65 for the blue auroral line at 427.8 nm, 0.90 for the green line at 557.7 nm, and 0.45 for the red line at 630.0 nm.

To prepare the images for labeling and network training, a four-step image processing procedure (see Figure 2) was performed on each image:

1. Rotate the images clockwise 90° to direct the geomagnetic pole toward the top of the image and flip the image along the east-west axis.
2. Filter each image with a 3 × 3 median filter to avoid bias effects from the location of stars, remove bad pixels, and reduce noise.
3. Bin the pixels by using a 2 × 2 averaging window to reduce the size of the images and thus speed up the training process.
4. Crop the images to the central 128 × 128 pixels of the binned image, corresponding to the size of the red frame in the left panel of Figure 2. Apart from further speeding-up the training process, there are three reasons for cropping the images. First, pixels that contain considerably distorted features from the fish-eye lens projection and atmospheric conditions are removed. Second, aurorae look similar toward the horizon and classification of aurorae at small elevation angles is not useful. Lastly, selecting a smaller field of view reduced the number of frames which included several aurora classes; this increased the accuracy of the network and made it easier to both label the images and to choose representative aurora labels.

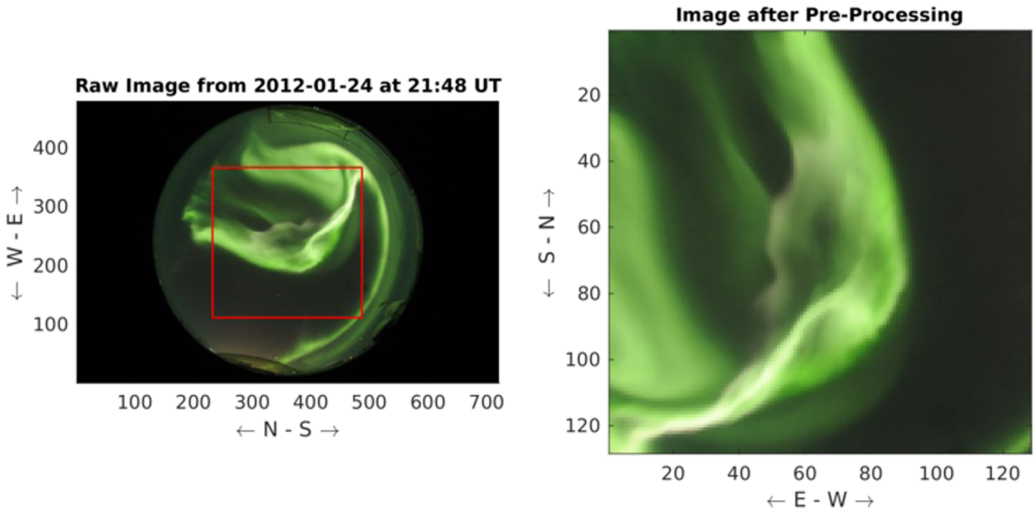


Figure 2. The image to the left depicts a raw (720 × 479 pixels) auroral image. The image to the right shows the same image after the four-step processing procedure. The processed image is rotated, filtered, binned, and cropped. The resolution of the processed image is 128 × 128 pixels. The red frame (left panel) indicates the size of the cropped area.

2.2. Aurora Classification

The processed images were labeled manually. The set of labels used in this study is described below and illustrated in Figure 3. The visual class definition, used for labeling the auroral images, is written in italic, followed by a short physical description and/or explanation. The proposed set of auroral labels was easy to identify under different geomagnetic conditions and applicable to most of the auroral images. In addition, most labels represent an exclusive production mechanisms or characterize a physical property of the aurora. Finally, the labels represent auroral forms which are recognizable at different viewing/elevation angles. Thus, the distribution of the classes is not dependent on the location of the auroral form in the image. This makes statistical interpretations of the results easier. The *edge aurora* class is an exception to the location independence by its definition.

Auroral breakup. *Auroral breakup is characterized by bright and large auroral forms which cover most of the frame.* Sample images labeled as *auroral breakup* are shown in Row 1 in Figure 3. *Auroral breakup* is an expansion of bright aurora that includes a variety of different large-scale features (e.g., Nishimura et al., 2010). It is characteristic to the substorm expansion phase, which is caused by dynamic processes in the magnetotail (e.g., Xing et al., 2010), leading to enhanced particle precipitation to the ionosphere.

Colored aurora. *An image is classified as colored aurora if the aurora, of any shape and form, is clearly not monochromatic green but has a prominent red, blue or purple emission.* Images classified as *colored aurora* are presented in Row 2 in Figure 3. The typical precipitation energy and the atmospheric composition in the altitude range of 90–130 km result in green (557.7 nm) being the dominant color in most aurorae. Distinct colored aurora occurs when the electron energy distribution has a pronounced low- or high-energy tail, changing the electron penetration depth into the Earth's ionosphere. Blue and purple auroral displays are usually seen when electrons penetrate deeper into the Earth's ionosphere, as compared to electrons causing the green aurora. Red aurora, however, is produced at higher altitudes and characterizes lower energy electron precipitation.

Auroral arcs. *Aurorae with the emission distributed in a single or multiple east-west aligned structure/structures spanning across the image are labeled as auroral arcs.* Row 3 in Figure 3 illustrates samples of *arcs*. *Auroral arcs* run parallel to the auroral oval and the magnetic latitudes (Karlsson et al., 2020). They result from quasi-static particle acceleration in a region close to the ionosphere (Lysak et al., 2020), and they magnetically map to the plasma sheet. *Arcs* are typically considered as quiet time auroral forms but exist at all magnetic activity levels as a basic element of the auroral displays.

Discrete-irregular. *Auroral emission appears in broken arcs, north-south aligned arcs, vortical structures, or a combination of several discrete shapes. Discrete-irregular auroral forms are not as bright and not as large as auroral breakup forms.* Sample images labeled as *Discrete-irregular* are presented in Row 4 in Figure 3. The *Discrete-irregular* class contains a mixture of different physical generation processes which are not easy to untangle.

Patchy aurora. *Patchy aurora is characterized by diffuse aurora consisting of irregular shapes which cover large portions of the image. The intensity of the auroral emission in this class is weak. Patchy aurora* images are shown in Row 5 in Figure 3. *Patchy aurora* mainly consists of different pulsating aurora structures (Nishimura et al., 2020). Diffuse patches are caused by pitch angle scattering of energetic electrons to the ionosphere. Different plasma waves play a key role in the scattering processes.

Edge aurora. *Images with auroral emission occurring only at the edge of the image are labeled as edge aurora.* Sample images are seen in Row 6 in Figure 3. *Edge aurora* can be any of the auroral classes above but information about the class is limited by an insufficient number of bright pixels and uncertainty of the form of the aurora outside the image frame. Thus, not attempting to classify these images as, for instance, *breakup*, *arcs*, or *discrete*, makes the classifier more robust. The *edge aurora* label was included as an additional subclass since the location of the aurora in a set of images is often valuable information for determining if aurora is drifting northward, southward, eastward, or westward.

Faint clear. *Images which are dark without clearly visible aurora are labeled as faint clear.* Images labeled as *faint clear* are presented in the bottom row of Figure 3. *Faint-clear* images indicate very weak electron precipitation and a quiet ionosphere-magnetosphere environment along the field lines overhead.

Images where a mixture of classes existed were labeled by the most dominant feature with priority given from top (highest priority) to bottom (lowest priority), in the auroral class description. Furthermore, two additional labels were used for classifying the entire data set: *unknown-complicated* and *rejected*. Images classified as *unknown-complicated* or *rejected* were not used for training the networks and are therefore not

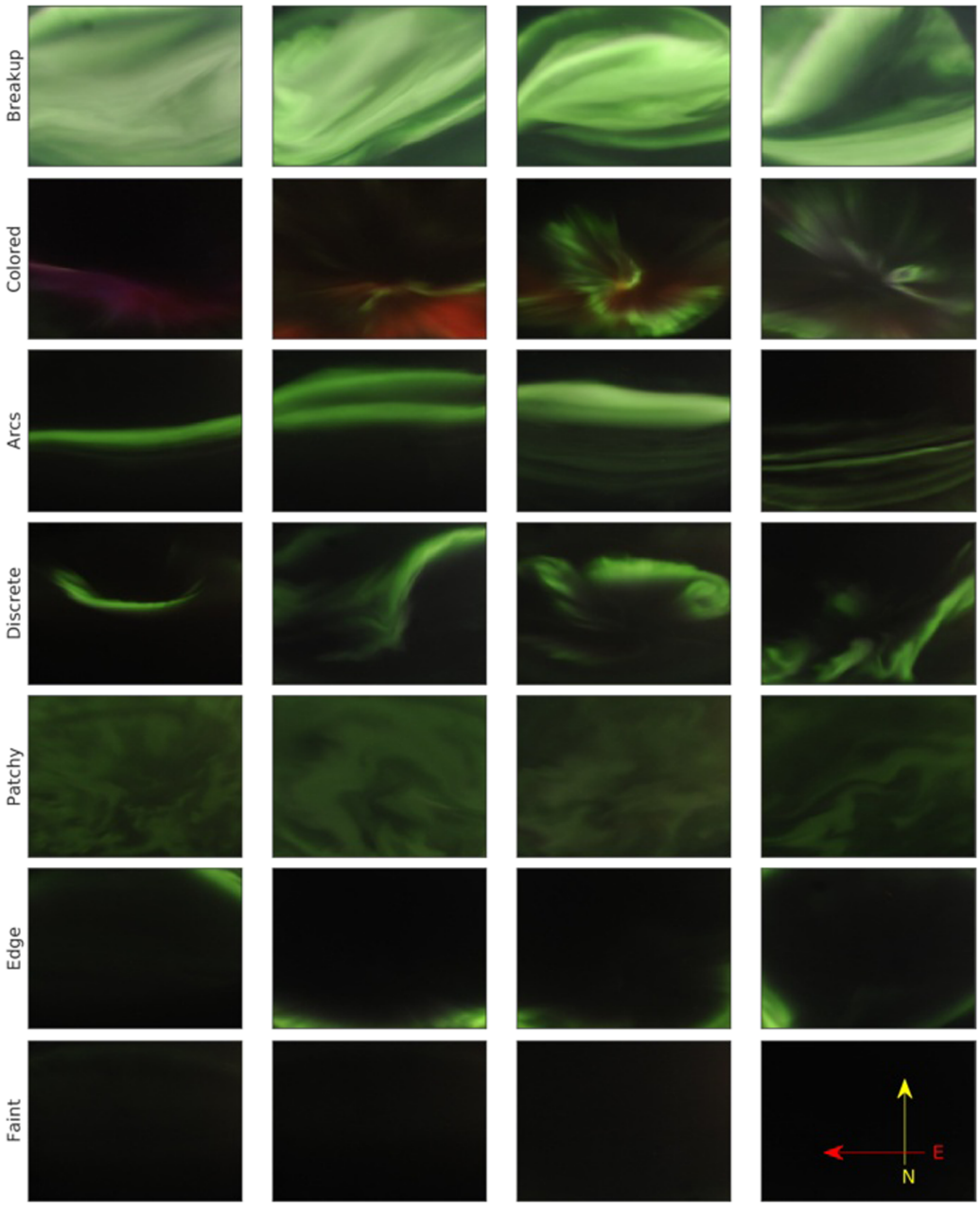


Figure 3. The figure depicts four sample images from each aurora class used for training and testing of the classifiers. The images are processed according to the four-step processing procedure described in section 2.1. The direction with respect to the magnetic pole is indicated by the arrows in the bottom right.

presented in Figure 3. The *unknown-complicated* class was used to exclude images with ambiguous auroral forms which fitted none or several of the auroral labels without a clearly dominant feature. The *rejected* class was used to exclude images with unwanted features, such as clouds and light pollution, which were not detected in the initial preprocessing stage; see section 2.1.

Two of the authors classified the entire data set, one of the authors labeled the images in consecutive order and the other in random order. Only the images with agreeing labels were used for training the network. This was done in order to reduce labeling bias and noise. Thus, a clean training and testing data set was produced by only using the images with agreeing auroral labels and excluding the images with ambiguous auroral forms (labeled as *unknown-complicated*), unwanted features (labeled as *rejected*), and disagreeing

Table 1
The Table Summarizes the Image Data Sets Used in This Study

Year	Number of images	Label	Training	Test	Number of images
2010	479	*Breakup*	113	33	146 (1.04%)
2011	1,506	*Colored*	97	36	133 (0.95%)
2012	2,498	*Arcs*	636	195	831 (5.92%)
2013	0	*Discrete*	150	44	194 (1.38%)
2014	1,684	*Patchy*	1,691	459	2,150 (15.32%)
2015	2,835	*Edge*	167	50	217 (1.55%)
2016	2,037	*Faint*	146	29	175 (1.25%)
2017	1,733	*Unknown-complicated*	—	—	2,630 (18.75%)
2018	996	*Rejected*	—	—	1,078 (7.68%)
2019	262	Disagreement	—	—	6,476 (46.16%)
Total	14,030	Total	3,000	846	14,030 (100.00%)

Note. To the left, the number of auroral images each year that satisfies the initial conditions described in section 2.1 (no clouds, moon below the horizon, and the Sun below −15°). To the right, the number of training, testing, and total images per class. Note that images labeled as *Unknown-complicated*, *Rejected*, or with disagreeing labels were not used for training and testing the classifiers.

labels. The clean data set contained 3,846 (27%) of the 14,030 images in the initial data set, spanning over the Years 2010 to 2019. Finally, the clean data were split into a training and test data set. The training set contains 3,000 images and was used for training the classifiers, while the testing set contains 846 images and was used as an independent test set to evaluate the performance of the classifier. The number of images each year and in each label is presented in Table 1 along with number of images in the training and test sets. Note that although 46% of the images had disagreeing labels, the most disagreement was whether or not an image was suitable for training/testing. The experts agreed on 95% of the labels on images that both experts considered suitable for training/testing (i.e., both experts labeled as *breakup*, *colored*, *arcs*, *discrete*, *patchy*, *edge*, or *faint*), as illustrated in Figure 3.

The motivation for constructing a clean data set is to avoid "confusing" the network by using ambiguous images during training and testing. Note that labels defined in auroral observers classification guides, such as the labels proposed in the International Auroral Atlas (IAA, C.D.W., 1964), were not applicable for training and testing of our data set. The studied data set is not large enough to reliably train and test networks with numerous (>10) and very specific auroral subclasses while still maintaining an acceptable number (∼100) of training images. In addition, labels defined by the temporal characteristics of the aurora cannot be implemented in our study since all considered classifiers are time invariant.

The labels containing less than ∼100 training images is at the lower limit for network training. Note that an insufficient number of training images might cause nonconvergence during training of the classifier. In the early stages of the study, some readjustment of the classes was necessary to ensure there were enough representative samples for the anticipated classes. There were also some false starts, where preselection (i.e., removal of images based on other criteria, such as the presence of the moon) caused some classes to be too few in number.

2.3. DNNs

In machine learning, an object is commonly described by a set of values referred to as features. For images, these features are the raw pixel values of the image. Machine learning algorithms seek to find patterns in the given set of features such that some task can be solved in an optimal way. Determining what features to present the algorithm is crucial in designing a good algorithm. Historically, the main approach has been to design methods that extract features from raw data that are assumed to provide good discriminative power (Guyon & Elisseeff, 2003). However, the design of such features can be challenging and might require a significant amount of domain knowledge. In contrast, DNNs automatically learn which features in the training data are important to solve the desired task (LeCun et al., 2015). The DNNs achieve this automatic extraction by transforming the data through a cascade of nonlinear transformations, which results in a representation that is suited for the problem at hand. Each transformation is commonly referred to as a hidden layer, which

contains parameters that must be optimized. These parameters are optimized using gradient descent, where the gradients are obtained through the back propagation algorithm (Theodoridis & Koutroumbas, 2009b).

Currently, there exists a vast amount of different DNN architectures. This study focuses on some widely used and well-known architectures, as the goal was to demonstrate that DNNs can be effective in auroral classification. The following DNNs were used for classifying the aurora images:

- VGG: A widely used family of convolutional neural networks (CNNs) that have demonstrated a high performance on a number of tasks (Simonyan & Zisserman, 2014). Dropout (Srivastava et al., 2014) was included to regularize the model. Later versions also included batch normalization (Ioffe & Szegedy, 2015). Different versions of the VGG can be created by adding more layers to the network. In this work, we evaluated two versions of the VGG. First, the 16-layer version titled the VGG-16, then the 19-layer version named VGG-19.
- AlexNet: Often considered the breakthrough of deep learning, AlexNet is a CNN consisting of five convolutional layers and three fully connected layers (Krizhevsky et al., 2012). Each layers is followed by a Rectified Linear Unit (ReLU) activation function. Similarly with VGG, Dropout was included to regularize the model.
- ResNet: Residual networks (ResNets) (He et al., 2016) include skip connections between the hidden layers of the networks. Such skip connections ease the flow of gradients in the network (Balduzzi et al., 2017) such that more hidden layers can be included in the network, a process that has shown to increase performance. As with the VGG, adding more layers results in different version of the ResNet. In this study, we evaluated two versions of the ResNet. First, the 18-layer version titled ResNet-18, then the 50-layer version named ResNet-50.
- Clausen and Nickisch (2018) approach: Additionally, we evaluated the performance of the approach used by Clausen and Nickisch (2018). Clausen and Nickisch (2018) used a pretrained deep learning model, an inception model (Szegedy et al., 2017), as a feature extractor that was combined with a linear classifier.

All models were trained using a cross-entropy loss and the Adam optimizer (Kingma & Ba, 2015). The models were developed using the deep learning framework Pytorch (Paszke et al., 2017) on a Tesla K80 GPU.

3. Results

This section presents the classification scores for the KNN method, the SVM method, and six different deep learning-based models on the aurora data set. Further, the confusion matrix and the presoftmax representation obtained by the highest performing model is visualized. Examples of images that were incorrectly classified by the highest performing model are shown and discussed. Lastly, a comparison between the class wise accuracy of a deep learning-based approach to a traditional machine learning approach is presented.

This study evaluates all classifiers by calculating the precision, recall, and F1 score. Precision is defined as follows:

$$\text{Precision} = \frac{\text{true positive}}{\text{true positive} + \text{false positive}}, \tag{1}$$

and measures the classifiers ability for not labeling positive samples as negative. A true positive is when the model correctly predicts the positive class. A false positive is when the model incorrectly predicts the positive class. Recall is defined as follows:

$$\text{Recall} = \frac{\text{true positive}}{\text{true positive} + \text{false negative}}, \tag{2}$$

and measures the classifiers ability to find positive samples. A false negative is when the model incorrectly predicts the negative class. F1 score is defined as follows:

$$\text{F1} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \tag{3}$$

and acts as a weighted average of precision and recall.

Table 2
The Table Summarizes the Precision, Recall, and F1 Score for Different Classifiers on the Aurora Data Set

Algorithm	Precision	Recall	F1 score	# of parameters (M)
3NN	0.84 ± 0.0	0.56 ± 0.0	0.58 ± 0.0	0
5NN	0.66 ± 0.0	0.54 ± 0.0	0.53 ± 0.0	0
Linear SVM	0.78 ± 0.0	0.70 ± 0.0	0.72 ± 0.0	≈0
Clausen and Nickisch (2018)	0.88 ± 0.01	0.87 ± 0.01	0.88 ± 0.01	43
VGG-16	0.84 ± 0.02	0.80 ± 0.03	0.81 ± 0.02	138
VGG-19	0.82 ± 0.04	0.78 ± 0.03	0.79 ± 0.03	143
AlexNet	0.88 ± 0.03	0.88 ± 0.03	0.87 ± 0.03	60
ResNet-18	$\mathbf{0.92 \pm 0.02}$	0.87 ± 0.05	0.89 ± 0.04	11
ResNet-50	$\mathbf{0.92 \pm 0.03}$	$\mathbf{0.89 \pm 0.04}$	$\mathbf{0.90 \pm 0.03}$	25

Note. The reported scores are the average over 10 runs with different random initialization. The last column lists the number of parameters, in millions, for each classifier, rounded to the nearest whole million. The bold entries mark the highest performance.

3.1. Comparison of the Classification Performances

The performance results of the six deep learning-based models described in section 2.3 are displayed in Table 2. Additionally, a baseline using two well-known machine learning classifiers is provided, namely, a KNN classifier (Theodoridis & Koutroumbas, 2009a) and a SVM classifier (Cortes & Vapnik, 1995). For both classifiers, the histogram of oriented gradients method (Dalal & Triggs, 2005) is used for extracting features from the images. For the KNN classifier, the results when considering the three and five nearest neighbors are reported. For the SVM, the results are for where a linear kernel is utilized. All deep learning-based models outperform the KNN and SVM baseline. The ResNet-50 achieved the highest score across all metrics.

Figure 4. The figure shows the confusion matrix for the ResNet-50 network on the test data. The diagonal displays the percentage of correctly classified images for each class, that is, images where the network automatically classified an auroral image similarly to the manually labeled test data. The off-diagonal elements show the percentage of images from a given class erroneously classified as another class. The results show that the model achieves a high accuracy on most classes, but has some difficulties with separating *colored aurora* and *discrete-irregular* aurora.

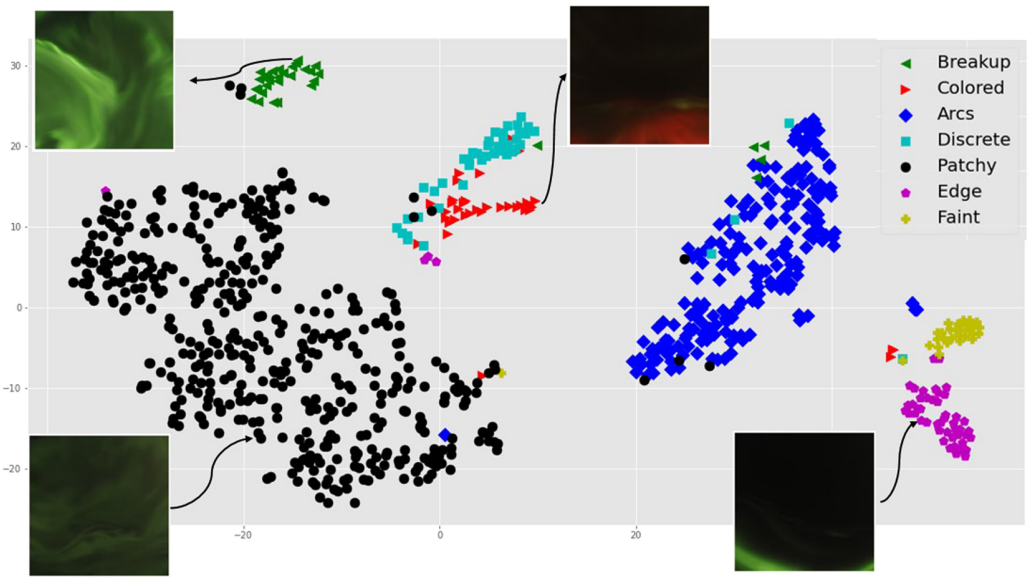


Figure 5. The figure presents the presoftmax representation of the aurora test data produced by the ResNet-50 network and projected down to two dimensions using T-SNE. Images from some of the classes are also displayed in the plot. The figure shows that the model has found a suitable representation where the classes are clustered together and are easily separable.

Figure 4 shows the confusion matrix for the ResNet-50 on the test set of the aurora data. From this figure, it is evident that some classes are well understood by the network. Specifically, almost all samples from the *auroral arcs*, *patchy aurora*, *edge aurora*, and *faint-clear* classes in the test set are classified correctly. On the other hand, the *colored aurora* is more challenging and is partly classified as *discrete-irregular* aurora. Also, both *auroral breakup* and *discrete-irregular* are partly classified as *auroral arcs*.

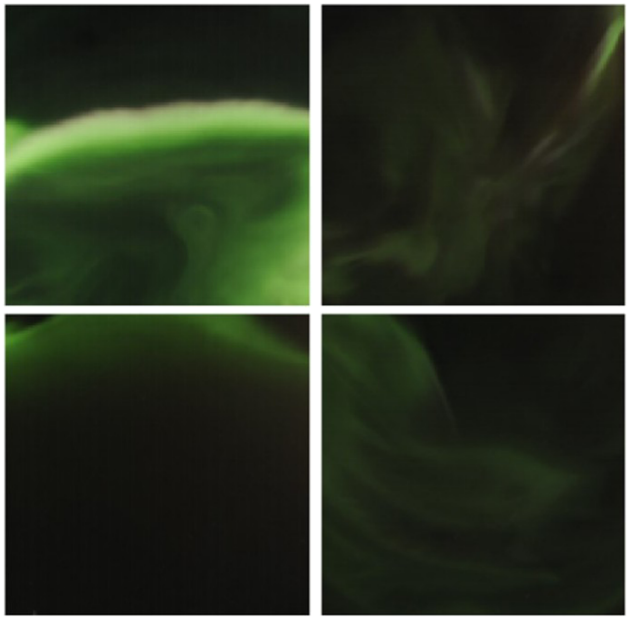


Figure 6. This figure depicts test images that were incorrectly classified by the ResNet-50 network. (top left) *Auroral breakup* classified as *arcs*. (top right) *Colored aurora* classified as *discrete-irregular*. (bottom left) *Edge aurora* classified as *arcs*. (bottom right) *Patchy aurora* classified as *discrete-irregular*. All images are processed according to the four-step processing procedure described in section 2.1.

3.2. Details of the Highest Performing Model (ResNet-50)

Figure 5 displays a two-dimensional presoftmax representation of the test data of the aurora data set, produced by the ResNet-50. The representation was projected down to two dimensions using the t-distributed stochastic neighbor embedding (T-SNE) dimensionality reduction technique (van der Maaten & Hinton, 2008). The figure also contains some examples of the actual images that are represented by the two-dimensional points. It is from this representation that the network determines what class to assign a sample image to. From this figure, it is clear that some classes are well separated from the other classes. For instance, the cluster of *edge aurora* and the cluster of *faint-clear* samples toward the rightmost part of the figure are compactly represented and well separated from the other classes. However, some classes are more mixed together, which corroborates the findings in Figure 4.

Figure 6 shows four examples of incorrectly classified images from the auroral test data. The incorrect classifications are generally sensible, and it can be seen how the algorithm may have opted for the incorrect classification. Possible interpretations are as follows: Figure 6 (top left) shows *auroral breakup* classified as *arcs*; although the breakup is imminent, the image is still dominated by the bright arc. Figure 6 (top right) shows *colored aurora* classified as *discrete-irregular* aurora; the distinction of color is subtle compared to other colored aurorae in the training set. Additionally, the strong non-east-west feature in the top right is found in *discrete-irregular* aurora cases. Figure 6 (bottom left) shows *edge aurora*

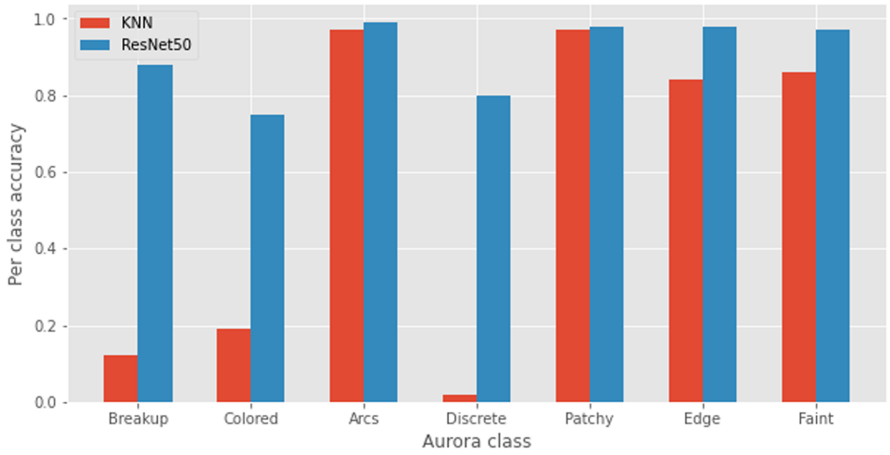


Figure 7. The figure shows the class wise accuracy comparison between a 3NN-classifier and the ResNet-50 network. The figure illustrates how the deep learning-based model outperform the traditional machine learning method in locating all classes. The ResNet-50 classifier handles particularly the breakup, colored, and discrete aurora classes much better than the KNN-classifiers.

classified as *arcs*; in fact, this is an arc, but it is on the edge of the field of view. The intensity of the arc is not high, so the brightness in the corners is not sufficient for the algorithm to opt for that class. Figure 6 (bottom right) shows *patchy aurora* classified as *discrete-irregular*; there is a sharp edge feature in the center-top of the image, which is generally not present in *patchy aurora* cases.

To examine what type of images the deep learning-based models are capable of recognizing compared to the traditional machine learning approach, we compare the class wise accuracy of the 3NN-classifier and the ResNet-50. Figure 7 displays the class-wise accuracy on the test images of the aurora data set. The accuracy of some classes is comparable between the two methods, but particularly for the *auroral breakup*, *colored aurora*, and *discrete-irregular* auroral classes it is clear that the deep learning approach is superior. These classes are typically more challenging to recognize, which seems to suggest that the deep learning-based approach is capable of identifying more complicated structures in auroral images.

4. Discussion

In this study, the DNNs generally outperformed the KNN and SVM techniques. Overall, from the presented results in Table 2 where several architectures, namely, the Clausen and Nickisch (2018) method, AlexNet, ResNet-18, and ResNet-50, achieved close to 90% average precision, infer that aurora classification is a suitable task for DNNs. It has been demonstrated that it is possible to classify specific auroral forms such as *auroral arcs*, *edge aurora*, *patchy aurora*, and *faint clear* with a high precision (>90%), as seen in Figure 4. The networks were in general less successful in classifying *auroral breakup*, *colored aurora*, and *discrete-irregular* aurora, likely reasons are too few and/or too ambiguous training images. In addition, from Figure 5 it is clear that colored aurora has an overlap with *discrete-irregular* aurora. This overlap occurs when the aurora does not have a clearly pronounced colored emission. Note that it might be possible to achieve a higher label separability using unsupervised clustering methods. However, an additional goal of this study was to define auroral labels that represents exclusive production mechanisms or characterize physical properties of the aurora when possible. The resulting clusters attained by an unsupervised clustering method will most likely not satisfy this goal. An unsupervised classifier might therefore not be applicable for most scientific purposes where the aim is to study the physics of the aurora and its production mechanisms.

The DNN architectures VGG-16 and VGG-19 had the worst performance. These models have a significantly higher amount of parameters, as presented in Table 2, and it might be that the data set is not large enough for training models of such a size. Also, the ResNet models are generally known to outperform AlexNet and the VGG-based models (He et al., 2016), often attributed to their ability to propagate gradients effectively even for very deep networks (Balduzzi et al., 2017).

Clausen and Nickisch (2018) used a pretrained DNN to extract features and then trained a linear classifier using these features. This means that the parameters of the pretrained DNN are not optimized for handling

auroral images but for other types of images. In contrast, the models used in this study, VGG, AlexNet, and ResNet, are only trained using auroral images. From Table 2 it is evident that the pretrained approach gives better performance than the traditional machine learning algorithms but does slightly worse than the best DNNs trained on only auroral images. This suggests that training the models specifically for classifying images of aurora might improve their capability to detect different types of aurora. However, a larger data set is needed to validate this proposition.

An analysis of the network performance on data from other cameras still remains to be done. Note that the color sensitivity of the detector, presented in Figure 1, is a camera-specific feature. Thus, the RGB images acquired by other cameras should be adjusted to account for the differences in the wavelength-dependent light response before being used by the classifier. In addition, the training and testing was done using images without clearly ambiguous images, as described in section 2. Hence, a proper analysis of how ambiguous images is classified also needs to be investigated. The network will ideally classify the most dominant auroral feature. However, an evaluation of the performance on ambiguous images will itself be subject to biases and subjective interpretations, as described by McKay and Kvammen (2020). It is irrational to expect a DNN to classify ambiguous images correctly if not even auroral experts can agree on what the correct label is. Thus, a common consensus about the auroral morphology and the criteria for each class needs to be introduced before progress can be made on classifying ambiguous images. Alternatively, one could interpret ambiguous auroral images as a mixture of more common classes (e.g., *auroral breakup*, *arcs*, and *patchy aurora*) and label ambiguous images with multiple classes.

Future endeavors in aurora classification with DNNs should investigate the dimension space using different cameras and auroral events, for instance, by T-SNE maps. Further improvements to the classifier can likely be achieved by including the time dimension. The time dimension information can be included by combining the CNNs with recurrent neural networks (RNNs) or by using the CNNs directly on a data set consisting of labeled stacks of consecutive images with, for example, ∼10 images in each stack, the appropriate number of images in each stack depends on the imaging frequency, the exposure time, and the field of view. The inclusion of the time dimension also allows for classifying labels with a distinctive temporal behavior (e.g., *pulsating aurora*). From the data set used in this study, it can be concluded that images labeled as *colored aurora* and *discrete-irregular* are not well separated by the classifiers. Other labeling sets should therefore be considered. It should be noted that classifying colored aurora is of particular interest for detecting exciting phenomena such as sunlit aurora (Shiokawa et al., 2019; Størmer, 1929), stable auroral red (SAR) arcs (Mendillo et al., 2016; Rees & Roble, 1975), and Strong Thermal Emission Velocity Enhancement (STEVE) (Gallardo-Lacourt et al., 2018; MacDonald et al., 2018), which are characterized by pronounced colored emissions. In addition, both *auroral arcs* and *patchy aurora* achieved a high precision and are common auroral forms. Subdividing these classes, for instance, *arcs* into *single arc* and *multiple arcs*, might therefore be advantageous. Furthermore, auroral omega bands are quite common and distinctive auroral forms (Partamies et al., 2017) which might be possible to classify with a high precision.

5. Conclusion

This paper presents the results of an extensive study of automatic aurora classification using different machine learning techniques. Seven auroral classes were considered: *auroral breakup*, *colored aurora*, *auroral arcs*, *discrete-irregular*, *patchy aurora*, *edge aurora*, and *faint-clear*. The classifiers were both trained and tested on clean colored (RGB) auroral images, without clearly ambiguous auroral forms and unwanted features, such as clouds and light pollution. Six DNN architectures were tested along with the well-known machine learning classifiers KNN and SVM. The ResNet-50 DNN architecture achieved the highest performance with an average classification precision of 92%.

Overall, the conclusion is that automatic auroral image classification is a suitable task for DNNs. The DNNs generally outperformed the KNN and SVM techniques. However, progress in this field of study is constrained by biases and subjective interpretations (McKay & Kvammen, 2020). It is irrational to expect the DNNs to classify an auroral image correctly if auroral researchers cannot agree on what the correct aurora label is. High precision (>90%) classification of clearly ambiguous auroral images can therefore not be readily achieved before a common consensus about the auroral morphology and the criteria for each class is established. The use of data without clearly ambiguous auroral forms for automatic aurora classification,

mainly labeled into physically meaningful definitions, might be the first step in sorting auroral structures in a morphological space.

Data Availability Statement

The image data archive are freely accessible at https://www2.irf.se/allsky/data.html, and the processed image data set and code used in this paper are available online (at https://dataverse.no/dataset.xhtml?persistentId=doi:10.18710/SSA38J).

Acknowledgments
The authors would like to thank Urban Brändström and the Swedish Institute of Space Physics for providing the original auroral image data. However, the users are obliged to contact the Kiruna Atmospheric and Geophysical Observatory before using the images. Andreas Kvammen is supported by the Tromsø Research Foundation. The work by Derek McKay is partly supported by the Academy of Finland project number 322535. The work by Noora Partamies is partly supported by the Research Council of Norway under contract 287427 and a CoE contract 223252. The authors would also like to thank Björn Gustavsson for valuable suggestions and comments.

References

Balduzzi, D., Frean, M., Leary, L., Lewis, J. P., Ma, K. W.-D., & McWilliams, B. (2017). The shattered gradients problem: If resnets are the answer, then what is the question? In D. Precup & Y. W. Teh (Eds.), *Proceedings of the 34th International Conference on Machine Learning* (Vol. 70, pp. 342–350), Proceedings of Machine Learning Research. International Convention Centre, Sydney, Australia: PMLR. Retrieved from http://proceedings.mlr.press/v70/balduzzi17b.html

C.D.W. (1964). International Auroral Atlas. Published for the International Union of Geodesy & Geophysics, Edinburgh (University Press), 1963. pp. 17; 2 Figures; 52 black and white, 4 colour plates. 45s. *Quarterly Journal of the Royal Meteorological Society*, *90*(386), 502–502. https://doi.org/10.1002/qj.49709038624

Clausen, L. B. N., & Nickisch, H. (2018). Automatic classification of auroral images from the Oslo Auroral THEMIS (OATH) data set using machine learning. *Journal of Geophysical Research: Space Physics*, *123*, 5640–5647. https://doi.org/10.1029/2018JA025274

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. https://doi.org/10.1023/A:1022627411411

Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* (Vol. 1, pp. 886–893). San Diego, CA, USA, USA: IEEE. https://doi.org/10.1109/CVPR.2005.177

Gallardo-Lacourt, B., Nishimura, Y., Donovan, E., Gillies, D. M., Perry, G. W., Archer, W. E., et al. (2018). A statistical analysis of STEVE. *Journal of Geophysical Research: Space Physics*, *123*, 9893–9905. https://doi.org/10.1029/2018JA025368

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, *3*, 1157–1182. Retrieved from http://dl.acm.org/citation.cfm?id=944919.944968

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 770–778). Las Vegas, NV, USA: IEEE. https://doi.org/10.1109/CVPR.2016.90

Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *Proceedings of the 32Nd International Conference on International Conference on Machine Learning - Volume 37* (pp. 448–456), ICML'15: JMLR.org. JMLR.org. Retrieved from http://dl.acm.org/citation.cfm?id=3045118.3045167

Karlsson, T., Andersson, L., Gillies, D. M., Lunch, K., Marghitu, O., Partamies, N., et al. (2020). Quiet, discrete auroral arcs—Observations. *Space Science Reviews*, *216*, 16.

Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*. Retrieved from http://arxiv.org/abs/1412.6980

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Proceedings of the 25th International Conference on Neural Information Processing Systems* (Vol. 1, pp. 1097–1105). NIPS'12. USA: Curran Associates Inc. http://dl.acm.org/citation.cfm?id=2999134.2999257

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature Cell Biology*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Lysak, R., Echim, M., Karlsson, T., Marghitu, O., Rankin, R., Song, Y., & Watanabe, T.-H. (2020). Quiet, discrete auroral arcs: Accelerations mechanisms. *Space Science Reviews*, *216*, 92.

MacDonald, E. A., Donovan, E., Nishimura, Y., Case, N. A., Gillies, D. M., Gallardo-Lacourt, B., et al. (2018). New science in plain sight: Citizen scientists lead to the discovery of optical structure in the upper atmosphere. *Science advances*, *4*(3), eaaq0030.

Mauer, C. (2009). Measurement of the spectral response of digital cameras with a set of interference filters. University of Applied Sciences Cologne. Department of Media- and Phototechnology.

McKay, D., & Kvammen, A. (2020). Auroral classification ergonomics and the implications for machine learning. *Geoscientific Instrumentation, Methods and Data Systems*, *9*(2), 267–273.

Mendillo, M., Baumgardner, J., & Wroten, J. (2016). SAR arcs we have seen: Evidence for variability in stable auroral red arcs. *Journal of Geophysical Research: Space Physics*, *121*, 245–262. https://doi.org/10.1002/2015JA021722

Nishimura, Y., Lessard, M., Katoh, Y., Miyoshi, Y., Grono, E., Partamies, N., et al. (2020). Diffuse and pulsating aurora. *Space Science Reviews*, *216*, 4.

Nishimura, Y., Lyons, L., Zou, S., Angelopoulos, V., & Mende, S. (2010). Substorm triggering by new plasma intrusion: THEMIS all-sky imager observations. *Journal of Geophysical Research*, *115*, A07222. https://doi.org/10.1029/2009JA015166

Partamies, N., Juusola, L., Whiter, D., & Kauristie, K. (2015). Substorm evolution of auroral structures. *Journal of Geophysical Research: Space Physics*, *120*(7), 5958–5972. https://doi.org/10.1002/2015JA021217

Partamies, N., Weygand, J. M., & Juusola, L. (2017). Statistical study of auroral omega bands. *Annales Geophysicae*, *35*(5), 1069–1083. https://doi.org/10.5194/angeo-35-1069-2017

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., et al. (2017). Automatic differentiation in PyTorch.

Rao, J., Partamies, N., Amariutei, O., SyrjÃ¤suo, M., & van de Sande, K. E. A. (2014). Automatic auroral detection in color all-sky camera images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *7*(12), 4717–4725. https://doi.org/10.1109/JSTARS.2014.2321433

Rees, M. H., & Roble, R. G. (1975). Observations and theory of the formation of stable auroral red arcs. *Reviews of Geophysics*, *13*(1), 201–242.

Rhodes, B. (2019). Skyfield v1.10. Retrieved from http://rhodesmill.org/skyfield/

Shiokawa, K., Otsuka, Y., & Connors, M. (2019). Statistical study of auroral/Resonant-Scattering 427.8-nm emission observed at subauroral latitudes over 14 years. *Journal of Geophysical Research: Space Physics*, *124*, 9293–9301. https://doi.org/10.1029/2019JA026704

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. CoRR, abs/1409.1556.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, *15*, 1929–1958. Retrieved from http://jmlr.org/papers/v15/srivastava14a.html

Störmer, C. (1913). On an auroral expedition to bossekop in the spring of 1913. *The Astrophysical Journal*, *38*, 311.
Størmer, C. (1929). The distribution in space of the sunlit aurora rays. *Nature*, *123*(3090), 82–83.
Syrjäsuo, M., & Donovan, E. (2002). Analysis of auroral images: Detection and tracking. *Geophysica*, *38*(1–2), 3–14.
Syrjäsuo, M. T., Donovan, E. F., Qin, X., & Yang, Y. H. (2007). Automatic classification of auroral images in substorm studies. In *8th International Conference on Substorms (ICS8)* (pp. 309–313).
Syrjäsuo, M. T., & Donovan, E. F. (2004). Diurnal auroral occurrence statistics obtained via machine vision. *Annales Geophysicae*, *22*(4), 1103–1113. https://doi.org/10.5194/angeo-22-1103-2004
Syrjasuo, M., & Partamies, N. (2011). Numeric image features for detection of aurora. *IEEE Geoscience and Remote Sensing Letters*, *9*(2), 176–179.
Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017). Inception-v4, Inception-ResNet and the impact of residual connections on learning. *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence* (pp. 4278–4284). Palo Alto, CA: AAAI Press.
Theodoridis, S., & Koutroumbas, K. (2009a). Chapter 2—Classifiers based on Bayes decision theory. In S. Theodoridis & K. Koutroumbas (Eds.), *Pattern recognition (fourth edition)* (4th ed., pp. 13–89). Boston: Academic Press. https://doi.org/10.1016/B978-1-59749-272-0.50004-9
Theodoridis, S., & Koutroumbas, K. (2009b). Chapter 4—Nonlinear classifiers. In S. Theodoridis & K. Koutroumbas (Eds.), *Pattern recognition (fourth edition)* (4th ed., pp. 151–260). Boston: Academic Press. https://doi.org/10.1016/B978-1-59749-272-0.50006-2
van der Maaten, L. J. P., & Hinton, G. E. (2008). Visualizing high-dimensional data using t-SNE. *Journal of Machine Learning Research*, *9*, 2579–2605.
Xing, X., Lyons, L., Nishimura, Y., Angelopoulos, V., Larson, D., Carlson, C., et al. (2010). Substorm onset by new plasma intrusion: THEMIS spacecraft observations. *Journal of Geophysical Research*, *115*, A10246. https://doi.org/10.1029/2010JA015528
Yang, Q., Liang, J., Hu, Z., & Zhao, H. (2012). Auroral sequence representation and classification using hidden Markov models. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(12), 5049–5060. https://doi.org/10.1109/TGRS.2012.2195667
Yang, Q., Tao, D., Han, D., & Liang, J. (2019). Extracting auroral key local structures from all-sky auroral images by artificial intelligence technique. *Journal of Geophysical Research: Space Physics*, *124*, 3512–3521. https://doi.org/10.1029/2018JA026119

PAPER II:
Auroral classification ergonomics and the implications for machine learning

McKay, D. and Kvammen, A.: Auroral classification ergonomics and the implications for machine learning, Geoscientific Instrumentation, Methods and Data Systems, 9, 267–273, doi: 10.5194/gi-9-267-2020, 2020

Geosci. Instrum. Method. Data Syst., 9, 267–273, 2020
https://doi.org/10.5194/gi-9-267-2020

Geoscientific Instrumentation Methods and Data Systems
Open Access
EGU

Auroral classification ergonomics and the implications for machine learning

Derek McKay[1] **and Andreas Kvammen**[2]

[1]NORCE Norwegian Research Centre AS, Tromsø, Norway
[2]Department of Physics and Technology, UiT – The Arctic University of Norway, Tromsø, Norway

Correspondence: Derek McKay (demc@norceresearch.no)

Received: 6 December 2019 – Discussion started: 28 January 2020
Revised: 6 June 2020 – Accepted: 16 June 2020 – Published: 9 July 2020

Abstract. The machine-learning research community has focused greatly on bias in algorithms and have identified different manifestations of it. Bias in training samples is recognised as a potential source of prejudice in machine learning. It can be introduced by the human experts who define the training sets. As machine-learning techniques are being applied to auroral classification, it is important to identify and address potential sources of expert-injected bias. In an ongoing study, 13 947 auroral images were manually classified with significant differences between classifications. This large dataset allowed for the identification of some of these biases, especially those originating as a result of the ergonomics of the classification process. These findings are presented in this paper to serve as a checklist for improving training data integrity, not just for expert classifications, but also for crowd-sourced, citizen science projects. As the application of machine-learning techniques to auroral research is relatively new, it is important that biases are identified and addressed before they become endemic in the corpus of training data.

1 Introduction

Each year, the all-sky cameras in the Arctic and Antarctic regions collect several million images of the sky. These contain a plethora of atmospheric and astronomical phenomena including, of particular interest to the authors, manifestations of the aurorae. Auroral emissions are excited when charged particles from the magnetosphere enter the ionosphere and collide with atoms, molecules and ions in the ionosphere. The flux of energetic charged particles entering the ionosphere is dependent on solar wind as well as magnetospheric and ionospheric conditions. Different conditions will result in different auroral features; thus, the auroral sky acts as a window into the otherwise obscure solar wind–magnetosphere–ionosphere environment. Having computer algorithms to pick out interesting features, or features that have potential for scientifically interesting phenomena, is helpful for scientists in auroral research.

Examination of what can be done using machine learning for such interests has been pursued, and there are other groups doing the same. Ideally, autonomous software would take a set of images and identify those which contain aurorae and, in these cases, which morphological types are present (break-up, arcs, discrete, patchy etc.).

Yet, although algorithms for the identification of visual features have made remarkable progress, these tend to be "exceptionally data-hungry". It is well-established that it is expensive and tedious to produce large, labelled training datasets, especially in cases in which expert knowledge is required (e.g. Yu et al., 2015).

Although initial attempts have been made to undertake automatic auroral classification, these have not been particularly successful (low prediction rates) or useful (high accuracy, but the categories are so broad as to not really be of significant benefit). Those programmes which have demonstrated success have focused on very specific subgroups (e.g. Yang et al., 2019).

Part of the problem with low success rates for prediction is the presence of prediction bias (e.g. Domingos, 2000). This can be attributed to various causes, such as the following.

- Noisy training data
- An incomplete feature set
- Strong regularisation
- Algorithmic errors
- Biased training samples

In order to address these issues, a programme was undertaken to improve the reliability of machine-learning results (Kvammen et al., 2020a). In addition to using more up-to-date machine-learning methods, attention was paid to the provision of a comprehensive dataset for training the algorithms. As a part of this process, it was deemed important to remove sources of bias in the classification of the training dataset. Following some preliminary work with small samples of both greyscale and colour images, a main classification run was undertaken.

As the differences between the classifications of informed researchers were significant, the findings are presented here. It is intended that they will serve as a reference point for other endeavours in the development of machine-learning training and test sets for auroral research, but also for any other field in which machine-learning image recognition is developed from specialised sets categorised by subject experts. The paper concludes with a discussion, including references to similar work from other disciplines and suggestions for future work.

2 Methodology

Images from the Kiruna all-sky camera (location: 67.84° N, 20.41° E; 425 m above mean sea level; operated by the Swedish Institute for Space Physics) were used. The camera is a Nikon D700 with a Nikon Nikkor 8 mm 1 : 2.8 lens giving an almost 180° field of view. The exposure time is 6 s taken automatically on each minute. To ease data transfer rates and processing the "quick-look summary" JPEG images (720×479 pixels) were used, rather than the full-resolution images. Approximately 300 000 of these images from nine winter seasons were filtered down to a set of 13 947 for human classification (by removing cloudy, moonlit and twilight images). A 3×3 median filter was applied to filter out points (stars, defective pixels, etc.), then the data were binned using a 2×2 averaging window to reduce the size of the images for neural network training. The central 128×128 pixels were then selected. This removes the horizon, where distortion is more pronounced and where light pollution and atmospheric effects are typically found. Additional information about the pre-processing and machine learning is reported in Kvammen et al. (2020a).

Two auroral physicists each classified these 13 947 processed images by hand using different software implementations (one using Python, the other using MATLAB). The motivation to do this was based on knowledge of the different systems (thus making self-maintenance of the code possible) and as a way of working independently to ensure robust results.

The classification was done according to nine possible classes. The labels are listed in Table 1 along with a brief description of each label; a more thorough description of the labels and the labelling procedure is available in Kvammen et al. (2020a). These classes were the result of several iterations of planning, whereby the two experts, together with a machine-learning researcher, identified categories which would be scientifically useful, possible to discern with a reasonable algorithmic network and suitable for the sample size available. Sample images (after pre-processing) illustrating each auroral label are presented in Fig. 1.

After comparison of results, it was found that the experts only agreed on 54 % of the images, with the most disagreement being on which images were suitable for training and which had an aurora with an unknown and/or complicated form. When it was agreed that the image was suitable for algorithm training, the experts agreed on 95 % of the labels. By only using the images with agreeing auroral labels (i.e. bother experts independently reached the same classification) and by excluding the images with ambiguous auroral forms, unwanted features and disagreeing labels, a clean training dataset was produced at the price of excluding approximately 73 % of the 13 947 images in the initial dataset. The experimental results of the labelled dataset, as derived from the machine-learning study, are not considered in this paper but are presented in Kvammen et al. (2020a).

3 Ergonomic categories

The comparison of the classifications for both the trials and the main classification run allowed for the identification of emerging biases based on the approach each researcher took to identify the aurorae in the images. These biases are a result of the levels of comfort (physical and cognitive) that exist during a classification process, leading to the term "classification ergonomics". Those identified as part of this study are shown in Table 2 and are discussed in the subsequent sections.

3.1 Physical comfort bias

The classification of the aurorae in the main study was a nine-class system. Given the designations, the number keys were the obvious choice, and the classification software used these either on the main keyboard (0–8) or the numeric keypad (KP0–KP8). In the case of a mistake, it was possible to go back to the previous image, and the backspace key was used to accomplish this. This key configuration is shown in Fig. 2.

The first bias that was noted was the inconvenience of the backspace for making corrections. This required moving the

Table 1. The set of aurora labels used for the classification run.

Label no.	Label	Description
0	Auroral break-up	Bright auroral forms that cover most of the image
1	Coloured aurora	The auroral emission is clearly not monochromatic green
2	Auroral arcs	Auroral structures with clear east–west-aligned form
3	Discrete–irregular	A combination of broken arcs, north–south-aligned arcs and vortical structures
4	Patchy aurora	Aurora appears as irregular blobs or stripes on a diffuse background
5	Edge aurora	Auroral emission only at the edge of the framed image
6	Faint–clear	Auroral emission is not clearly visible
7	Unknown or ambiguous	Aurora does not fit any of the labels above or is a mixture of several labels
8	Rejected	The image is not suitable for training due to e.g. light pollution, clouds, noise

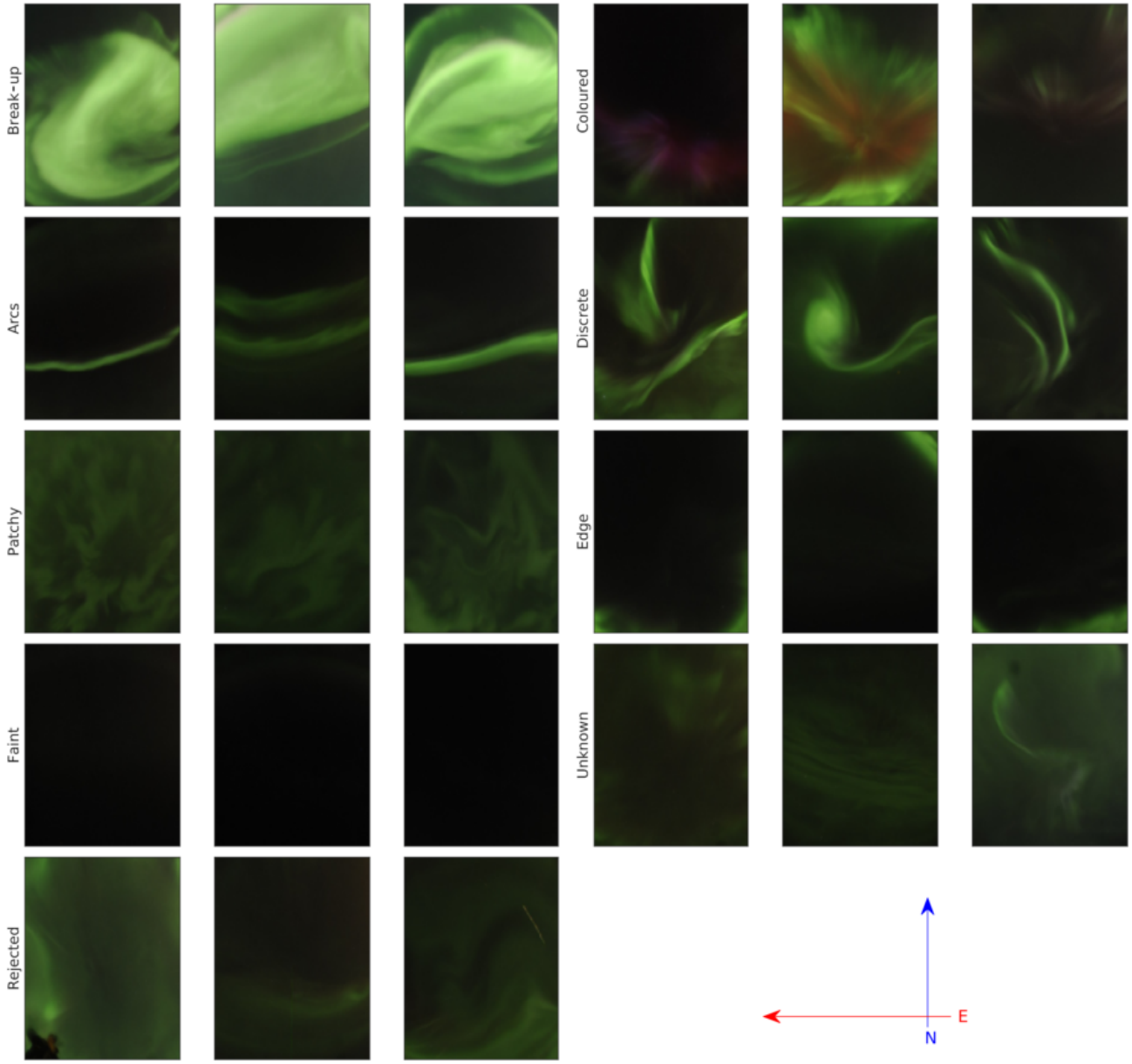


Figure 1. Sample images of each label. The direction with respect to the magnetic pole is indicated by the arrows at the bottom right. The rejected images at the bottom left are rejected due to, from left to right, a person in the bottom left corner, lidar emission and passing aircraft.

Table 2. A set of biases that may affect user classification.

Physical comfort bias
Data contrast bias
Environment contrast bias
Repetition bias
Learning bias
Feature bias
Ambiguity bias
Expert knowledge bias

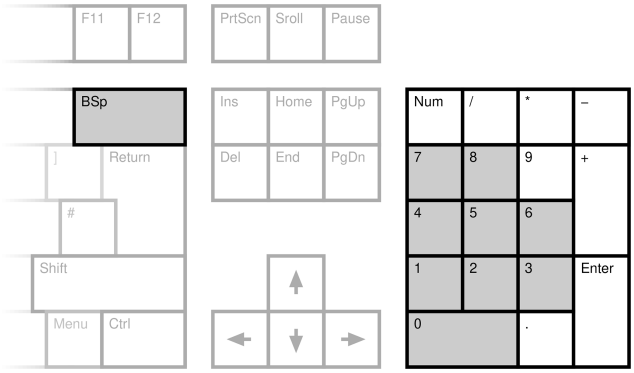


Figure 2. The original right-hand key set for classification. The grey-shaded keys were used for the classification.

right hand completely away from the rest position where the fingers hover over the KP4, KP5 and KP6 keys on the keypad. As this was awkward, there was a perceptible reluctance to make corrections. Thus, the KP-DECIMAL (to the right of the KP0 key) was used as an alias.

After several hundred classifications, discomfort was experienced, even with the keyboard rotated 10–20° anticlockwise to make the keys suit the angle of the right hand. As a result, some testing was also done with more comfortable key arrangements. This resulted in a basic WASD configuration being used. WASD refers to the directional (move forward, backward, left and right) keys as used in FPS (first-person shooter) computer games.

This configuration is shown in Fig. 3, where the coloured circles show the at-rest position of the fingers (keys A, S and D), with the arrows showing easy-reach positions. The left thumb rests on the space bar. The little finger typically can reach the shift and control keys (as a modifier; in FPS games this might be e.g. run and crouch) but were not used here. The actual keys that were used for the classification are shaded in grey.

Additionally, the keyboard was rotated 10–15° clockwise to match the natural angle of the left wrist and hand, as shown in Fig. 4. This was used for most of the classification work, and no discomfort was experienced.

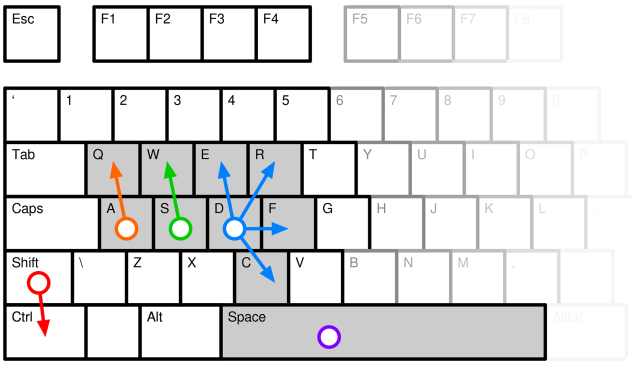


Figure 3. A left-hand key set for classification. The coloured circles show the at-rest position of the fingers, with the arrows showing easy-reach positions. The grey-shaded keys were used for the classification. See also Fig. 4.

Figure 4. Hand position for the left-hand key set, with the keyboard at an angle of 10–15° to minimise finger reach strain.

3.2 Data contrast bias

If the classifier has just seen a faint, patchy aurora, then a following faint, patchy aurora is likely to be classified the same. If the preceding image was a bright break-up, then it is more likely for the faint, patchy aurora to be classified as blank. In the initial parts of the study, attempts were made to mitigate this by normalising the image scale of all images. This was not readily achieved with colour images and thus not pursued. In the study, the classifications were done both randomly and chronologically by the two experts. Repeated, random classification would be best but was not possible within the time limitation of the project. The chronological classification allows knowledge of the substorm process to be applied to obtain a more reliable result, although this introduces anticipation of the phenomenon rather than an objective evaluation of each image on its own characteristics (Sect. 3.8).

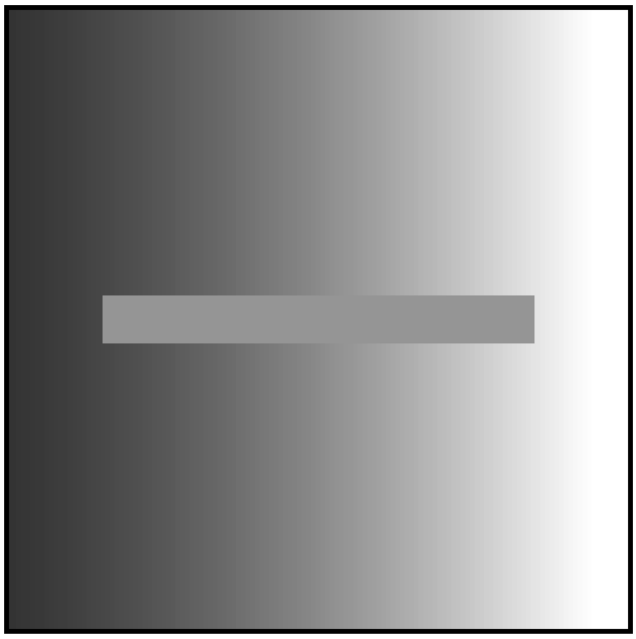


Figure 5. Example of environment contrast. The central bar may seem to be a gradient, but it is uniform in shade. Cover the surrounding (real) gradient with paper of uniform colour to demonstrate this.

3.3 Environment contrast bias

Humans naturally retain perceptual constancy. This allows visual features to be discerned against a noisy or changing background: a trait that is useful to all animals in a hunter–prey scenario, for instance. However, this human trait of retaining perceptual constancy results in optical illusions. Colour constancy and brightness constancy will cause an illusion of colour or contrast difference when the luminosity or colour of the area surrounding an object is changed. The eye partly does this as a result of compensating for the overall lighting (change in the iris aperture), but the brain also compensates for subtle changes within the field of view. An example of this is shown in Fig. 5.

Originally, the software presented the image with a white border (the default for the plotting software). However, the contrast made it difficult to discern the difference between features which were faint but still recognisable and those which were sub-threshold for visual identification. Hence, the figure background was changed to black. This made it easier to discern the borderline cases.

The environmental conditions beyond the computer screen were also significant, with differences in the ambient lighting and room brightness being an issue. This was noted, and consistency in the arrangements for the process is likewise recommended for future studies.

3.4 Repetition bias

It is more comfortable to press the same button twice than to press two different buttons. Additionally, if a mistake is made, it is extra effort to go back and correct it. This "laziness" accumulates during the classification process, making long sessions problematic.

For example, if there are 10 similar images in a row, the chance of classifying number 11 in the same way is higher than if there were 10 random images first. In the study single-repetition bias was 27 %, rising to 40 % for double-repetition bias. It can be mitigated with randomisation (but see Sect. 3.2), whereby different experts are presented with a different random sequence of images. Discrepancy between classifications can then be investigated or the images discarded.

3.5 Learning bias

If there are lots of categories, the classifier may not necessarily hold all of them in mind. Thus, some "sectors" of the classification may have a higher activation energy than others. For example, when classifying hundreds of arcs and patchy aurorae and then getting a discrete case, the classifier may subconsciously think "it is not patchy, so it must be an arc", thereby inadvertently omitting the thought of a different class. This is a recency effect (whereby a new classification is biased toward the set of most recently used labels) which has been reported in the biological sciences (Culverhouse, 2007). Randomisation can be used, but re-classification can also be employed to test for variation due to learning.

3.6 Feature bias

The classifier is more likely to get the classification of a prominent feature correct than faint or diffuse features. This leads to a form of confusion bias; e.g. what to do with a bright discrete aurora (Class 3) on a background of diffuse patchy aurora (Class 4).

There is also positivity bias, whereby identification is biased by prior expectations (Culverhouse, 2007). In auroral classification, the substorm progression (development of the auroral display) makes it possible to anticipate the next image. This is partially mitigated by randomising the samples, but this can lead to contrast bias (Sect. 3.2).

3.7 Ambiguity bias

Ambiguity bias occurs when there is confusion as to what a particular image may be. This is exacerbated by feature bias (Sect. 3.6). However, even in cases in which there is no dominant feature, the classifier will tend to subconsciously identify some feature and latch onto it to the exclusion of other features in the image. For the main study, it was decided that ambiguous images should be rejected to make the learning environment clearer for the machine-learning algo-

rithms – thus, such ambiguities were undesirable. What was noted was that, especially early in the study, the users would tend to try to classify the auroral image rather than reject it. When it was clear that there was no shortage of data, this tendency was reduced.

Nevertheless, it is recommended that there is a clarification of classification rules, making it clear what the user should do in a case of mixed features. If there is a precedence or priority of forms, then that should also be made very clear. Even so, ambiguities and borderline cases will remain. When data volume allows, these could be discarded.

3.8 Expert knowledge bias

Differences in "expert knowledge" that affect the results have also been seen. For example, although the two expert researchers involved are knowledgeable in auroral physics and its optical manifestation in general, one had done research on auroral arcs, whereas the other had not. The specialist was more picky on the arc classification (classifying 15 % fewer), partially as a result of having a deeper understanding of the underlying physics but also partly in terms of having seen many more images prior to approaching the classification task. This led to a higher level of discernment on that particular category. This can be mitigated by establishing clear guidelines and "recognition cards" to assist the classification process. In cases in which there is ambiguity or disagreement, revision of the characteristics being used to do the classification can be carried out and re-classification done where necessary.

4 Discussion

The application of machine learning to auroral classification is an area in which only a few studies have been carried out. However, development is now progressing rapidly and it is likely that it will be applied much more and become an important part of auroral research in the future. Therefore, it is vitally important to properly address the ergonomics and biases sooner rather than later in order to avoid inadvertently introducing errors and biases early in the establishment of this new area of science.

Discrepancy between expert classifiers has been reported before. A previous auroral study had two experts that agreed on the class in about 70 % of the images, and the experts chose the unknown class in almost 50 % of all images (Syrjäsuo et al., 2007). However, an analysis of potential reasons for the discrepancy was not included. Similarly, a biological study found that trained personnel achieve 67 % to 83 % self-consistency and 43 % consensus in expert taxonomic labelling tasks, with those routinely engaged in particular discriminations returning accuracies in the range of 84 % to 95 % (Culverhouse et al., 2003).

It is surmised that, in addition to ambiguity over the content, there is an ergonomic factor that contributes to classification bias. In any general image classifications (e.g. car vs. house or tree vs. dog), common knowledge, massive samples of people doing the training, and clear-cut distinctions between the objects make it easier (although not completely) to avoid subjective bias or even prejudice. But when the classification is being done by a small number of experts with built-in knowledge and subject background, then the training set can readily become subject to inadvertent bias. However, as a specialist field, there may be no choice. The general public may not be able to know the difference between auroral types (at least not without some training, itself subject to interpretation).

Four key human traits that affect classification performance are (a) a short-term memory limit of five to nine items, (b) boredom and fatigue, (c) recency effects whereby a new classification is biased toward the set of most recently used labels, and (d) positivity bias, wherein identification is biased by prior expectations (Culverhouse, 2007). Ambient noise, high ambient temperature, difficulty of discerning auroral features, and lack of sleep decrease performance. Additionally, attention should be paid to error analysis and associated quality metrics to weight not just algorithms, but also human-based classification according to performance (Zhu et al., 2014).

As a recommendation for future studies, classifying images in random and chronological order could be supplemented by a classification in reverse chronological order, with an examination of the results for potential hysteresis in the category selected. Specifically designed experiments could also be devised to test different biases in isolation. Ideally, these would be more generic, applying to machine-vision training more generally rather than the specific aurora case presented here.

5 Conclusions

Ergonomics refers to the design factors intended to improve productivity by reducing the fatigue and discomfort of the user. As part of the ongoing study, the trade-off between user fatigue and scientific bias is considered. When considering the training of a classification scheme, it is important to reconcile the aspects of the task which cause scientific bias but which improve overall efficiency. Given the nature of large classification programmes, removing sources of repetitive and cognitive strain not only serve to improve the working condition of the user, but also assist in ensuring that no work environment bias is injected into training datasets that are later used for classification. The items presented in Table 2 serve as a checklist for researchers who are working in machine learning.

This is of particular use for specialist fields (such as auroral research), in which it is necessary to use a small number

of experts to train algorithms. Consensus on any given classification is important in reducing errors in the training set, yet it is typical for experts to operate in very small teams or even alone. Addressing these issues will help future studies find a balance between the statistical effectiveness of large samples and the potential for scientific bias which may result from inappropriate ergonomic design that facilitates large sample classifications. This is particularly important for auroral research, in which the application of machine learning is relatively new, and there is much potential for misguided research on the grounds of biased input data.

Data availability. Original source images are publicly available from the Swedish Institute for Space Physics archive at http://www2.irf.se/allsky/data.html (Brändström, 2020), and the processed image dataset and code used in this paper are available at https://doi.org/10.18710/SSA38J (Kvammen et al., 2020b). However, users are obliged to contact the Kiruna Atmospheric and Geophysical Observatory before usage of the images. The classification from these images by the authors has been submitted as a Supplement with this paper.

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/gi-9-267-2020-supplement.

Author contributions. DM and AK jointly established the classification definitions. DM and AK independently developed their classification tools and applied them to the dataset. The text and figures for this paper were produced by DM, with comments and editing by AK.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors wish to thank Kristoffer Wickstrøm, Noora Partamies, Charlie Negri, Tuula Paavilainen and Alvin Panther for their contributions to this work. In addition, the authors would like to thank Urban Brändström and the Swedish Institute of Space Physics for providing the original auroral image data. Andreas Kvammen is supported by the Tromsø Research Foundation.

Financial support. Andreas Kvammen is supported by the Tromsø Research Foundation.

Review statement. This paper was edited by Flavia Tauro and reviewed by two anonymous referees.

References

Brändström, U.: Kiruna All-Sky Camera, Swedish Institute of Space Physics, available at: http://www2.irf.se/allsky/data.html, last access: 7 July 2020.

Culverhouse, P. F.: Human and machine factors in algae monitoring performance, Ecol. Inform., 2, 361–366, https://doi.org/10.1016/j.ecoinf.2007.07.001, 2007.

Culverhouse, P. F., Williams, R., Reguera, B., Herry, V., and González-Gil, S.: Do experts make mistakes? A comparison of human and machine indentification of dinoflagellates, Mar. Ecol.-Prog. Ser., 247, 17–25, https://doi.org/10.3354/meps247017, 2003.

Domingos, P.: A unified bias-variance decomposition, in: ICML '00: Proceedings of the Seventeenth International Conference on Machine Learning, edited by: Langley, P., Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 231–238, 2000.

Kvammen, A., Wickstrøm, K., McKay, D., and Partamies, N.: Auroral Image Classification with Deep Neural Networks, https://doi.org/10.1002/essoar.10501968.1 online first, 2020a.

Kvammen, A., Wickstrøm, K., McKay, D., and Partamies, N.: Replication Data for: Auroral Image Classification with Deep Neural Networks, DataverseNO, https://doi.org/10.18710/SSA38J, 2020b.

Syrjäsuo, M., Donovan, E., Qin, X., and Yang, Y.: Automatic classification of auroral images in substorm studies, in: 8th International Conference on Substorms (ICS8), 309–313, available at: http://aurora.phys.ucalgary.ca/donovan/pdfs/syrjaesuo_ics8.pdf (last access: 7 July 2020), 2007.

Yang, B., Spanswick, E., Liang, J., Grono, E., and Donovan, E.: Responses of Different Types of Pulsating Aurora in Cosmic Noise Absorption, Geophys. Res. Lett., 46, 5717–5724, https://doi.org/10.1029/2019GL083289, 2019.

Yu, F., Seff, A., Zhang, Y., Song, S., Funkhouser, T., and Xiao, J.: LSUN: Construction of a Large-scale Image Dataset using Deep Learning with Humans in the Loop, available at: https://arxiv.org/pdf/1506.03365.pdf (last access: 7 July 2020), 2015.

Zhu, T., Johnson, A. E. W., Behar, J., and Clifford, G. D.: Crowd-Sourced Annotation of ECG Signals Using Contextual Information, Ann. Biomed. Eng., 42, 871–884, https://doi.org/10.1007/s10439-013-0964-6, 2014.

PAPER III:
Lumikot: Fast Auroral Transients During the Growth Phase of Substorms

McKay, D., Paavilainen, T., Gustavsson, B., Kvammen, A., and Partamies, N.: Lumikot: Fast auroral transients during the growth phase of substorms, Geophysical Research Letters, 46, 7214–7221, doi: https://doi.org/10.1029/2019GL082985, 2019

Geophysical Research Letters

RESEARCH LETTER
10.1029/2019GL082985

Lumikot: Fast Auroral Transients During the Growth Phase of Substorms

D. McKay[1,2], T. Paavilainen[3], B. Gustavsson[2], A. Kvammen[2], and N. Partamies[4,5]

[1]NORCE Norwegian Research Centre AS, Tromsø, Norway, [2]Department of Physics and Technology,UiT-The Arctic University of Norway, Tromsø, Norway, [3]Teaching and Learning Services,University of Helsinki, Helsinki, Finland, [4]Department of Arctic Geophysics,The University Centre in Svalbard, Longyearbyen, Norway, [5]Birkeland Centre for Space Science, Bergen, Norway

Key Points:
- Small, transient auroral features (Lumikot) have been observed in the growth phase of substorms
- They move bidirectionally along the boundary between the bright arc and the diffuse aurora
- They map magnetically to the boundary region between the plasmasheet and outer radiation belt

Supporting Information:
- Supporting Information S1

Correspondence to:
D. McKay,
derek.mckay@norut.no

Citation:
McKay, D., Paavilainen, T., Gustavsson, B., Kvammen, A., & Partamies, N. (2019). Lumikot: Fast auroral transients during the growth phase of substorms. *Geophysical Research Letters*, *46*, 7214–7221. https://doi.org/10.1029/2019GL082985

Received 23 MAR 2019
Accepted 20 JUN 2019
Accepted article online 25 JUN 2019
Published online 9 JUL 2019

Abstract The development of a magnetospheric substorm may be classified into three phases: growth, expansion, and recovery. The growth phase is important as it includes processes that lead to the expansion. In a recent growth-phase study, a type of fast discrete auroral transient phenomena—referred to as Lumikot—were observed. The Lumikot are several kilometers across and move in the high-energy precipitation region, parallel to the main growth-phase arc, with both east-west and west-east directions of travel during the same event. Their apparent transverse movement and quasi-stable intensity make them distinct from cooccurring optical pulsating aurorae. Comparison to other studies show that they occur in the cosmic noise absorption region and it is likely that the Lumikot are colocated with high-energy particle populations on the boundary between the outer radiation belt and the plasmasheet.

1. Introduction

An auroral substorm takes place in a se quence of three phases: the growth, expansion, and recovery phase. The initial stage of the substorm is the growth phase, lasting 1–2 hr (McPherron, 1970). During this time, a narrow, east-west aligned growth-phase arc drifts equatorward and brightens prior to the poleward expansion at the substorm onset. The expansion phase is accompanied by an increase in high-energy particle precipitation (Wing et al., 2013), which originates from the plasma sheet (Wing et al., 2013). Lower-energy electrons (<10 keV) cause the optical emission, while higher-energy electrons penetrate deeper in the atmosphere and can be detected via cosmic (radio) noise absorption (CNA).

The main growth-phase arc is discrete aurora. However, a broad region of diffuse emission often exists on the equatorward side of this arc, shown to be proton aurora (Ono et al., 1987). The relationship between the discrete and diffuse emission was studied by McKay et al. (2018). They found an arc of CNA equatorward of the optical emission and contained within diffuse aurora. The offset between the growth phase and the CNA arc is nearly uniform during the equatorward progression of the growth-phase arc. McKay et al. (2018) also observed optical pulsating aurora between the diffuse emission and the main growth-phase arc, colocated with CNA. In the same study, fast-moving auroral patches were seen to traverse the sky, in the same region where the optical pulsations later appeared. These patches are referred to as *Lumikot* (singular: *Lumikko*, from the Finnish word for weasel).

This paper characterizes the Lumikot precipitation and their appearance. In particular, the particle precipitation source region is inferred by estimating the altitude and position of the Lumikot with respect to the growth phase and CNA arc. Furthermore, the plasmasheet projection of the Lumikot is approximated by employing the Tsyganenko T96 magnetospheric model. The association between the Lumikot and the optical aurora pulsations is discussed as well.

2. Instrumentation

For this study, two instruments of the Finnish Meteorological Institute's All-Sky Camera (FMI-ASC) network were used, located at Kilpisjärvi (20.78°E, 69.05°N, elevation 480 m) and Abisko (18.82°E, 68.36°N, elevation 370 m). The FMI-ASCs (Sangalli et al., 2011) use electron multiplication CCDs and filter wheels for the main auroral emission lines at 427.8, 557.7, and 630.0 nm. A standard imaging mode of 20 images per minute also includes images using background filters, dark frames, and panchromatic images. The cadence of any single wavelength is irregular. Exposure times for 427.8- and 557.7-nm images are 1.2 and 0.8 s,

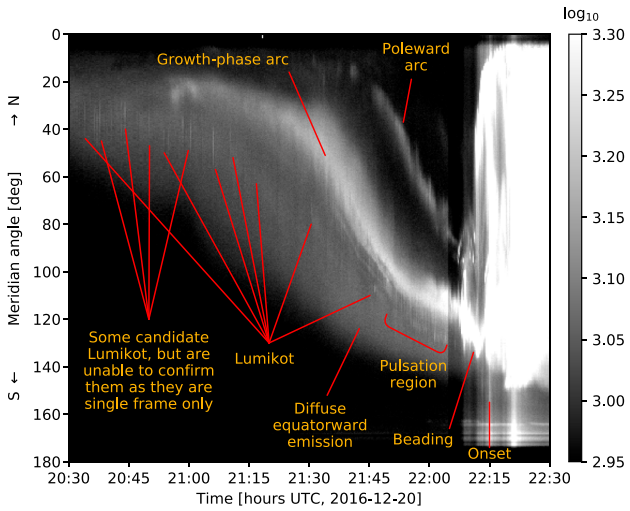


Figure 1. Keogram from the Kilpisjärvi all-sky camera showing the features of the substorm and the location of the Lumikot and Lumikot candidates therein.

respectively. Displayed grayscale levels for all images are in arbitrary analog-to-digital units (ADU) but are linear to the received luminance. The cameras use fish-eye optics to generate all-sky images with 512 × 512 pixels. At auroral altitudes of 110 km, this results in a near-zenith resolution of less than 1 km per pixel. The alignment of the FMI-ASCs at Kilpisjärvi and Abisko, respectively, is 7.53° and 6.50° east of geographic north. Magnetic north is ≈9.5° east of north for the observations in this paper. The camera lens' projection is an equidistant one (linear to pixels), and after geometric calibration the pointing error is less than 0.25° for all pixels.

3. Observations and Results

Details about the geomagnetic conditions of the substorm from 20 December 2016 (commencing at 20 UT, and continuing to 21 December 2016 03 UT) can be found in McKay et al. (2018). This isolated event was preceded by a quiet-time period of 20 hr. The planetary K index was $K_p = 1$ and the solar wind speed from 18:00–21:00 UT was slow ($v = 367 \pm 5$ km/s). The interplanetary magnetic field was steadily negative with $B_z = -3.2 \pm 0.6$ nT during the entire growth phase. Although these parameters describe the overall activity level, the magnetic field measurements (even the local ones) do not capture the disturbances related to the small-scale transient auroral signatures.

A broad diffuse arc formed over Kilpisjärvi at 19:22 UT with a bright growth-phase arc appearing at 20:58 UT on the northern edge of the diffuse emission. The arc system drifted equatorward at the speed of 71±7 m/s. During the evolution of the growth phase, small patches of aurora (Lumikot) were seen just south of the growth-phase arc, rapidly translating parallel to the arc across the field of view. The classification criteria is that the patch must be 3σ, finite in extent, moving laterally, and seen in three consecutive frames (in order to confirm this movement). This region of the Lumikot corresponds to high-energy precipitation, as indicated by CNA measured by the Kilpisjärvi Atmospheric Imaging Receiver Array (KAIRA; Figures 2 and 3 in McKay et al., 2018).

Multiple Lumikot were observed between 19:25 UT (when the cloud cover first cleared) and 22:10 UT (onset) within the CNA region, shown in Figure 1.

Both east-to-west and west-to-east directions of travel were observed. Two examples, #1 and #2, are shown in Figures 2 and 3, respectively. In each figure a sample all-sky frame is shown at the top. A region around each Lumikko, marked with a red box, has been extracted from each individual frame. These are presented below in chronological order, with the time, filter wavelength, and exposure time included. Because of the imaging sequence, the time between slices is nonuniform.

In addition to the 20 December 2016 event, six other substorms used in the previous study were investigated and Lumikot were found in two of them: 8 February 2015 and 18 February 2017. A total of 20 E→W Lumikot and 17 W→E Lumikot were detected, which represents a lower limit of the occurrence rate. Lumikot-like features were seen in many single frames, but it was not possible to discern if they travel too fast for the cadence time or if they appear and disappear without positional movement. These cases were discarded.

3.1. Altitude Determination

The simultaneous imaging and the overlapping fields of view of the two cameras permitted altitude determination of the Lumikot. A background subtraction was needed to isolate the Lumikko emission. To estimate the background for each Lumikko, a four-step procedure was used: (1) filtering the image using a 3 × 3 median filter; (2) manually defining an area that encloses the Lumikko; (3) cutting out the enclosed area; (4) performing a 2-D interpolation of the empty area. The resulting image is essentially identical to the original image but with the Lumikko emission removed. The Lumikko was thereafter isolated by subtracting the interpolated background image from the original image. This procedure is only effective if the Lumikko is positioned on a smooth background, such as the diffuse aurora is in this case (Figures 2 and 3).

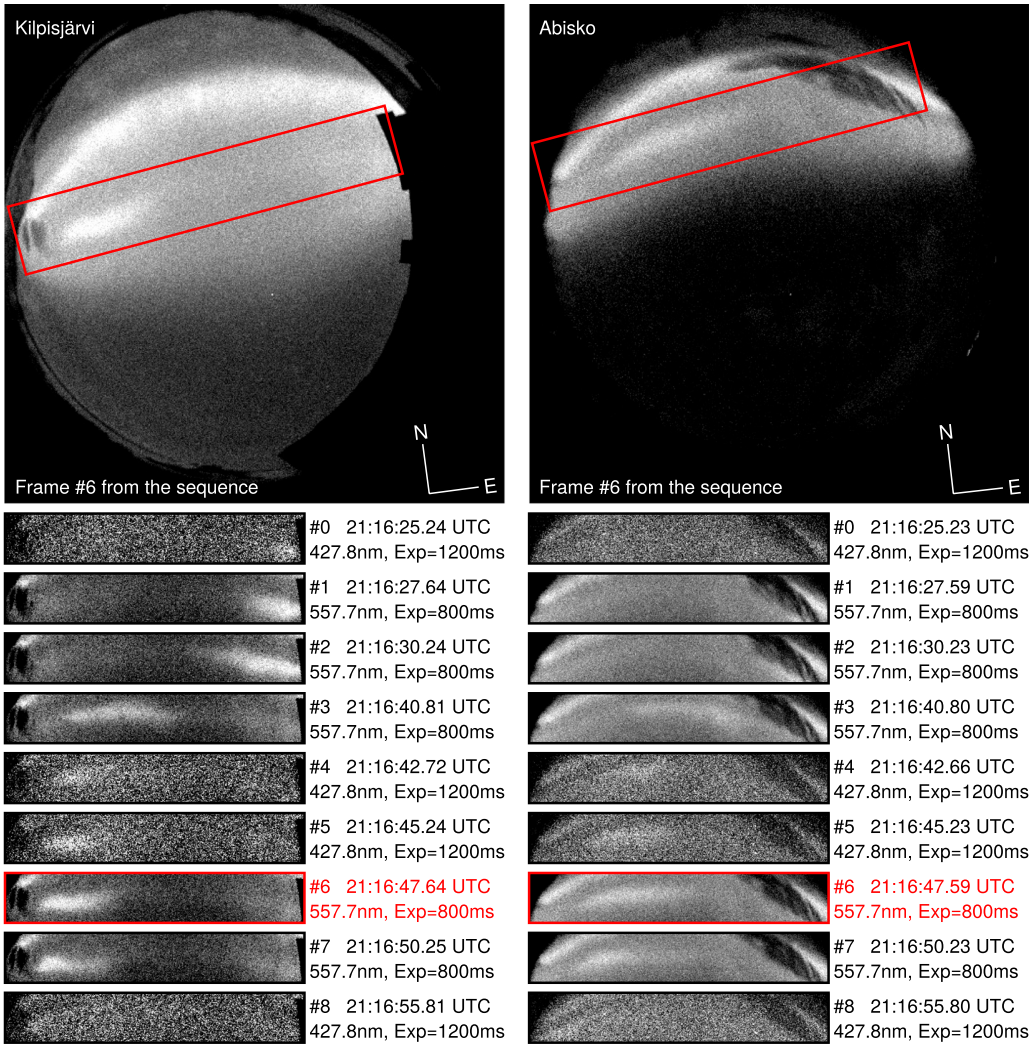


Figure 2. Example Lumikko #1, for the (left) Kilpisjärvi and (right) Abisko all-sky cameras, with apparent east-to-west motion. The top panel in each column is an all-sky image. Individual, time-ordered slices at green and blue wavelengths are shown below.

The coordinates of each Lumikko were determined by searching for an altitude, z, for which the Lumikko projections from Abisko and Kilpisjärvi were overlapping and thus minimizing the least squares error function, $E(z)$.

$$E(z) = \arg\min_{z} \left(\frac{1}{\alpha(z)} \sum_{i,j} [P_A(z,(i,j)) - P_K(z,(i,j))]^2 \right) \quad (1)$$

where $P_A(z,(i,j))$ and $P_K(z,(i,j))$ are the normalized Lumikko projections to altitude z from the background-subtracted images at Abisko and Kilpisjärvi; (i,j) are the image pixel coordinates in longitude and latitude; $\alpha(z)$ is an altitude-dependent normalization factor, proportional to the total intensity of the projected Lumikko.

The projection of the images to the longitude-latitude shell at the altitude z relies on precise geometric calibration of the camera fields of view. This is achieved by identifying stars in the image and fitting camera parameters (Gustavsson et al., 2008). Typically, detection of 200 stars results in 1σ errors of approximately ± 0.2 pixels.

The geographic coordinates of the Lumikot midpoints were similar in the Abisko and Kilpisjärvi projections when $E(z)$ was near the minimum. However, the shapes of the Lumikot projections were not identical,

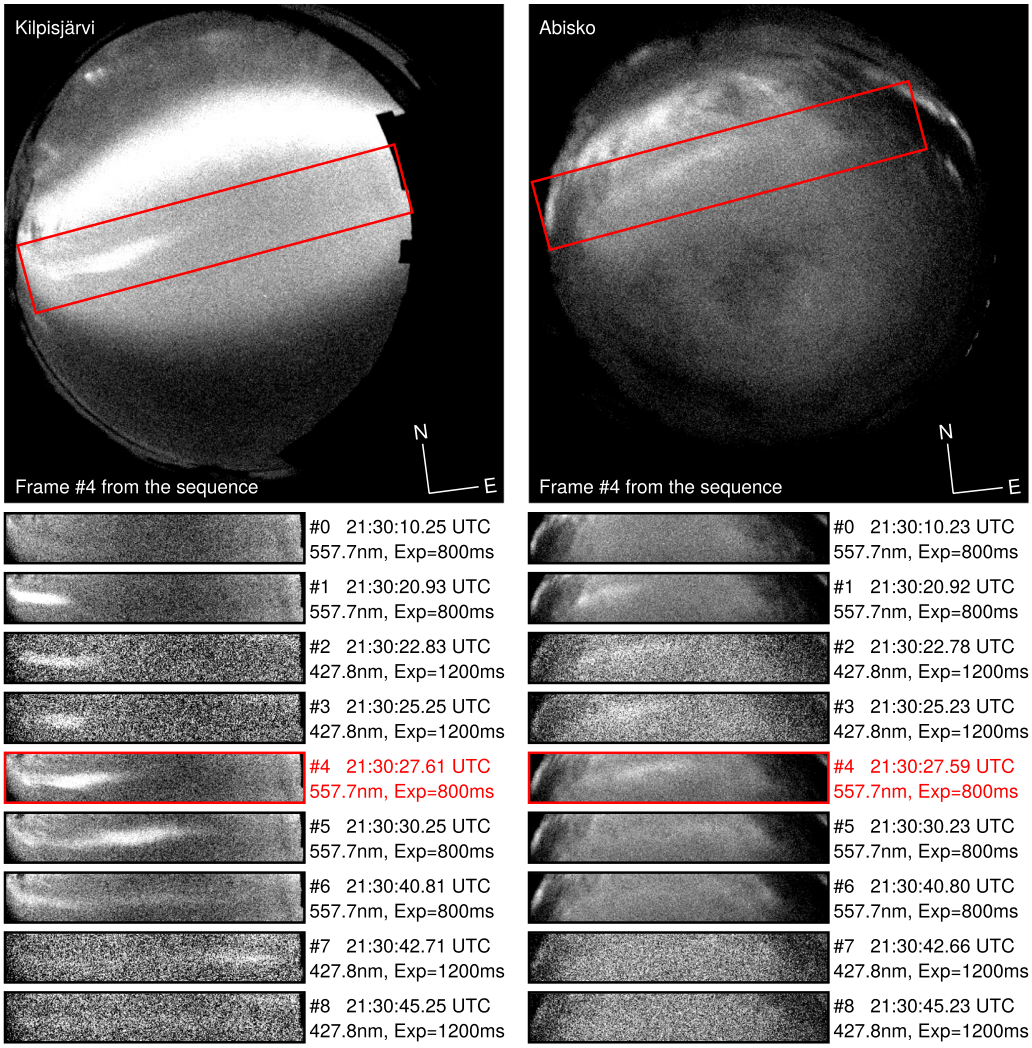


Figure 3. Example Lumikko #2, with apparent west-to-east motion.

as each Lumikko is not a horizontal emission layer, but a 3-D volume. The layer approximation results in a latitude difference between the projections. Nevertheless, the shell projection method is preferred as it utilizes information from the entire Lumikko, rather than one triangulation point. The disadvantage of the method is that the error function ($E(z)$) does not provide direct estimates of the coordinate deviation around the best-fit altitude. The projections corresponding to the altitudes with errors above 20% of the best-fit value show clear discrepancy, both below and above the best-fit altitude. A conservative confidence interval of altitudes within 20% of the best-fit error value was therefore used. The Lumikot image pairs in

Table 1
The Best-Fit Coordinate Values for the Sample Lumikot With Error Ranges

Event	#1a	#1b	#2
time (UTC)	21:16:48	21:16:51	21:30:28
x [km]	-170 [−166, −176]	−201 [−194, −207]	−132 [−133, -137]
y [km]	31 [28, 35]	30 [27, 33]	−9 [−11, −8]
z [km]	118 [114, 124]	120 [115, 125]	115 [110, 121]

Note. The xy coordinate ranges correspond to the z coordinate ranges and are therefore not necessarily ordered from the lowest to highest value.

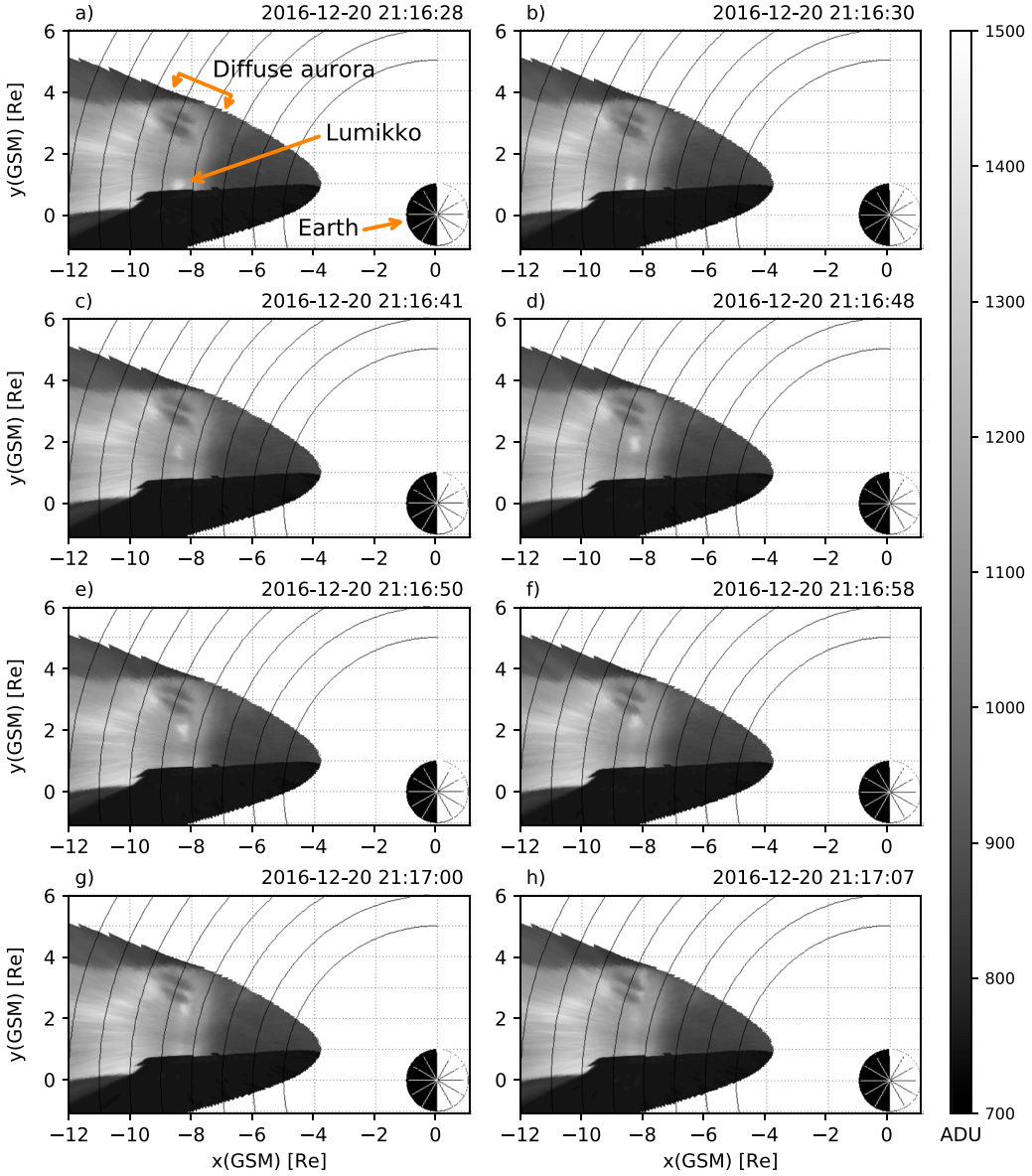


Figure 4. (a–h) The Kilpisjärvi images (projected to the plasmasheet) place the Lumikot at $x \approx -8\,R_e, y \approx 2.2\,R_e$. The growth-phase arc maps to the distance of approximately $10\,R_e$ from the Earth. The dark structure at $y < 1\,R_e$ is not an auroral feature but a structure at the Kilpisjärvi site.

Figure 2 at 21:16:48 UTC (#1a) and at 21:16:51 UTC (#1b) and in Figure 3 at 21:30:28 UTC were used in the altitude determination. The altitude and horizontal coordinate results are summarized in Table 1. The Lumikot *xy* coordinates in Table 1 are transferred from degrees in longitude and latitude to kilometers north and east of Kilpisjärvi so that the velocity of Lumikko #1 can then be estimated in appropriate units. The Lumikko *xy* midpoint is defined as the mean center of two fitted asymmetrical 2-D Gaussians, where one 2-D Gaussian is fitted to the Abisko projection and the other is fitted to the Kilpisjärvi projection.

Lumikot take approximately 30–45 s to cross the sky. The apparent velocity of example #1 based on the coordinates of #1a and #1b becomes $\bar{v}_{(x,y,z)} \approx [-12, 0, 1]$ km/s $\pm$ [3, 2, 2] km/s; that is, westward, with no appreciable change in the latitudes or heights (Table 1).

3.2. Projection of Lumikot Onto the Plasmasheet

The example Lumikko from Figure 2 was projected to the location along the magnetic field lines, which is the farthest distance from Earth, which should be the estimated location of the cross-tail current sheet, or

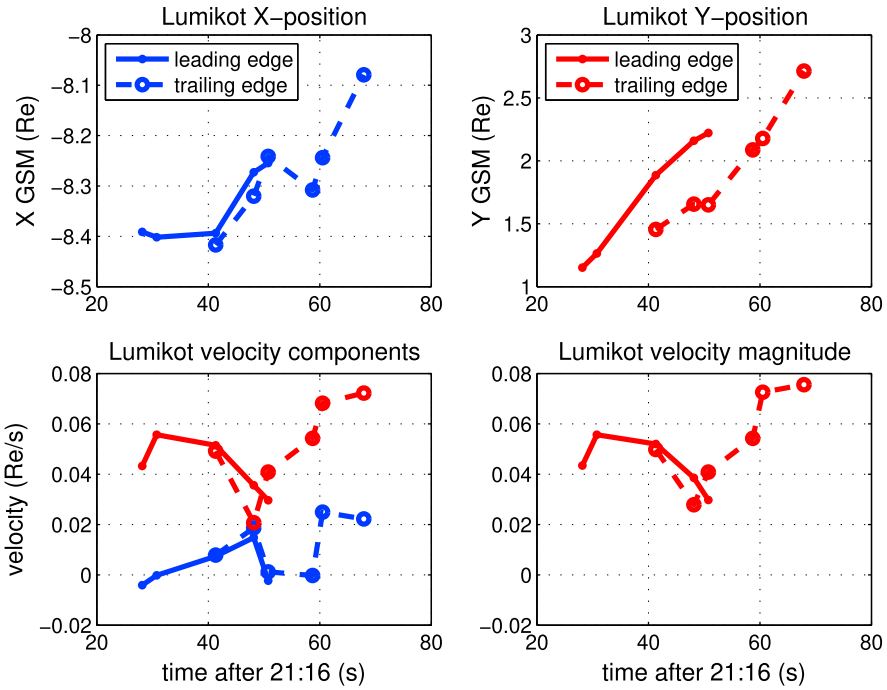


Figure 5. The magnetospheric X and Y coordinates of Lumikot at 21:16–21:17 UT are shown in the upper panels. The X and Y velocity components are plotted in the bottom left while the speed is shown in the bottom right panel.

plasmasheet, by first projecting the images to an initial altitude of 112 km. Magnetic coordinates were calculated using `aacgmv2` (Shepherd, 2014) Python libraries for the epoch 20 December 2016, giving Kilpisjärvi = 102.39°E, 66.06°N and Abisko = 100.35°E, 65.42°N. The magnetic field lines were then traced from these locations into the magnetosphere using the Tsyganenko T96 magnetospheric model (Tsyganenko & Stern, 1996) with the solar wind magnetic field GSM components, B_y and B_z of 7.96 and −1.84 nT, solar wind density of 12.05 cm^{-3}, and solar wind velocity of 368.5 km/s. As fluctuations in the solar wind were minor in this case, and it was still in the early stage of the growth phase, Tsyganenko produces representative mapping to the inner magnetosphere.

The mapped position of the example Lumikko is just tailward of X_{GSM} 8 R_e moving from Y_{GSM} 1.1 to 2.2 R_e over a ≈ 10-s period as illustrated in Figure 4. This gives speeds ranging from 1.8 to 4.8 × 10^5 m/s with an average of 3.4 × 10^5 m/s (Figure 5). This is comparable to a typical Alfvén speed in the inner plasmasheet ($v_A \approx 5.5 \times 10^5$ m/s).

4. Discussion

To the best of our knowledge, this is the first report of randomly bidirectional and convection-independent auroral transients in the substorm growth phase. Lumikot associate with quasi-stable growth-phase arcs with a low occurrence rate; probably overlooked in the less extensive growth-phase studies since the work of Lessard et al. (2007). However, the event in this paper was not unique as other examples were also found. Presumably, as faster cadence times have been implemented for auroral imaging, the prevalence of Lumikot becomes more apparent. One observational limitation may be the narrow latitude band.

The sizes of individual Lumikot are in the range of a few kilometers in the north-south and 10–50 km in the east-west direction, that is, meridionally thinner than beads and azimuthally of the order of bead size (Motoba et al., 2012). The z coordinate estimates of 115–120 km agree with the pulsating aurora heights reported by Partamies et al. (2017), where 557.7-nm median height was ≈118 km before the pulsating started and dropped to ≈108 km during the pulsating aurora. This places the enhancement of Lumikot precipitation energy between those of average recovery-phase aurora and pulsating aurora, embedded in the region of CNA.

The Lumikot includes enhanced emission embedded in a region of high-energy precipitation, as indicated by CNA of ≈0.3 dB measured by KAIRA. No variation in CNA (to a limit of ±0.05 dB) was measured when the Lumikot passed, which may be due to the limited spatial resolution of KAIRA (15° full width at half maximum at 38.1 MHz).

As the magnetotail stretches during the growth phase, the auroral arc moves equatorward until the buildup of energy is released at the substorm onset (Lessard et al., 2007). The equatorward arc maps to the inner edge of the plasmasheet where significant gradients in plasma properties may exist (Fukunishi, 1975). However, an inner magnetospheric instability is unlikely to be responsible for bidirectional propagation. Chorus waves have been shown to be responsible for the precipitation during pulsating aurora (Kasahara et al., 2018). While some pulsating patches have been observed to follow the convection (Yang et al., 2017), others, such as the amorphous pulsating aurora (Grono et al., 2017) do not. Lumikot neither follow the convection speed nor the consistency in the direction.

The propagation speed of our sample Lumikot (Example #1) was found to be ≈12 km/s. This structure was visible in three consecutive images. A majority of other Lumikot were also observed in 3–4 consecutive images, suggesting that our sample speed estimate is representative for the studied cases. The data do not allow the assertion of a preferred direction of motion. So, unlike other previously reported auroral structures, such as streaming auroral pulsating patches (Yamamoto & Oguti, 1982), azimuthally propagating extensions of north-south aligned arcs (Nishimura et al., 2010), azimuthally moving auroral enhancements (Lyons et al., 2015) or auroral brightening called auroral horn (Koskinen et al., 1990), Lumikot do not show an observable preference in their propagation direction.

The Lumikot phenomenon defies theoretical models which would generate bidirectional propagation of small-scale aurora.

Although the projection of the Lumikot from the ASC images to the *E* region is quantitatively accurate, the mapping to the plasmasheet should only be seen as qualitatively representative since the Tsyganenko T96 model does not take the dynamics of the magnetosphere and magnetosphere-solar-wind interaction into account. For substorm conditions, this can mean discrepancies between the Tsyganenko model and reality.

According to Nishimura et al. (2016), auroral beads propagate both eastward and westward. Their appearance is fundamentally different from Lumikot as they grow periodically within the growth-phase arc during the minutes prior to the substorm onset and initiate the brightening before the breakup. The relation of Lumikot to the substorm onset instability is much less obvious. However, the beading shows that there are mechanisms producing both propagation directions for small-scale aurora during the growth phase.

Optical pulsating aurora are seen during the growth phase of the substorm both before and after the Lumikot (McKay et al., 2018), suggesting that Lumikot are not "prototype" pulsating aurora. It is not clear whether Lumikot exist in the recovery phase of the substorm, as the complexity of the postbreakup aurora preclude their reliable detection. More detailed relationship between pulsating aurora and Lumikot should be investigated in the future. With the advances in camera sensitivity and sampling rate it has become possible to conduct more comprehensive surveys of transient auroral phenomena. Additionally, new techniques being developed for recovery-phase pulsating aurora (Grono et al., 2017) could be applied to growth-phase research.

5. Conclusion

Fast auroral transients, referred to as Lumikot, have been observed during the substorm growth phases. They appear in the region between diffuse aurora and the optically bright growth-phase arc, with pulsating aurora and high-energy particle precipitation leading to CNA. The Lumikot are too small spatially and temporally to greatly modify CNA regardless whether the Lumikot precipitation is energetic enough or not. This region corresponds to CNA and pulsating aurora and, thus, high-energy particle precipitation. Tracing the magnetic field lines maps this band to the border between the outer radiation belt and the plasmasheet.

The Lumikot are several kilometers across and travel east-west parallel to the main growth-phase arc. Both directions of travel have been observed during the same event. They are significant because such bidirectional transients prior to the expansive onset of the auroral breakup challenges the current understanding of substorm instabilities. The magnetospheric dynamics causing the Lumikot are a novel hitherto unknown

phenomena that can now be used to help understand preonset magnetospheric processes. Lumikot are significant because such bidirectional transients prior to the onset of the auroral breakup may provide important insights to the processes leading to the onset.

Acknowledgments
We thank the Finnish Meteorological Institute for their operation of the FMI-ASC network. FMI-ASC data are available as keograms at http://www.space.fmi.fi/MIRACLE/ASC/ and http://gaia-vxo.org. Full resolution images can be requested from Kirsti Kauristie (`kirsti.kauristie@fmi.fi`). KAIRA was funded by the University of Oulu and FP7 European Regional Development Funds and is operated by Sodankylä Geophysical Observatory and the University of Tromsø, and data are available on request from the PI at http://www.sgo.fi/KAIRA/. Magnetic coordinates and directions were determined using the `aacgmv2` Python libraries and the U.S. Department of Commerce, NOAA National Geophysical Data Center online services, https://www.ngdc.noaa.gov/geomag-web/. Ancillary data are from the Japanese World Data Centre (http://wdc.kugi.kyoto-u.ac.jp/). The authors thank T. Antonsen and J. Vierinen for useful discussions as well as the suggestions of the anonymous reviewers. D. McKay was supported by the Tromsø Research Foundation. N. Partamies is supported by the Research Council of Norway under CoE Contract 223252.

References

Fukunishi, H. (1975). Dynamic relationship between proton and electron auroral substorms. *Journal of Geophysical Research*, *80*, 553–574. https://doi.org/10.1029/JA080i004p00553

Grono, E., Donovan, E., & Murphy, K. R. (2017). Tracking patchy pulsating aurora through all-sky images. *Annales Geophysicae*, *35*, 777–784. https://doi.org/10.5194/angeo-35-777-2017

Gustavsson, B., Kosch, M., Wong, A., Pedersen, T., Heinselman, C., Mutiso, C., et al. (2008). First estimates of volume distribution of HF-pump enhanced emissions at 6300 and 5577 Å: A comparison between observations and theory. *Annales Geophysicae*, *26*, 3999–4012. https://doi.org/10.5194/angeo-26-3999-2008

Kasahara, S., Miyoshi, Y., Yokota, S., Mitani, T., Kasahara, Y., Matsuda, S., et al. (2018). Pulsating aurora from electron scattering by chorus waves. *Nature*, *554*, 337–340. https://doi.org/10.1038/nature25505

Koskinen, H. E. J., Pulkkinen, T. I., & Pellinen, R. J. (1990). Mapping of the auroral horn into the magnetotail. *Planetary and Space Science*, *38*, 1179–1186. https://doi.org/10.1016/0032-0633(90)90025-L

Lessard, M. R., Lotko, W., LaBelle, J., Peria, W., Carlson, C. W., Creutzberg, F., & Wallis, D. D. (2007). Ground and satellite observations of the evolution of growth phase auroral arcs. *Journal of Geophysical Research*, *112*, A09304. https://doi.org/10.1029/2006JA011794

Lyons, L. R., Nishimura, Y., Gallardo-Lacourt, B., Nicolls, M. J., Chen, S., Hampton, D. L., et al. (2015). Azimuthal flow bursts in the inner plasma sheet and possible connection with saps and plasma sheet earthward flow bursts. *Journal of Geophysical Research: Space Physics*, *120*, 5009–5021. https://doi.org/10.1002/2015JA021023

McKay, D., Partamies, N., & Vierinen, J. (2018). Pulsating aurora and cosmic noise absorption associated with growth-phase arcs. *Annales Geophysicae*, *36*(1), 59–69. https://doi.org/10.5194/angeo-36-59-2018

McPherron, R. L. (1970). Growth phase of magnetospheric substorms. *Journal of Geophysical Research*, *75*, 5592. https://doi.org/10.1029/JA075i028p05592

Motoba, T., Hosokawa, K., Kadokura, A., & Sato, N. (2012). Magnetic conjugacy of northern and southern auroral beads. *Geophysical Research Letters*, *39*, L08108. https://doi.org/10.1029/2012GL051599

Nishimura, Y., Lyons, L., Zou, S., Angelopoulos, V., & Mende, S. (2010). Substorm triggering by new plasma intrusion: THEMIS all-sky imager observations. *Journal of Geophysical Research*, *115*, A07222. https://doi.org/10.1029/2009JA015166

Nishimura, Y., Yang, J., Pritchett, P. L., Coroniti, F. V., Donovan, E. F., Lyons, L. R., et al. (2016). Statistical properties of substorm auroral onset beads/rays. *Journal of Geophysical Research: Space Physics*, *121*, 8661–8676. https://doi.org/10.1002/2016JA022801

Ono, T., Hirasawa, T., & Meng, C. I. (1987). Proton auroras observed at the equatorward edge of the duskside auroral oval. *Geophysical Research Letters*, *14*, 660–663. https://doi.org/10.1029/GL014i006p00660

Partamies, N., Whiter, D., Kadokura, A., Kauristie, K., Nesse Tyssøy, H., Massetti, S., et al. (2017). Occurrence and average behavior of pulsating aurora. *Journal of Geophysical Research: Space Physics*, *122*, 5606–5618. https://doi.org/10.1002/2017JA024039

Sangalli, L., Partamies, N., Syrjäsuo, M., Enell, C.-F., Kauristie, K., & Mläkinen, S. (2011). Performance study of the new EMCCD-based all-sky cameras for auroral imaging. *International Journal of Remote Sensing*, *32*, 2987–3003. https://doi.org/10.1002/2017JA024039

Shepherd, S. G. (2014). Altitude-adjusted corrected geomagnetic coordinates: Definition and functional approximations. *Journal of Geophysical Research: Space Physics*, *119*, 7501–7521. https://doi.org/10.1002/2014JA020264

Tsyganenko, N. A., & Stern, D. P. (1996). Modeling the global magnetic field of the large-scale Birkeland current systems. *Journal of Geophysical Research*, *101*, 27,187–27,198. https://doi.org/10.1029/96JA02735

Wing, S., Gkioulidou, M., Johnson, J. R., Newell, P. T., & Wang, C.-P. (2013). Auroral particle precipitation characterized by the substorm cycle. *Journal of Geophysical Research: Space Physics*, *118*, 1022–1039. https://doi.org/10.1002/jgra.50160

Yamamoto, T., & Oguti, T. (1982). Recurrent fast motions of pulsating auroral patches. I—A case study on optical and quantitative characteristics during a slightly active period. *Journal of Geophysical Research*, *87*, 7603–7614. https://doi.org/10.1029/JA087iA09p07603

Yang, B., Donovan, E., Liang, J., & Spanswick, E. (2017). A statistical study of the motion of pulsating aurora patches: Using the THEMIS all-sky imager. *Annales Geophysicae*, *35*, 217–225. https://doi.org/10.5194/angeo-35-217-2017

PAPER IV:
The 3-D Distribution of Artificial Aurora Induced by HF Radio Waves in the Ionosphere

Kvammen, A., Gustavsson, B., Sergienko, T., Brändström, U., Rietveld, M., Rexer, T., and Vierinen, J.: The 3-D Distribution of Artificial Aurora Induced by HF Radio Waves in the Ionosphere, Journal of Geophysical Research: Space Physics, 124, 2992–3006, doi: https://doi.org/10.1029/2018JA025988, 2019

JGR Space Physics

RESEARCH ARTICLE
10.1029/2018JA025988

The 3-D Distribution of Artificial Aurora Induced by HF Radio Waves in the Ionosphere

A. Kvammen[1], B. Gustavsson[1], T. Sergienko[2], U. Brändström[2], M. Rietveld[1,3], T. Rexer[1], and J. Vierinen[1]

[1]Department of Physics and Technology, University of Tromsø-The Arctic University of Norway, Tromsø, Norway, [2] Swedish Institute of Space Physics, Kiruna, Sweden, [3]EISCAT Scientific Association, Ramfjordbotn, Norway

Key Points:
- The first 3-D estimates of induced emission at 8,446 Å are presented along with 3-D estimates of the enhanced emission at 6,300 and at 5,577 Å
- The altitude distribution of the resulting excitation rates is inconsistent with excitation rate predictions
- We observe that the emission enhancements are strongly dependent on the pump frequency proximity to the double resonance

Correspondence to:
A. Kvammen,
andreas.kvammen@uit.no

Citation:
Kvammen, A., Gustavsson, B., Sergienko, T., Brändström, U., Rietveld, M., Rexer, T., & Vierinen, J. (2019). The 3-D distribution of artificial aurora induced by HF radio waves in the ionosphere. *Journal of Geophysical Research: Space Physics*, *124*, 2992–3006. https://doi.org/10.1029/2018JA025988

Received 10 AUG 2018
Accepted 12 MAR 2019
Accepted article online 21 MAR 2019
Published online 6 APR 2019

Abstract We present 3-D excitation rate estimates of artificial aurora in the ionospheric *F* layer, induced by high-frequency radio waves from the European Incoherent Scatter heating facility. Simultaneous imaging of the artificial aurora was done with four separate Auroral Large Imaging System stations, permitting tomography-like 3-D auroral reconstruction of the enhanced atomic oxygen emissions at 6,300, 5,577, and 8,446 Å. Inspection of the 3-D reconstructions suggests that the distribution of energized electrons is less extended in altitude than predicted by transport calculations of electrons accelerated to 2–100 eV. A possible reason for this discrepancy is that high-frequency pumping might induce an anisotropic distribution of energized electrons.

Plain Language Summary Auroral lights can be artificially generated by transmitting high-frequency radio waves with high power into the upper atmosphere. In this article, we use multiple viewpoint imaging of artificially produced aurora to estimate the 3-D distribution of the auroral lights by employing tomography-like techniques. The 3-D distribution is estimated in the red, green, and infrared auroral emission lines with wavelengths of 630.0, 557.7, and 844.6 nm, respectively. These emissions are excited by energetic electrons, which have been accelerated through interaction processes between the transmitted radio waves and plasma in the upper atmosphere, at an altitude of about 220–250 km. We observe that the estimated 3-D auroral distributions are less extended in altitude than indicated by previous theoretical work. A possible reason for this disagreement is that the radio wave-plasma interaction processes might lead to a direction dependent electron acceleration.

1. Introduction

When powerful high-frequency (HF) radio waves reach the ionosphere, several wave-plasma interactions are excited and most of the HF wave energy is dissipated by the plasma (Senior et al., 2012), inducing observable phenomena, such as electron temperature enhancements (Honary et al., 1995; Rietveld et al., 2003; Robinson, 1989), production of electron density striations (Milikh et al., 2008), artificial ionization (Bernhardt et al., 2016; Pedersen et al., 2009), stimulated electromagnetic emissions (Leyser, 2001), and enhancement of optical emissions (Brändström et al., 1999; Gustavsson et al., 2005). At auroral latitudes, it is postulated that incident ordinary mode HF radio waves excite upper-hybrid (Kosch et al., 2002), lower-hybrid (Djuth et al., 2005), and Langmuir turbulences (Djuth et al., 2004) as well as electron Bernstein waves (Stubbe et al., 1994) within magnetic field-aligned plasma striations in the ionosphere. The strong wave-plasma interaction region will have an upper border at the reflection altitude and a rough lower border at the altitude where the pump frequency is in resonance with the upper-hybrid frequency (Eliasson & Papadopoulos, 2015), typically a few kilometers below the reflection altitude for *F* region heating (Gustavsson et al., 2005).

The net result of these plasma processes is energization of electrons and increased plasma pressure within plasma striations (Gurevich & Zybin, 2006). This triggers an instability, known as the resonance instability (Gurevich, 2007), where the striations continue to expand as the plasma pressure increases, which causes self-focusing and increased HF radiation flux into the striations (Eliasson & Papadopoulos, 2015; Istomin & Leyser, 2003; Gondarenko et al., 2005). The striations will expand from a few meters to hundreds of meters during the first 10–30 s after heating onset before stabilizing, the expansion is primarily in the plane perpendicular to the magnetic field (Coster et al., 1985; Milikh et al., 2008). After that point, close to 100% of the

HF wave energy is dissipated by the plasma within the interaction region, provided that the pump power flux exceeds 30 μW/m^2 (Senior et al., 2012).

Observations of both enhanced plasma lines (Carlson et al., 1982) and enhanced optical emissions (Brändström et al., 1999) during heating can only be explained by high-energy electrons. Understanding how the electron energy distribution is modified during HF pumping is one of the central questions in ionospheric heating research. Mantas (1994) and Mantas and Carlson (1996) attempted to explain observations of enhanced emission intensities at 6,300 Å, denoted I_{6300}, during HF heating by $O(^1D)$ excitation (threshold 1.96 eV) from a purely thermal electron energy distribution. However, a thermal electron energy distribution cannot explain observations of enhanced emissions from states with higher excitation energies, such as I_{5577} from $O(^1S)$ (threshold 4.17 eV; Haslett & Megill, 1974), I_{8446} from $O(3p^3P)$ (threshold 10.99 eV; Gustavsson et al., 2005), and I_{4278} from $N_2^+(1NG)$ (threshold 18.75 eV; Holma et al., 2006). There are not enough high-energy electrons in a thermal population to induce observable enhancements at these wavelengths (Gustavsson et al., 2002). The observations are therefore in line with an accelerated, suprathermal, electron energy distribution (Bernhardt et al., 1989).

Gustavsson et al. (2005) and Gustavsson and Eliasson (2008) combined optical images of I_{6300}, I_{5577}, I_{8446}, and I_{4278} and IS radar observations to estimate the energy and altitude distribution of electrons accelerated to 2–100 eV during heating. Gustavsson and Eliasson (2008) achieved this by employing a two-stream electron transport model with isotropic electron acceleration by upper- and lower-hybrid waves within a narrow altitude range, taking electron-ion and electron-neutral collisions into account. Hysell et al. (2012) and Hysell et al. (2014) obtained similar results of the suprathermal electron energy distribution versus energy and altitude using spectrographic measurements and electron transport from the multistream SAMI2-PE model (Varney et al., 2012). These results provide a method for calculating the altitude distribution of HF-enhanced optical emissions by employing the appropriate excitation cross sections and transition probabilities. So far, there have been no attempts at comparing the resulting electron energy distributions and the corresponding excitation rates to volumetric emission rate estimates. The 3-D emission rate estimates of heating induced I_{6300} were for the first time estimated by Gustavsson et al. (2001) using the European Incoherent Scatter (EISCAT) heating facility and simultaneous multistation imaging at three Auroral Large Imaging System (ALIS) imaging stations. Gustavsson et al. (2001) achieved 3-D emission rate reconstruction by using tomography-like inversion methods. The same method was later employed by Gustavsson et al. (2008) to estimate the volume emission rates in both I_{6300} and in I_{5577} using the High-Frequency Active Auroral Research Program facility and two imaging stations. Shindin et al. (2018) estimated the 3-D emission rates of I_{6300} at midlatitudes, induced by the Sura heating facility, using two imaging stations.

In this paper, we use multistation optical observations of I_{6300}, I_{5577}, and I_{8446} to test predicted excitation-altitude profiles, calculated using the method described in Gustavsson and Eliasson (2008). This is achieved by comparing the projections of reconstructed 3-D aurora models to simultaneous multistation images. The first results of 3-D artificial auroral reconstruction from as many as four imaging stations in I_{6300} and in I_{5577} are presented along with the first published estimates of the I_{8446} volume distribution.

2. Experiment and Observations

The EISCAT Scientific Association heating facility (Rietveld et al., 2016), the EISCAT ultrahigh frequency (UHF) incoherent scatter radar (Rishbeth & Van Eyken, 1993), and ALIS (Brändström, 2003) were operating simultaneously on 16 February 2015. Heating array 3 was employed to transmit right-hand circular polarized HF waves, that is, ordinary mode waves in the ionospheric plasma, antiparallel to the magnetic field. A frequency of 6.200 MHz was used from 16:00 to 16:50 UT, that is, heating not close to a gyroharmonic resonance, and a frequency of 5.423 MHz from 16:51 UT and onward, that is, heating in proximity to the fourth gyroharmonic frequency. The HF pumping was operating in a 150-s heating on and 85-s heating off cycle modulation, making it possible to measure the sky background between the heating pulses and to estimate the decay time of the 6,300-Å emission. The growth and decay of the enhanced intensities in the 6,300-Å emission line, I_{6300}, are shown in Figure 1 for one heating cycle. The growth and decay time of I_{6300} are dependent on the effective $O(^1D)$ lifetime. The observed $O(^1D)$ lifetimes are presented in Figure 5 and are more thoroughly discussed in section 3.

The frequency-dependent effective radiated power was approximately 138.2 MW at 6.200 MHz and 115.9 MW at 5.423 MHz, corresponding to effective power fluxes of 143 and 120 μW/m^2 at the heating altitude (Rietveld et al., 1993). The modeled beam patterns are shown in Figure 2.

The EISCAT UHF incoherent scatter radar was operated with a meridional scan pattern throughout the experiment, providing plasma parameter measurements of the heated volume with approximately 3-km altitude resolution and 5-s time resolution. The electron temperature was enhanced to approximately 3300 K at the resonance height during 6.200-MHz heating and to about 2500 K when pumping at 5.423 MHz. The electron density and ion drift velocity remained stable throughout the time interval of interest, that is, no natural auroral activity and no ion upflow. The Beata pulse-coding scheme was employed, and plasma parameters were extracted from the backscattered power spectrum using the Guisdap analysis tool (Lehtinen & Huuskonen, 1996), version 8.8.

Optical imaging of the artificial aurora was made with ALIS in the emission lines at 6,300, 5,577, 8,446, and 4,278 Å; however, significant enhancements were only observed in 6,300, 5,577, and 8446 Å. The UHF electron temperature measurements and the enhanced emission intensities are presented in Figure 3. The electron temperature panel clearly shows a decrease in the pump enhanced temperatures when going from 6.200- to 5.423-MHz pumping, that is, from heating away from a gyroharmonic to heating near the fourth gyroharmonic. The scatter points in the optical intensity plots represent the mean intensity of a (11 × 11) pixel grid around the peak enhancement in the images. I_{6300} was strongest during 6.200-MHz heating; the enhancement was reduced by a factor 3 when changing pump frequency to 5.423 MHz. I_{5577} remained quite similar at 6.200- and at 5.423-MHz pumping, whereas significant I_{8446}, above a noisy background, was only detected during 5.423-MHz heating. Imaging was done at four ALIS stations, in Abisko, Kiruna, Silkimuotka, and Tjautjas. An exposure time of 6 s was chosen at all stations and for all filters, providing images approximately every 10 s. Clouds before 16.37 UT and clouds and interference after 17.00 obstructed optical observations. Tomographic reconstruction was therefore done between 16.37 and 17.00 UT. No images were taken in the 8,446-Å filter at Kiruna; hence, tomographic reconstruction of I_{8446} was done using only three imaging stations. In addition, note that the periodic electron temperature modulation, seen in the top panel of Figure 3 during heating on, is due to the scanning of the UHF radar beam through the heated volume. The UHF radar scanned over positions between 7.8°S and 15.5°S of zenith. The electron temperature enhancement peaked when the UHF beam was antiparallel to the magnetic field, although there is an asymmetry in the temperature enhancement between the bottom and the top of the meridional scan. It is not clear if there is a physical reason for the asymmetry or if it is solely a geometric effect.

The optical data acquired on 16 February 2015 were particularly suitable for 3-D reconstruction of artificial aurora for two reasons: (1) Enhanced emissions were observed from as many as four separate imaging stations, making the tomography-like inversion more reliable than in previous reconstruction attempts. (2) The artificial aurora was observed to have a simple continuous shape at all imaging stations and in all emission lines; see, for example, the 6,300-Å emission in Abisko in Figure 1. This enabled us to approximate the emission distribution by using simple Gaussian distribution functions. In contrast, a more structured artificial aurora with many small-scale striations, as seen during some High-Frequency Active Auroral Research Program experiments (e.g., Bernhardt et al., 2016; Kendall et al., 2010), or a drifting emission pattern, as seen by, for example, Grach et al. (2017), would cause the reconstruction problem to be much more complex and ambiguous.

3. Aurora Modeling

The 3-D artificial aurora reconstruction was achieved by adjusting a 3-D parameterized aurora model so that the model projections fitted the observed images. The 3-D aurora models were constructed with a two-dimensional Gaussian in the horizontal plane combined with three different profiles along the magnetic field: (1) Gaussian distribution, (2) Gustavsson-Eliasson profiles, and (3) modified Gustavsson-Eliasson profiles. The 3-D Gaussian distribution gave us an empirical model with search parameters for the peak excitation rate, I_0, the center coordinates, (x_0, y_0, z_0), and the horizontal and vertical widths (σ_{xy}, σ_z), in total of six model parameters. The Gustavsson-Eliasson profiles model, denoted the G-E model, and the modified Gustavsson-Eliasson profiles model, denoted the modified G-E model, used excitation rate profiles, calculated for a range of electron acceleration altitudes by employing the method described in Gustavsson and Eliasson (2008). This gave us a set of five model parameters: I_0, x_0, y_0, z_0, and σ_{xy}, for the physical G-E

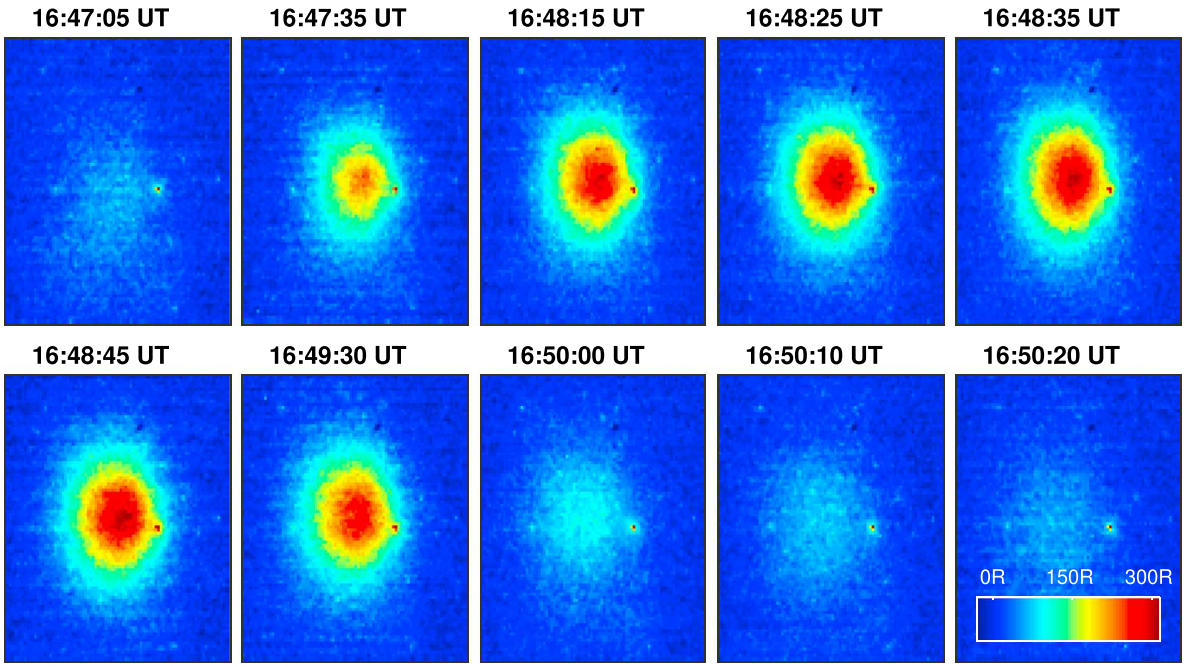


Figure 1. Images of the 6,300-Å emission in Rayleigh units [*R*] from one heating cycle as observed in Abisko. Images during heating are shown in the top row, and images after heating offset are in the bottom row. Note that the images are background reduced and flat field corrected, and the bright speckles are stars.

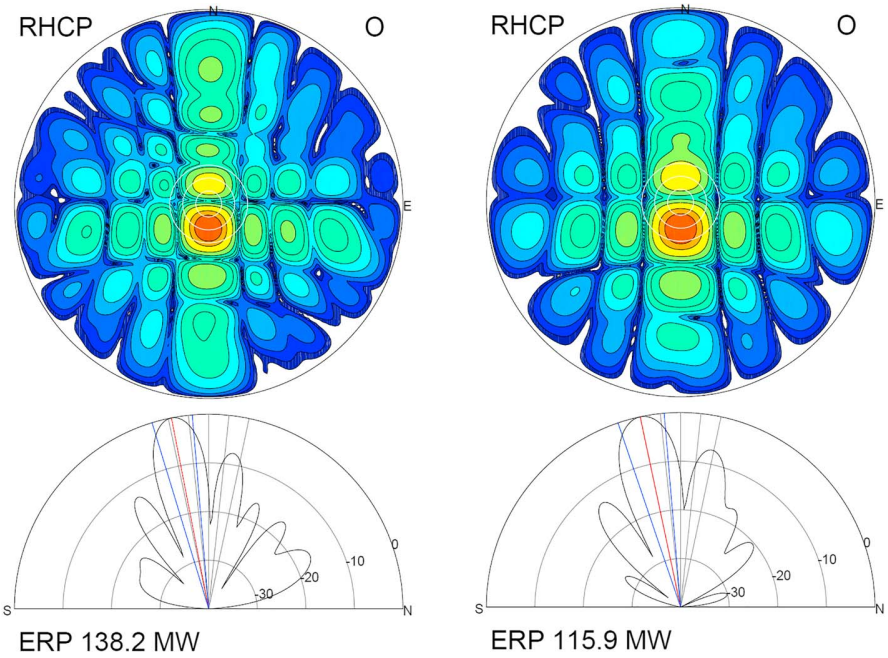


Figure 2. The figure shows the modeled beam patterns during 6.200-MHz heating to the left and during 5.423-MHz heating to the right. At the top, the 2-D radiation pattern where the white rings encircle zenith. At the bottom, the beam pattern in the meridional plane. The heating beam was pointed 12°S of zenith, approximately antiparallel to the magnetic field. RHCP = right-hand circular polarized; ERP = effective radiated power; O = ordinary mode.

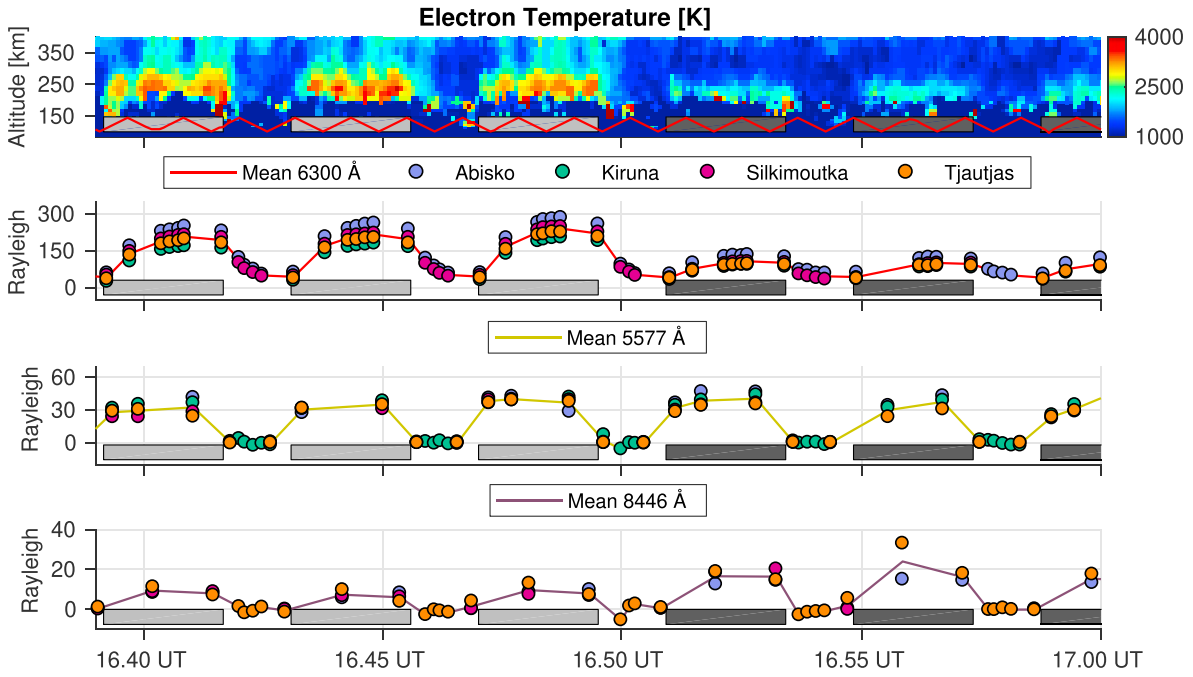


Figure 3. The European Incoherent Scatter ultrahigh frequency electron temperature observations and the enhanced emission intensities. The light gray boxes represent heating on during 6.200-MHz pumping and the dark gray boxes heating on at 5.423-MHz pumping. I_{6300} is not synchronized to the pump periods due to the long $O(^1D)$ lifetime. The scanning of the ultrahigh frequency radar beam through the heated volume is represented by the red line in the electron temperature panel.

model. An additional fitting parameter, γ, was used for adjusting the excitation rate profiles in the modified G-E model. Thus, six model parameters were used in the modified G-E model; $I_0, x_0, y_0, z_0, \sigma_{xy}$, and γ. Notice that the modified G-E model can be considered as a hybrid model, using the excitation rate profiles as the physical G-E model but is allowed for adjustment using the empirical γ parameter. The volume excitation rate model functions and the corresponding fitting parameters are described in more detail in Appendix A. It should be noted that increasing the number of fitting parameters, for example, allowing the excitation rates to be asymmetrical in the horizontal plane and adding more shape parameters along the magnetic field, would improve the fit to the observed data. However, we aimed at parameter fitting of unambiguous 3-D models with low parameter correlation and therefore employed few 3-D modeling parameters.

3.1. Parameter Fitting

A regularized version of the least squares error function, equation (1), was used in an iterative fitting process to determine the best fit 3-D aurora model parameters. The best fit parameters constructed 3-D emission models, which minimized the difference between the observed and the modeled images. The model images were produced by projecting the 3-D emission distribution down to the location of the imaging stations. An example of a best fit 3-D aurora model is presented in Figure 4 along with the comparison of the observed and modeled images for each of the four imaging stations: Abisko, Kiruna, Silkimuotka, and Tjautjas. The coordinates of the image station in Kiruna was chosen as origin. The coordinates of the 3-D model in Figure 4 are therefore given with respect to the Kiruna imaging station. The dot-projection algorithm, described in Rydesäter and Gustavsson (2000), was employed to calculate the projections of the 3-D aurora models. Several start guesses were used in the fitting process to avoid parameter searching in local minima. The model parameters at each iteration step was determined by the Nelder-Mead simplex search algorithm (Lagarias et al., 1998).

$$E(\mathbf{V}_t) = \arg\min_{\mathbf{V}_t}\left(\frac{1}{N_s}\left(\sum_s \sum_{i,j} [p_s(i,j) - \tilde{p}_s(i,j,I(\mathbf{V}_t))]^2 + W \sum_s |\mathbf{m_s} - \tilde{\mathbf{m}}_\mathbf{s}|^2\right)\right) \quad (1)$$

$E(\mathbf{V}_t)$ is the error function of the fitted parameter vector $\mathbf{V}_t$ at time t. N_s is the number of imaging stations, s is the station index, and (i,j) is the pixel index. $p_s(i,j)$ is the observed image in station s after

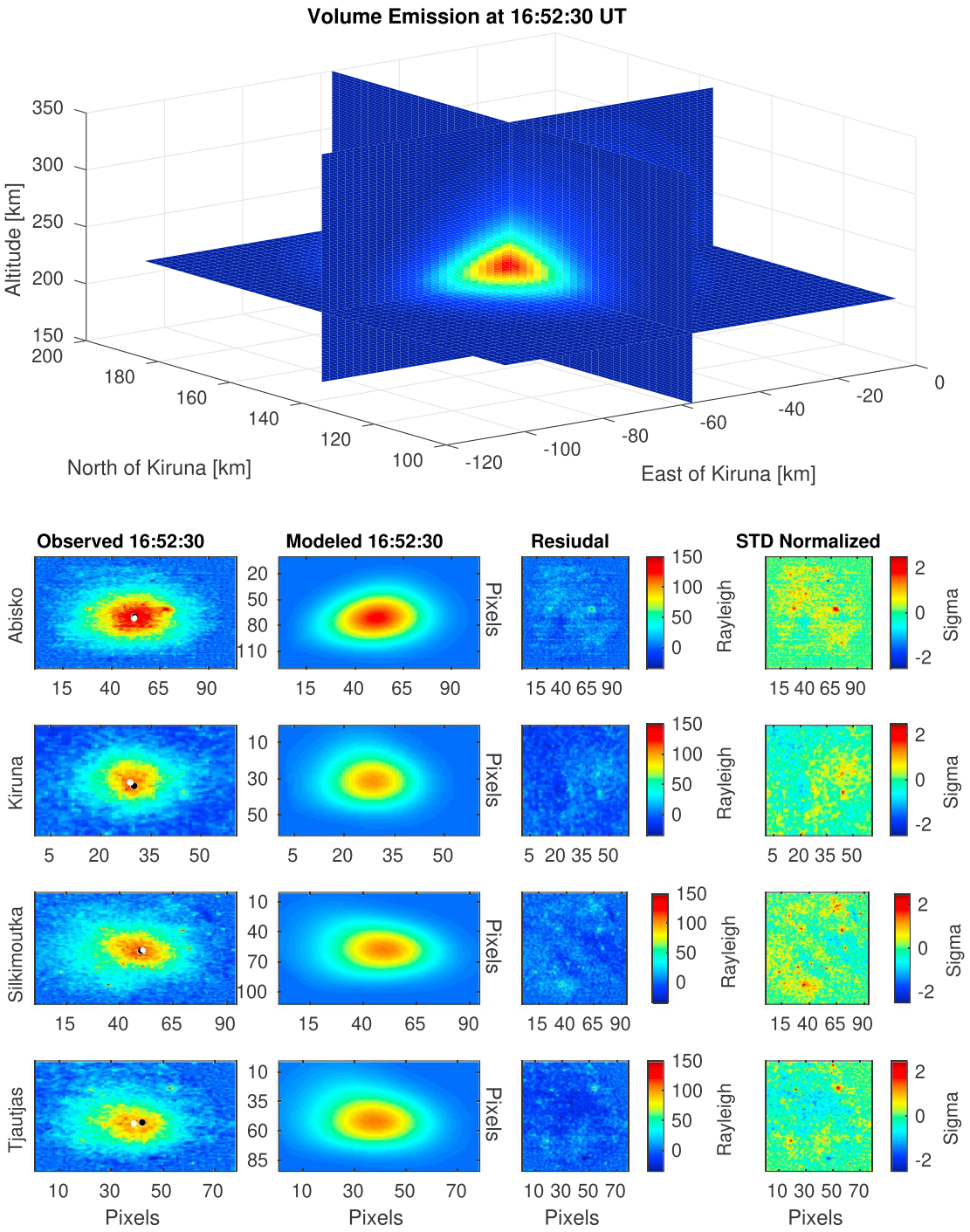


Figure 4. At the top, the best fit 3-D model of the 6,300-Å emission distribution at 16.52.30 UT. The 3-D distribution was constructed with the modified G-E model. At the bottom, the comparison between the observed and the modeled images. The first column shows the observed images for each of the four Auroral Large Imaging System stations. The second column shows the matching projections of the 3-D model. The third column shows the residual, that is, the pixel values in the observed images subtracted from the pixel values in the modeled images. The fourth column also shows the residual, but with the pixel values normalized with the standard deviation (STD) of the observations.

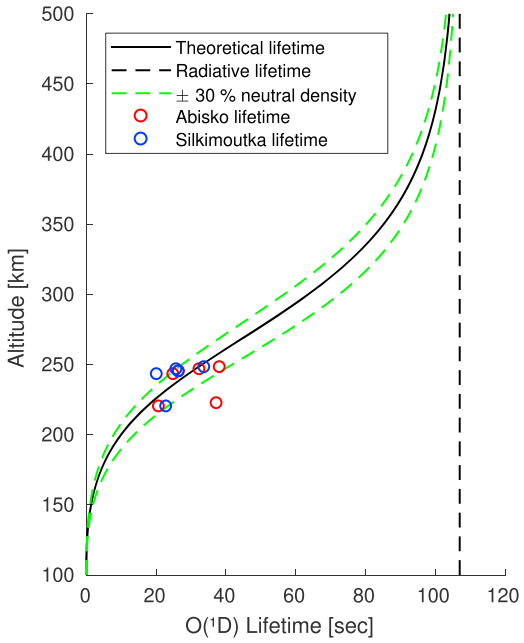


Figure 5. The figure depicts a comparison between the observed $O(^1D)$ lifetime and the theoretical $O(^1D)$ effective lifetime. Theoretical $O(^1D)$ effective lifetime estimates using ±30% of the MSIS-2000 neutral densities are included in the plot. The spread in observed $O(^1D)$ lifetime is large but agrees reasonably well with the theoretical $O(^1D)$ effective lifetime within the ±30 % neutral density confidence.

background reduction, and $\tilde{p}_s(i, j, I(\mathbf{V}_t))$ is the corresponding modeled image, the projection of the parametrized 3-D emission model, $I(\mathbf{V}_t)$, down to imaging station s. $[\mathbf{m_s} - \tilde{\mathbf{m}}_\mathbf{s}]$ is the image distance between the pixel value maxima in the observed and modeled images, marked with a black and a white dot in the observed images in Figure 4. W is a weight factor controlling the significance of the maxima coordinate difference.

To obtain accurate projections, both the field of view and the sensitivity of the cameras need to be known to high accuracy. Line-of-sight calibration was achieved by identifying stars in the image with the corresponding stars in the Yale Bright Star Catalog (Hoffleit & Jaschek,), the applied calibration method is described further in Gustavsson et al. (2008). Absolute intensity calibration factors from Wang (2011) were used to convert the Charge-Coupled Device counts to Rayleighs for images in 5,577 and in 6,300 Å. The Charge-Coupled Device sensitivity factors in the 8,446-Å emission line were determined by the irradiance spectra of the identified stars as given in the Pulkovo spectrometric catalog (Alekseeva et al., 1996).

3.2. Emission Model

The excitation-emission process was implemented in the aurora modeling in order to make the 3-D emission distributions comparable to the Gustavsson and Eliasson (2008) excitation profiles. The 5,577- and the 8,446-Å emission distributions were considered to be directly proportional to the excitation rate distributions. This is justified by the short radiative lifetime (0.7 s) of the $O(^1S)$ state and the spontaneous emission from the $O(3p^3P)$ state (Gustavsson et al., 2008). The effects of collisional de-excitation, quenching, can therefore be ignored. The effects of drift and diffusion are insignificant for $O(^1S)$ and negligible for $O(3p^3P)$, taking into account the short radiative lifetimes and that the size of the heated blob is about 20 km with drift speeds only up to a few hundred meters per second (Bernhardt et al., 2000; Gustavsson et al., 2001).

The $O(^1D)$ state has a radiative lifetime of 107 s. The $O(^1D)$-neutral collision frequency is higher than the $O(^1D)$ radiation frequency at lower F region altitudes. Most of the excited $O(^1D)$ states will therefore never emit 6,300-Å radiation before relaxation. Thus, the quenching needs to be accounted for in the I_{6300} 3-D aurora modeling. The quenching rate is predominantly dependent on the neutral density. The effective $O(^1D)$ lifetime can therefore be considered as a function of altitude. The $O(^1D)$ continuity model from Gustavsson and Eliasson (2008) was employed to calculate the altitude-dependent effective $O(^1D)$ lifetimes. Neutral temperatures and densities from the MSIS-2000 model (Picone et al., 2002) and electron temperatures and densities from the IRI model (Bilitza et al., 2014) were used in the effective $O(^1D)$ calculation. The resulting theoretical $O(^1D)$ lifetimes are presented in Figure 5 along with the observed $O(^1D)$ lifetimes from Abisko and Silkimuotka. The observed $O(^1D)$ lifetime is estimated from the 6,300-Å emission decay after heating offset.

Although the quenching considerably shortens the $O(^1D)$ lifetime, it is still relatively long, 30 s according to the theoretical estimate at a typical peak emission altitude of 245 km. Thus, the model function described in Gustavsson et al. (2001) was employed to account for the effects of the horizontal drift, diffusion, and the altitude-dependent intensity reduction from the quenching; see equations 3 and 7 in Gustavsson et al. (2001). The values of the $O(^1D)$ horizontal drift and diffusion are unknown; hence, the I_{6300} 3-D modeling requires three additional fitting parameters for all modeling methods: wind along the west-east direction, u_x, wind along the south-north direction, u_y, and diffusion, D. The vertical drift was neglected in the modeling, and it was assumed that the horizontal wind and the $O(^1D)$ diffusion were uniform within the heated volume.

4. Results

The best fit projections of the I_{6300} 3-D aurora models at 16.48.45 UT are presented in Figure 6 along with the comparison to the observed images at the matching time step. Note that an additional I_{6300} modified

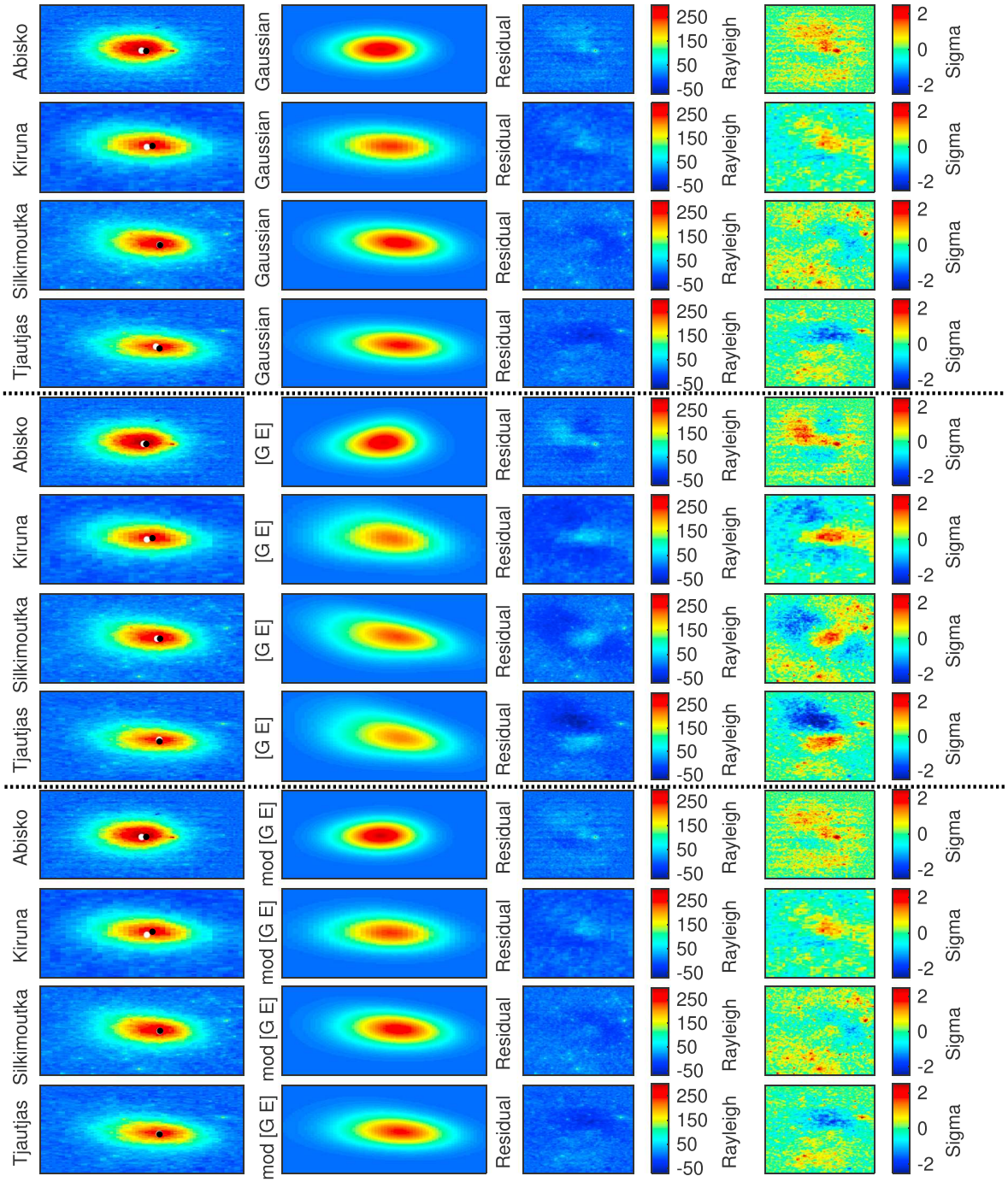


Figure 6. Rows 1-4 are the best fit Gaussian model projections, rows 5–8 are the best fit Gustavsson-Eliasson (G-E) model projections, and rows 9–12 are the best fit modified G-E model projections. The projections of the the Gaussian model and the modified G-E model are similar and in better agreement with the observed images than the projections of the G-E model. The pixel grids are similar to the pixel grids in Figure 4.

G-E model comparison, during 5.423-MHz heating, is presented in Figure 4. The remaining successful 3-D reconstruction results, 41 red aurora models, 15 green aurora models, and 2 infrared aurora models, for each of the three modeling methods, will not be presented in this article due to limited space.

Figure 6 clearly depicts a trend that is seen for all I_{6300} 3-D modeling results; the best fit projections of the the Gaussian and the modified G-E models are similar and in better agreement to the observed images

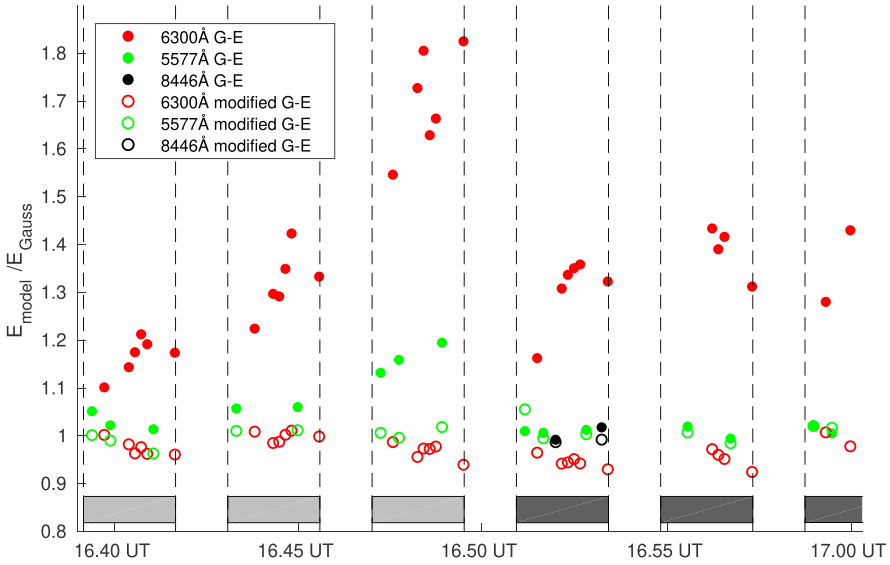


Figure 7. The Gustavsson-Eliasson (G-E) modeling produce I_{6300} projections with higher error values, E_{GE}, than both the Gaussian, E_{Gauss}, and the modified G-E, E_{modGE}, modeling methods, as seen by the error value ratios E_{GE}/E_{Gauss} and E_{modGE}/E_{Gauss}. The same trend is seen in the I_{5577} projections during 6.200-MHz pumping. As before, the light gray boxes represent heating on during 6.200-MHz pumping and the dark gray boxes heating on at 5.423-MHz pumping.

than the best fit G-E model projections. The same trend is also seen in the I_{5577} model projections during 6.200-MHz heating. There are, however, no significant deviation in the projections of the best fit models during 5.423-MHz pumping in I_{5577} and I_{8446}. The error values, from the error function in equation (1), of the best fit models were used to underline this result. Figure 7 presents the error value ratio between the G-E models and the Gaussian models, E_{GE}/E_{Gauss}, and the ratio between the modified *G-E* models and the Gaussian models E_{modGE}/E_{Gauss} at different time steps, t, and for all enhanced emissions. Figure 7 shows that the G-E modeling method produced model projections with less of an agreement to the observed images than the Gaussian model projections. On average, the G-E modeling produce I_{6300} projections with ~40% higher error values during 6.200- and 5.423-MHz heating and I_{5577} projections with ~10% higher error values during 6.200-MHz heating. The modified G-E modeling produce similar error values to the Gaussian modeling for all emission lines. It should be noted that the signal-to-noise ratio is much higher for I_{6300} than for I_{5577} and I_{8446}; see Rayleigh enhancement values in Figure 3. The discrepancy between the models might only be seen when the signal-to-noise ratio is high, which could explain why there are no clear deviations between the I_{5577} and I_{8446} model projections during 5.423-MHz pumping.

Mean excitation-altitude profiles were calculated to study the discrepancy between the physical G-E model and the empirical 3-D Gaussian and semiempirical modified G-E models. The resulting $O(^1D)$, $O(^1S)$, and $O(3p^3P)$ excitation-altitude profiles, for all modeling methods, are presented in Figure 8. Notice in Figure 8 that the Gaussian excitation rate profiles are in good agreement with the modified G-E excitation rate profiles for all excitation states and for both pump frequencies. Also note that Figure 8 only depicts the excitation rate distributions as a function of altitude. The deviations of the resulting 3-D aurora models will predominantly be along the altitude axis since all aurora models are created similarly in the horizontal plane, using a 2-D Gaussian distribution, but have different distribution functions along the magnetic field line. Normalized mean excitation-altitude profiles were used to produce the curves in Figure 8 in order to make the shape of excitation-altitude profiles comparable between all heating pulses. The normalized mean $O(^1D)$, $O(^1S)$, and $O(3p^3P)$ excitation rates in Figure 8 were calculated using 3-D reconstruction results during near steady state conditions in I_{6300}, I_{5577}, and I_{8446}, respectively. The 3-D reconstruction results from 16 I_{6300} models and 7 I_{5577} models during 6.200-MHz pumping and 7 I_{6300} models and 6 I_{5577} models during 5.423-MHz heating were used to produce the $O(^1D)$ and $O(^1S)$ excitation rate curves. Tomographic reconstruction in I_{8446} was successfully done at only two time steps. $O(3p^3P)$ excitation rate curves were therefore produced using two reconstruction results, both during 5.423-MHz pumping. The width of the curves in Figure 8 equals the $\pm\sigma$

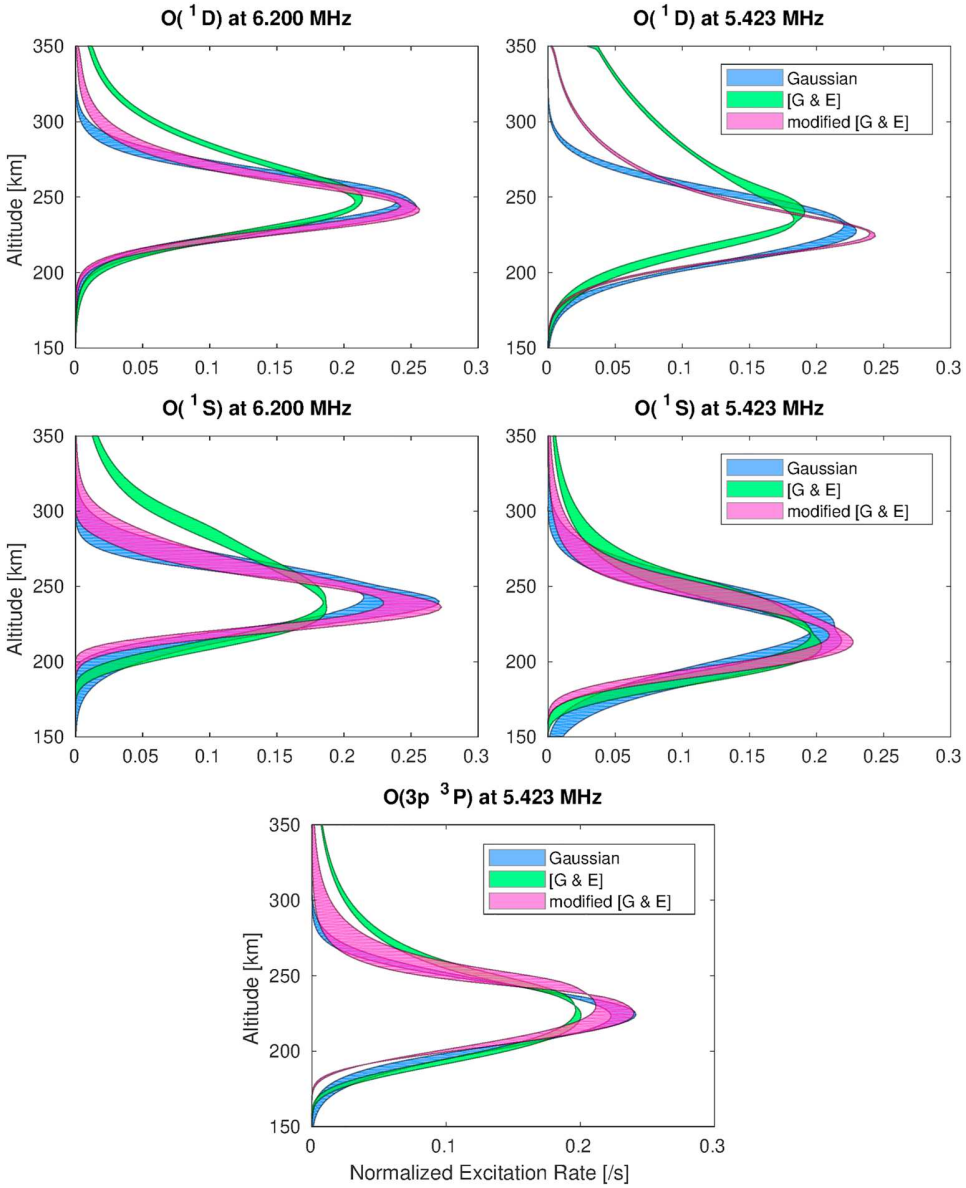


Figure 8. The figure depicts the resulting normalized mean excitation rates as a function of altitude for the three different 3-D aurora models; the empirical 3-D Gaussian model, the physical G-E model and hybrid modified G-E model. The best fit models from the the Gaussian and modified G-E methods have a smaller altitude range than predicted by the physical G-E model in $O(^1D)$ and $O(^1S)$ during 6.200-MHz pumping and in $O(^1D)$ during 5.423-MHz heating.

width for the $O(^1D)$ and $O(^1S)$ excitation rates. The edges of the $O(3p^3P)$ excitation rate curves are defined by the two resulting excitation profiles.

In addition, the best fit Gaussian model parameter values in altitude, z_0, horizontal width, σ_{xy}, and field-aligned width, σ_z, are presented in Figure 9 to depict the evolution of the $O(^1D)$, $O(^1S)$, and $O(3p^3P)$ volume excitation rates. The altitude of the excitation center drops from about 240 to 220 km when changing pump frequency from 6.200 to 5.423 MHz. This is a consequence of the electron density gradient in the lower F region. HF radio waves at a lower frequency will interact with a lower density plasma; the wave-plasma resonance altitude will therefore drop. The horizontal excitation width is about 12 km at 6.200-MHz heating and increases to around 16 km at 5.423-MHz heating. This increase is larger than the inferred ~1-km increase in the horizontal excitation width from the modeled beam patterns, as seen before in Figure 2.

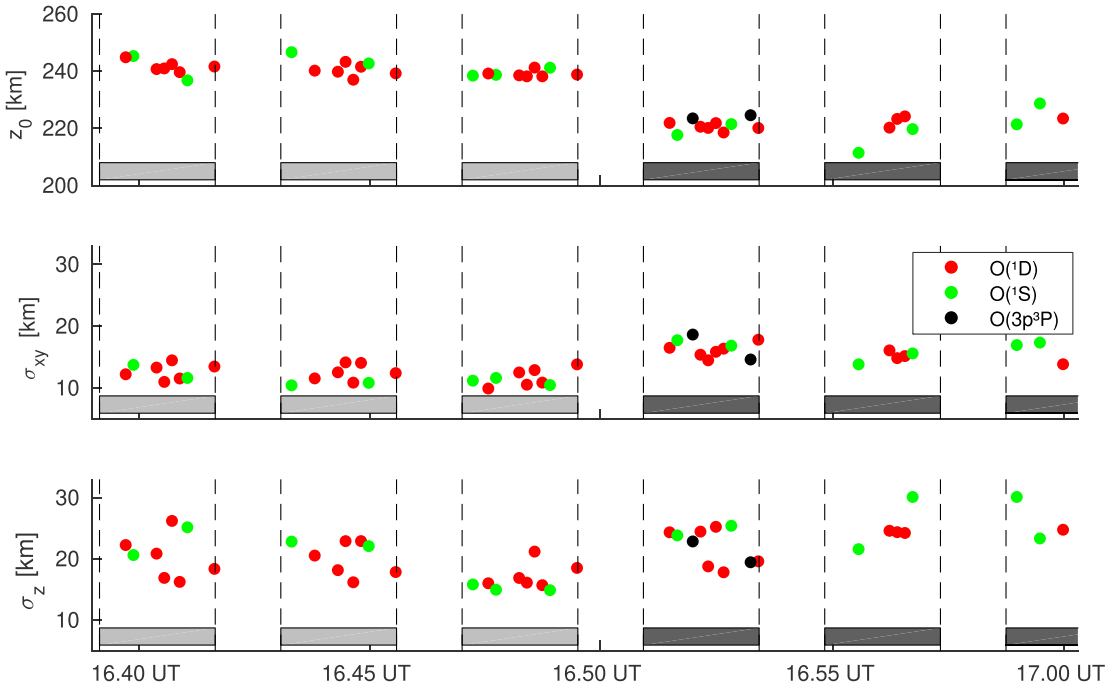


Figure 9. Gaussian best fit parameters, from top to bottom: Excitation center altitude, horizontal width, and width along the magnetic field. The light gray boxes represent heating on during 6.200-MHz pumping and the dark gray boxes heating on at 5.423-MHz pumping.

Additionally, in Figure 9, there is an apparent small increase in width along the magnetic field line when changing pumping frequency from 6.200 to 5.423 MHz, although the spread is somewhat large. There are no clear deviations in excitation center altitudes and widths between the different species.

5. Discussion

The main result in this article is that there are deviations between the physical G-E reconstructions and the empirical Gaussian and semiempirical modified G-E reconstructions. In particular, Figure 8 suggests that $O(^1D)$ and $O(^1S)$ excitation during 6.200-MHz heating and $O(^1D)$ excitation during 5.423-MHz pumping are induced within a smaller altitude range than predicted by Gustavsson and Eliasson (2008). This conclusion is based on three observations: (1) The model fits, an example is seen in Figure 6, are adequate to be meaningful fits. (2) The Gaussian and the modified G-E methods produce aurora models, which are significantly better than the best fit G-E model projections, as seen in Figure 7. (3) The excitation-altitude distributions of the G-E models are statistically inconsistent with the distributions of the Gaussian and the modified G-E models; see Figure 8.

An additional result is seen in Figure 3; I_{6300} and the electron temperature enhancement decreased when going from 6.200- to 5.423-MHz pumping, while the high threshold energy I_{8446} enhancement increased. Only a ~30% I_{6300} reduction can be explained by increased quenching at lower heating altitudes and about 15% electron temperature and I_{6300} reduction is caused by reduced power flux during 5.423-MHz pumping. A similar trend was reported in Gustavsson et al. (2006) by observing electron temperature enhancements, I_{6300}, I_{5577}, and the high threshold energy I_{4278} during a heating experiment with frequency stepping around the fourth gyroharmonic. Gustavsson et al. (2006) observed that I_{4278} was insignificant during heating far above or below the gyroharmonic and increased when heating a few tens of kilohertz above the fourth double resonance. This indicates, by comparison to Gustavsson et al. (2006), that heating occurred away from a gyroharmonic during 6.200-MHz pumping and just above the fourth gyroharmonic during 5.423-MHz pumping, although not close enough to the fourth double resonance to induce significant I_{4278}. The observations in Figure 3 and in Gustavsson et al. (2006) are in agreement with the Vlasov simulation results of accelerated electrons presented in Najmi et al. (2017). Najmi et al. (2017) argued that the electron energy distribution during heating just above the fourth gyroharmonic will have a more pronounced high-energy

tail than heating away from the gyroharmonic. The reason for this is that HF pumping just above the gyroharmonic will energize fewer electrons, but to higher energies through resonance acceleration by excited upper-hybrid waves. On the contrary, HF heating away from the gyroharmonic mainly enhances electron energization through bulk heating from electron Bernstein waves.

Three possible explanations to the confined excitation region, both during heating away and close to the fourth gyroharmonic, will be discussed: (1) anisotropic electron acceleration, (2) underestimated neutral densities, and (3) shortcomings in the Gustavsson and Eliasson (2008) electron transport model. Of these three, anisotropic electron acceleration appears to be the only viable explanation.

5.1. Anisotropic Electron Acceleration

The Gustavsson and Eliasson (2008) model assumes an isotropic electron acceleration, as assumed in other attempts to find the electron energy-altitude distribution during ionospheric F region heating at high latitudes (e.g., Gustavsson et al., 2005; Hysell et al., 2012, 2014). This assumption may, however, not be valid. Grach (1999) predicted that an anisotropic electron acceleration, with an enhanced population of energized electrons perpendicular to the magnetic field, would be induced when pumping in resonance with both the upper-hybrid frequency and an harmonic of the electron gyrofrequency; that is, $(f_h \simeq f_u \simeq nf_e)$ where n is an integer. Hence, the narrow excitation rate profiles during 5.423-MHz pumping could possibly be explained by double resonance heating in proximity to the fourth gyroharmonic frequency; that is, $(f_h \simeq f_u \simeq 4f_e)$. The deviation between the resulting excitation rate profiles at 6.200-MHz heating can, however, not be explained by proximity to the fourth double resonance point. Grach (1999) argued that the anisotropic component of the electron velocity distribution could be neglected during heating away from the double resonance due to isotropization of the accelerated electrons by scattering in the acceleration region.

The discrepancy between excitation rate profiles from the isotropic G-E modeling method and the excitation rate profiles from the empirical Gaussian and semiempirical modified G-E modeling can therefore only partly be related to the predictions of Grach (1999). Either the anisotropic component is more prominent than predicted during heating away from the double resonance or other mechanisms are responsible for the deviations between the predicted and empirically reconstructed excitation rate profiles.

5.2. Underestimated Neutral Density

The neutral densities from the MSIS-2000 model were used to calculate the Gustavsson and Eliasson (2008) profiles. An underestimation of the neutral densities would cause the Gustavsson and Eliasson (2008) profiles to miscalculate the electron-neutral collision frequency and favor a larger altitude range for the excitation. However, a crude underestimation of the neutral density would also cause a systematic difference between the theoretical and observed $O(^1D)$ lifetimes. No such difference is observed in Figure 5, and we therefore consider an underestimation of the neutral densities to be an unlikely cause for the discrepancy between the predicted and empirically reconstructed excitation rate profiles.

5.3. Shortcomings in the Gustavsson and Eliasson (2008) Transport Model

We cannot ignore the possibility of shortcomings in the two-stream electron transport model to be the cause of the deviations. However, none of the parameters in the Gustavsson and Eliasson (2008) transport model, apart from the neutral density, can cause significant altitude confinement of the excitation rates. In addition, the resulting electron energy distributions as a function of energy and altitude from Gustavsson and Eliasson (2008) are in apparent agreement with the results from Hysell et al. (2012, 2014), which employed a different electron transport model, the SAMI2-PE model (Varney et al., 2012). This indicates that the deviations between the predicted and empirically reconstructed excitation rate profiles do not likely originate from shortcomings or errors in the applied electron transport model.

6. Summary and Conclusions

In this paper, we have presented a method for comparing electron energy-altitude distributions to multiview point images by employing tomography-like inversion methods and excitation rate profiles, calculated using the method described in Gustavsson and Eliasson (2008). Modeling of the artificial aurora was done in the enhanced 6,300-, 5,577-, and 8,446-Å emission lines, providing estimates of the $O(^1D)$, $O(^1S)$, and $O(3p^3P)$ volume excitation rates. The resulting excitation rates, see Figure 8, indicate that the excitation rate distributions are less extended in altitude than predicted by the excitation rate profiles from Gustavsson and Eliasson (2008). The Gustavsson and Eliasson (2008) transport model assumes an isotropic electron accel-

eration. A possible explanation for this discrepancy might therefore be that HF heating both away from and just above the fourth double resonance induces an anisotropic distribution of energized electrons. Grach (1999) analytically predicted an anisotropic electron distribution, with an enhanced population perpendicular to the magnetic field, during HF heating near a double resonance point, but argued for isotropization of the accelerated electrons during HF pumping away from a double resonance point. A remaining open question is whether an anisotropic electron acceleration can lead to excitation rate profiles which are consistent with our observations. This conjecture could be investigated using a multistream electron transport model where all elastic and inelastic collisions are taken into account in addition to an anisotropic electron energization term. Additionally, the electron temperature enhancements and the emission enhancements during heating away and just above the fourth double resonance point are in agreement with the electron energy distributions obtained by Vlasov simulations of accelerated electrons, as presented in Najmi et al. (2017).

Appendix A: Aurora Models

A1. The Gaussian Model

The 3-D Gaussian aurora models are constructed using Gaussian distributions along all axes, $G(x, y, z, \sigma_{xy}, \sigma_z)$:

$$f(x,y,z) = I_0 \cdot \exp\left(-\left(\frac{X^2}{2\sigma_{xy}^2} + \frac{((Y + Z\sin(\theta))^2}{2\sigma_{xy}^2} + \frac{Z^2}{2\sigma_z^2}\right)\right) \tag{A1}$$

where

$$X, Y, Z = x - x_0, y - y_0, z - z_0 \tag{A2}$$

I_0 is the maximum excitation rate, located at the center of the excitation volume (x_0, y_0, z_0). The excitation distribution was assumed to have a symmetrical Gaussian shape in the horizontal plane; σ_{xy} is the horizontal $(1/e)$ width radius. θ is the magnetic field angle from zenith, 12°S at EISCAT Ramfjordmoen. σ_z is the width in altitude. The constructed excitation blob is allowed to be asymmetric ($\sigma_z \neq \sigma_{xy}$).

A2. The G-E and The Modified G-E Models

The G-E and the modified G-E modeling used excitation rate profiles, calculated using the method described in Gustavsson and Eliasson (2008), to construct the excitation-altitude distributions. $I_{z0}(z)$ is the normalized excitation rate profile as a function of altitude, z, with an excitation center at altitude z_0. A shape parameter, γ, was used to modify the shape of the normalized excitation-altitude profiles, $I_{z0}^{\gamma}(z)$. This yields the volume excitation function, $f(x, y, z)$:

$$f(x,y,z) = I_0 \cdot I_{z0}^{\gamma}(Z) \cdot \exp\left(-\left(\frac{X^2}{2\sigma_{xy}^2} + \frac{(Y + Z\sin(\theta))^2}{2\sigma_{xy}^2}\right)\right) \tag{A3}$$

The G-E modeling is defined by $\gamma = 1$; that is, the G-E excitation rate profiles were used directly. In the modified G-E modeling, the excitation rates were allowed to be modified by adjusting the γ value. A high gamma value ($\gamma > 1$) will sharpen the excitation rate profiles around the peak excitation altitude, z_0, whereas a low gamma value ($\gamma < 1$) will make the profiles more blunt.

A3. The Modeling Search Parameters

The required auroral modeling parameters, depending on the modeling method, are summarized in Table A1. The horizontal neutral wind and diffusion parameters are only necessary for the 6,300-Å

Table A1
The Search Parameters

Search parameter	Gaussian	G-E	Modified G-E
Maximum excitation	I_0	I_0	I_0
Excitation midpoint	x_0, y_0, z_0	x_0, y_0, z_0	x_0, y_0, z_0
Excitation radius/size	σ_{xy}, σ_z	σ_{xy}	σ_{xy}
Shape parameter			γ
(Neutral wind)	(u_x, u_y)	(u_x, u_y)	(u_x, u_y)
(Diffusion)	(D)	(D)	(D)
Number of parameters	6 (9)	5 (8)	6 (9)

emission modeling; these parameters are therefore within parenthesis. The vertical wind was assumed to be negligible.

Acknowledgments
ALIS is supported by the Swedish Research Council. EISCAT is an international association supported by research organizations in China (CRIRP), Finland (SA), Japan (NIPR and STEL), Norway (NFR), Sweden (VR), and the United Kingdom (NERC). A. Kvammen and J. Vierinen are supported by the Tromsø Research Foundation. The ALIS multistation images from the 16 February 2015 are available in the ALIS data archive (http://www.alis.irf.se/stdnames/2015/02/16/), and the EISCAT data are available in the Madrigal database. MatLab functions from the Auroral Image Data Analysis tools (AIDA-tools); http://www.alis.irf.se/~bjorn/AIDA_tools/, were used for camera calibration, data pre-processing and projection calculations from 3D volume emission rates to modeled image pixel intensities. The authors would like to thank the reviewers for valuable comments and suggestions.

References

Alekseeva, G., Arkharov, A., Galkin, V., Hagen-Thorn, E., Nikanorova, I., Novikov, V., et al. (1996). The Pulkovo spectrophotometric catalog of bright stars in the range from 320 to 1080 NM. *Baltic Astronomy*, *5*, 603–838.

Bernhardt, P. A., Siefring, C. L., Briczinski, S. J., McCarrick, M., & Michell, R. G. (2016). Large ionospheric disturbances produced by the HAARP HF facility. *Radio Science*, *51*, 1081–1093. https://doi.org/10.1002/2015RS005883

Bernhardt, P. A., Tepley, C. A., & Duncan, L. M. (1989). Airglow enhancements associated with plasma cavities formed during ionospheric heating experiments. *Journal of Geophysical Research*, *94*(A7), 9071–9092.

Bernhardt, P., Wong, M., Huba, J., Fejer, B. G., Wagner, L., Goldstein, J., et al. (2000). Optical remote sensing of the thermosphere with HF pumped artificial airglow. *Journal of Geophysical Research*, *105*(A5), 10,657–10,671.

Bilitza, D., Altadill, D., Zhang, Y., Mertens, C., Truhlik, V., Richards, P., et al. (2014). The International Reference Ionosphere 2012—A model of international collaboration. *Journal of Space Weather and Space Climate*, *4*(A07), 12.

Brändström, U. (2003). The Auroral Large Imaging System: Design, operation and scientific results.

Brändström, B. U. E., Leyser, T. B., Steen, Å., Rietveld, M. T., Gustavsson, B., Aso, T., & Ejiri, M. (1999). Unambiguous evidence of HF pump-enhanced airglow at auroral latitudes. *Geophysical Research Letters*, *26*(23), 3561–3564.

Carlson, H. C., Wickwar, V. B., & Mantas, G. P. (1982). Observations of fluxes of suprathermal electrons accelerated by HF excited instabilities. *Journal of Atmospheric and Terrestrial Physics*, *44*(12), 1089–1100.

Coster, A. J., Djuth, F. T., Jost, R. J., & Gordon, W. E. (1985). The temporal evolution of 3-m striations in the modified ionosphere. *Journal of Geophysical Research*, *90*(A3), 2807–2818.

Djuth, F. T., Isham, B., Rietveld, M. T., Hagfors, T., & La Hoz, C. (2004). First 100 ms of HF modification at Tromsø, Norway. *Journal of Geophysical Research*, *109*, A11307. https://doi.org/10.1029/2003JA010236

Djuth, F. T., Pedersen, T. R., Gerken, E. A., Bernhardt, P. A., Selcher, C. A., Bristow, W. A., & Kosch, M. J. (2005). Ionospheric modification at twice the electron cyclotron frequency. *Physical Review Letters*, *94*(12), 125001.

Eliasson, B., & Papadopoulos, K. (2015). Numerical study of anomalous absorption of O mode waves on magnetic field-aligned striations. *Geophysical Research Letters*, *42*, 2603–2611. https://doi.org/10.1002/2015GL063751

Gondarenko, N. A., Ossakow, S. L., & Milikh, G. M. (2005). Generation and evolution of density irregularities due to self-focusing in ionospheric modifications. *Journal of Geophysical Research*, *110*, A09304. https://doi.org/10.1029/2005JA011142

Grach, S. M. (1999). On kinetic effects in the ionospheric F-region modified by powerful radio waves. *Radiophysics and Quantum Electronics*, *42*(7), 572–588.

Grach, S., Sergeev, E., Mishin, E., & Shindin, A. V. (2017). Dynamic properties of ionospheric plasma turbulence driven by high-power high-frequency radiowaves. *Physics-Uspekhi*, *59*(11), 1091.

Gurevich, A. V. (2007). Nonlinear effects in the ionosphere. *Physics-Uspekhi*, *50*(11), 1091–1121.

Gurevich, A. V., & Zybin, K. P. (2006). Strong field aligned scattering of UHF radio waves in ionospheric modification. *Physics Letters A*, *358*(2), 159–165.

Gustavsson, B., Brändström, B. U. E., Steen, Å., Sergienko, T., Leyser, T. B., Rietveld, M. T., et al. (2002). Nearly simultaneous images of HF-pump enhanced airglow at 6300 Åand 5577 Å. *Geophysical Research Letters*, *29*(24), 2220. https://doi.org/10.1029/2002GL015350

Gustavsson, B., & Eliasson, B. (2008). Hf radio wave acceleration of ionospheric electrons: Analysis of HF-induced optical enhancements. *Journal of Geophysical Research*, *113*, A08319. https://doi.org/10.1029/2007JA012913

Gustavsson, B. J., Kosch, M., Wong, A., Pedersen, T., Heinselman, C., Mutiso, C., et al. (2008). First estimates of volume distribution of HF-pump enhanced emissions at 6300 and 5577 å: A comparison between observations and theory.

Gustavsson, B., Leyser, T. B., Kosch, M., Rietveld, M. T., Steen, Å., Brändström, B. U. E., & Aso, T. (2006). Electron gyroharmonic effects in ionization and electron acceleration during high-frequency pumping in the ionosphere. *Physical Review Letters*, *97*(19), 195002.

Gustavsson, B., Sergienko, T., Kosch, M., Rietveld, M., Brändström, B., Leyser, T., et al. (2005). The electron energy distribution during HF pumping, a picture painted with all colors. *In Annales Geophysicae*, *23*, 1747–1754.

Gustavsson, B., Sergienko, T., Rietveld, M. T., Honary, F., Steen, A., Brändström, B. U. E., et al. (2001). First tomographic estimate of volume distribution of HF-pump enhanced airglow emission. *Journal of Geophysical Research*, *106*(A12), 29,105–29,123.

Haslett, J. C., & Megill, L. R. (1974). A model of the enhanced airglow excited by RF radiation. *Radio Science*, *9*(11), 1005–1019.

Hoffleit, D., & Jaschek, C.The bright star catalogue. New Haven, Conn.: Yale University Observatory, c1991, 5th rev. ed., edited by Hoffleit, Dorrit; Jaschek, Carlos.

Holma, H., Kaila, K. U., Kosch, M. J., & Rietveld, M. T. (2006). Recognizing the blue emission in artificial aurora. *Advances in Space Research*, *38*(11), 2653–2658.

Honary, F., Stocker, A., Robinson, T., Jones, T., & Stubbe, P. (1995). Ionospheric plasma response to HF radio waves operating at frequencies close to the third harmonic of the electron gyrofrequency. *Journal of Geophysical Research*, *100*(A11), 21,489–21,501.

Hysell, D. L., Miceli, R. J., Kendall, E. A., Schlatter, N. M., Varney, R. H., Watkins, B. J., et al. (2014). Heater-induced ionization inferred from spectrometric airglow measurements. *Journal of Geophysical Research: Space Physics*, *119*, 2038–2045. https://doi.org/10.1002/2013JA019663

Hysell, D. L., Varney, R. H., Vlasov, M. N., Nossa, E., Watkins, B., Pedersen, T., & Huba, J. D. (2012). Estimating the electron energy distribution during ionospheric modification from spectrographic airglow measurements. *Journal of Geophysical Research*, *117*, A02317. https://doi.org/10.1029/2011JA017187

Istomin, Y. N., & Leyser, T. (2003). Electron acceleration by cylindrical upper hybrid oscillations trapped in density irregularities in the ionosphere. *Physics of Plasmas*, *10*(7), 2962–2970.

Kendall, E., Marshall, R., Parris, R. T., Bhatt, A., Coster, A., Pedersen, T., et al. (2010). Decameter structure in heater-induced airglow at the High frequency Active Auroral Research Program facility. *Journal of Geophysical Research*, *115*, A08306. https://doi.org/10.1029/2009JA015043

Kosch, M. J., Rietveld, M. T., Kavanagh, A. J., Davis, C., Yeoman, T. K., Honary, F., & Hagfors, T. (2002). High-latitude pump-induced optical emissions for frequencies close to the third electron gyro-harmonic. *Geophysical Research Letters*, *29*(23), 2112. https://doi.org/10.1029/2002GL015744

Lagarias, J. C., Reeds, J. A., Wright, M. H., & Wright, P. E. (1998). Convergence properties of the Nelder–Mead simplex method in low dimensions. *SIAM Journal on Optimization*, *9*(1), 112–147.

Lehtinen, M. S., & Huuskonen, A. (1996). General incoherent scatter analysis and GUISDAP. *Journal of Atmospheric and Terrestrial Physics*, *58*(1), 435–452.

Leyser, T. B. (2001). Stimulated electromagnetic emissions by high-frequency electromagnetic pumping of the ionospheric plasma. *Space Science Reviews*, *98*(3-4), 223–328.

Mantas, G. P. (1994). Large 6300-Åairglow intensity enhancements observed in ionosphere heating experiments are excited by thermal electrons. *Journal of Geophysical Research*, *99*(A5), 8993–9002.

Mantas, G. P., & Carlson, H. C. (1996). Reinterpretation of the 6300-Åairglow enhancements observed in ionosphere heating experiments based on analysis of Platteville, Colorado, data. *Journal of Geophysical Research*, *101*(A1), 195–209.

Milikh, G., Gurevich, A., Zybin, K., & Secan, J. (2008). Perturbations of GPS signals by the ionospheric irregularities generated due to HF-heating at triple of electron gyrofrequency. *Geophysical Research Letters*, *35*, L22102. https://doi.org/10.1029/2008GL035527

Najmi, A., Eliasson, B., Shao, X., Milikh, G., Sharma, A. S., & Papadopoulos, K. (2017). Vlasov simulations of electron acceleration by radio frequency heating near the upper hybrid layer. *Physics of Plasmas*, *24*(10), 102904.

Pedersen, T., Gustavsson, B., Mishin, E., MacKenzie, E., Carlson, H., Starks, M., & Mills, T. (2009). Optical ring formation and ionization production in high-power heating experiments at HAARP. *Geophysical Research Letters*, *36*, L18107. https://doi.org/10.1029/2009GL040047

Picone, J. M., Hedin, A. E., Drob, D. P., & Aikin, A. C. (2002). NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues. *Journal of Geophysical Research*, *107*(A12), 1468.

Rietveld, M. T., Kohl, H., Kopka, H., & Stubbe, P. (1993). Introduction to ionospheric heating at Tromsø I. Experimental overview. *Journal of atmospheric and terrestrial physics*, *55*(4-5), 577–599.

Rietveld, M. T., Kosch, M. J., Blagoveshchenskaya, N. F., Kornienko, V. A., Leyser, T. B., & Yeoman, T. K. (2003). Ionospheric electron heating, optical emissions, and striations induced by powerful HF radio waves at high latitudes: Aspect angle dependence. *Journal of Geophysical Research*, *108*(A4), 1141. https://doi.org/10.1029/2002JA009543

Rietveld, M. T., Senior, A., Markkanen, J., & Westman, A. (2016). New capabilities of the upgraded EISCAT high-power HF facility. *Radio Science*, *51*, 1533–1546. https://doi.org/10.1002/2016RS006093

Rishbeth, H., & Van Eyken, A. (1993). EISCAT: Early history and the first ten years of operation. *Journal of atmospheric and terrestrial physics*, *55*(4-5), 525–542.

Robinson, T. (1989). The heating of the high lattitude ionosphere by high power radio waves. *Physics reports*, *179*(2-3), 79–209.

Rydesäter, P., & Gustavsson, B. (2000). Investigation of smooth basis functions and an approximated projection algorithm for faster tomography. *International journal of imaging systems and technology*, *11*(6), 347–354.

Senior, A., Rietveld, M. T., Yeoman, T. K., & Kosch, M. J. (2012). The dependence of F-region electron heating on HF radio pump power: Measurements at EISCAT Tromsø. *Journal of Geophysical Research*, *117*, A04309. https://doi.org/10.1029/2011JA017267

Shindin, A., Klimenko, V., Kogogin, D., Beletsky, A., Grach, S., Nasyrov, I., & Sergeev, E. (2018). Spatial characteristics of the 630-nm artificial ionospheric airglow generation region during the Sura facility pumping. *Radiophysics and Quantum Electronics*, *60*(11), 849–865.

Stubbe, P., Stocker, A. J., Honary, F., Robinson, T. R., & Jones, T. B. (1994). Stimulated electromagnetic emissions and anomalous HF wave absorption near electron gyroharmonics. *Journal of Geophysical Research*, *99*(A4), 6233–6246.

Varney, R. H., Swartz, W. E., Hysell, D. L., & Huba, J. D. (2012). SAMI2-PE: A model of the ionosphere including multistream interhemispheric photoelectron transport. *Journal of Geophysical Research*, *117*, A06322. https://doi.org/10.1029/2011JA017280

Wang, Z. (2011). Absolute calibration of ALIS cameras.