Machine learning analysis of Jupiter's far-ultraviolet auroral morphology

J. D. Nichols[1], A. Kamran[1], and S. E. Milan[1]

[1]Department of Physics and Astronomy, University of Leicester, University Road, Leicester, LE1 7RH, UK

Key Points:

- PCA of Jupiter's FUV auroras indicates variation in the emission poleward of the statistical oval on the dawn side is most recurrent.
- DBSCAN objectively classifies auroral images into six repeatable morphological classes.
- One morphological class exhibiting bright main and poleward dusk emissions is identified with solar wind compressions.

Corresponding author: J. D. Nichols, `jdn4@le.ac.uk`

Abstract

We present the first principal component analysis of Jupiter's far-ultraviolet auroras, in order to identify the most repeatable sources of variation in the auroral morphology. We show that the most recurrent source of variance is emission just poleward of the statistical oval on the dawn side. Further significant repeatable sources of variance are localised expansions of the main emission on the dawn or dusk sides and poleward emission near noon and along the dusk side. We go on to show using a DBSCAN clustering analysis that the most significant auroral components form six repeatable auroral morphological classes. One class, exhibiting bright main and poleward dusk emissions, occurs solely during solar wind compressions. This presents an important new tool for diagnosing magnetospheric compressions at Jupiter.

1 Introduction

The Hubble Space Telescope (HST) has revealed Jupiter's FUV auroras to exhibit a complex morphology with a number of different components, including the main auroral emission, low latitude patches and arcs, and variable polar emissions (e.g. Grodent, Clarke, Waite Jr, et al., 2003; Grodent, Clarke, Kim, et al., 2003; Clarke et al., 2004, 2009; Nichols, Clarke, Gérard, Grodent, & Hansen, 2009; Nichols, Clarke, Gérard, & Grodent, 2009; Radioti et al., 2009; Dumont et al., 2015; Bonfond et al., 2008; Gray et al., 2016; Bonfond et al., 2017; Nichols et al., 2017; Grodent et al., 2018). Briefly, the satellite footprints are magnetically linked to the Galilean satellites, the main emission (ME) is thought to be driven by breakdown of corotation of iogenic plasma in the middle magnetosphere and associated magnetosphere-ionosphere coupling current system, while the high-latitude polar emissions map to the outer magnetosphere and magnetotail. The main emission is occasionally superimposed by bright patches thought to be associated with plasma injections in the middle magnetosphere, and the dawn side of the main emission is on occasion observed to brighten to very high intensities in events known as 'dawn storms'. Immediately poleward of the main emission, usually most evident on the dawn side, is a dark polar region, and poleward of this lies a highly dynamic region of transient emission known as the 'swirl' region. A dynamic and sometimes extremely bright region near noon known as the 'active' region, and poleward of the main emission on the dusk side lie polar dusk arcs, which are most evident when the magnetosphere is compressed by the solar wind. An example image of Jupiter's auroras as obtained by the Space Tele-

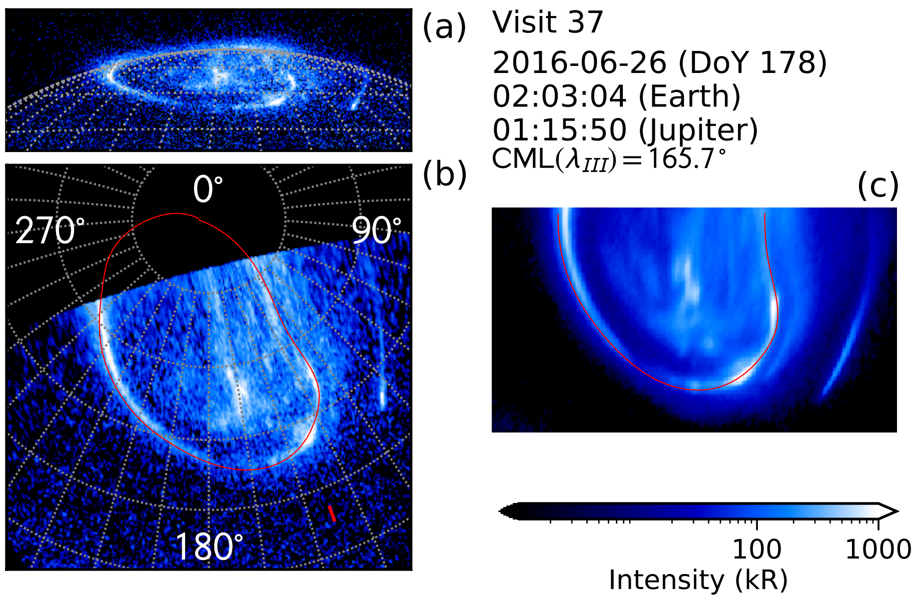

Figure 1. Plot showing (a) an example of an unprojected image with a 10×10° graticule overlaid; (b) the same image projected onto a latitude-longitude grid, presented on an equal area azimuthal Lambert projection with System III longitudes labelled, a 10×10° graticule in grey, and the (Nichols et al., 2017) statistical oval shown in red; and (c) the resulting co-added image for this interval processed as discussed in the text. Also shown are the time of the observation and the ML value.

scope Imaging Spectrograph (STIS) onboard HST, indicating many of the principle auroral features, is shown in Fig. 1a.

The behaviour of these various individual auroral components has been discussed extensively previously, e.g. in the aforementioned studies, but it is also interesting to consider the auroral morphology as a whole, as has been explored by e.g. Clarke et al. (2009); Nichols, Clarke, Gérard, Grodent, and Hansen (2009); Nichols et al. (2017); Grodent et al. (2018), for example in relation to the response of the magnetosphere to changes in the conditions in the interplanetary medium. The response of Jupiter's auroras to the solar wind is complex; metrics such as auroral power are of limited use. For example, while the auroral power from some pre-defined regions (e.g. poleward of the ME on the dusk side) vary with interplanetary conditions, in general auroral power exhibits only weak correlation with interplanetary parameters (Nichols et al., 2017). Hence, a more nuanced discussion of the variation of the complex auroral morphology is warranted. Variations in Jupiter's overall auroral morphology have been discussed qualitatively; for example Grodent et al. (2018) (hereafter G18) divided the morphology of the auroras observed over the first few months of the Juno mission into six families A-F based on a qualitative description of the overall state of the auroras. Such analysis is very helpful for providing magnetospheric context for analysis of in situ spacecraft data, and a natural question arises as to whether a more objective quantitative technique can be used to identify different morphological families. The aim of this paper is to address this question using the first application of machine learning methods to the study of outer planetary auroras. Specifically, we employ Principal Component Analysis (PCA) and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) techniques to objectively identify classes of auroral morphology and determine whether there exists a relation between auroral class and interplanetary conditions.

Machine learning methods often rely on techniques to reduce the dimensionality of a problem, i.e. to reduce the number of independent features to be analysed, in order to facilitate computational tractability. Principal component analysis achieves this by decomposing a data set into an orthogonal basis set that reveals the covariance within the data (Jolliffe, 2002). Hence, a data set of images (classically, pictures of faces for application to facial recognition) can be reduced from a series of independent pixels to a much smaller subspace of principal components that represent most of the variation within the images (Turk & Pentland, 1991). The PCA technique has been recently employed to study a number of different aspects of the terrestrial magnetosphere (Kim et al., 2012; Cousins et al., 2015; Milan et al., 2015, e.g.). Images of the Earth's auroras have been studied using supervised deep learning classifiers (e.g. Clausen & Nickisch, 2018; Yang et al., 2019), and while that paradigm was useful for that much larger data set, here we employ the unsupervised DBSCAN clustering algorithm (Ester et al., 1996) to identify clusters within the projections provided by PCA. We find that this classifier successfully identifies repeatable morphological classes, and we associate one with solar wind compression regions. This presents an important new tool for diagnosing magnetospheric compressions at Jupiter.

2 Data and Analysis

2.1 Hubble Space Telescope Data

We consider HST/STIS images of Jupiter's northern auroras, initially focusing on those images obtained during the Juno approach phase in 2016 as the interval for which there exists an extended set of accompanying interplanetary observations against which to compare the auroral classes. This program and the data reduction has been discussed previously (Nichols et al., 2017), and an example unprojected image extracted from the timetag data with a 100 s integration time is shown in Fig. 1a. The corresponding Lambert equal area azimuthal map projection, as viewed with 180° System-III longitude oriented toward the bottom is shown in Fig. 1b, along with the Nichols et al. (2017) statistical main oval shown in red. The PCA technique requires input vectors to be independent of one another. While the polar auroras in particular exhibit changes on timescale of tens of seconds (e.g. Grodent, Clarke, Kim, et al., 2003), the auroral morphology as a whole is often broadly unchanged during the course of one 45 minute period of visibility during each HST orbit. The morphology does, however, change from one Earth day to the next, representing over 2 jovian rotations. In order to both remove noise due to short term variability and produce independent images, we co-add projected images extracted with 100 s time resolution to build up average intensity maps for each orbit. The planet rotates during each 45-min exposure, such that parts of the auroral region rotate into or out of view. In order to analyse as much auroral region as possible while avoiding the introduction of artefacts owing to this changing visibility, we rotate each projection by 23° westward such that the most equatorward extent of the oval (around 160° System-III) is toward the bottom, modestly clip the top edge of each projected image, and only employ those images with central meridian longitudes (CML) between 140° and 180°. However, we note that this CML criterion imposes a rather strict CML bias in our image selection, and while for simplicity below we discuss features that are 'dawnward' or 'duskward' it should be borne in mind that such features are also present over a limited range of longitudes in these images. The overall aim of the PCA technique is to reduce the dimensionality of the problem. Each clipped map projection grid comprises 432×240 pixels of size $\sim$ 140 km, thus consisting of a total of 103,680 elements, or 'features'. Because we are interested in the broad morphology of each image we can reduce the number of elements even before applying PCA by simply rebinning each average image with a factor of 2 reduction in the number of pixels on each axis, such that the rebinned pixels are ~ 280 km in size. This reduces the dimensionality by a factor of 4 without losing much information regarding the broad auroral morphology. The resulting co-added image for the example shown in Fig. 1a,b is shown in Fig. 1c, and the overall result is a set of 29 (defining $m = 29$) images of size 216×120 (thus with $n = 25,920$ elements).

2.2 Principal Component Analysis

Each image is flattened to form an n-dimensional vector $\mathbf{I}$. It is standard practice in machine learning applications to mean-centre and normalise the data set in some form. For each vector $\mathbf{I}$ we subtract the mean image and divide by the image standard deviation. All vectors are then stacked to form a two dimensional $n \times m$ matrix $\mathbf{X}$. The covariance matrix $\mathbf{\Sigma}$ of $\mathbf{X}$ is then calculated as $\mathbf{\Sigma} = \frac{1}{m}\mathbf{X}^{\mathrm{T}}\mathbf{X}$, where $\mathbf{X}^{\mathrm{T}}$ is the transpose of $\mathbf{X}$. Eigendecomposition of the covariance matrix $\mathbf{\Sigma}$ is then performed, yielding (for $m < n$) m n-element eigenvectors $\mathbf{A}_i$ and their corresponding eignenvalues λ_i. The eigenvectors $\mathbf{A}_i$, termed eigenimages, are the principal components of the data set, with those corresponding to the largest eigenvalues describing the directions in n-space which contain the greatest variation in the data set. Physically, they describe morphological features consistently present in the data set. The 16 most significant eigenimages are shown in Fig. 2, while the proportion of the variance explained, given by $\lambda_i / \sum_{j=1}^{m} \lambda_j$, is shown by the circles in Fig. 3. The cumulative variation explained is shown by the solid line. It is evident that the first eigenimage $\mathbf{A}_1$ corresponds to $\sim$21% of the variation in the data set, while the first 11 eigenimages together explain around $\sim$80% of the variation. A commonly-used criterion to determine how many components to keep is the Scree test (Cattell, 1966), which retains any eigenvector whose eigenvalue rises above a straight line fitted to the lower eigenvalues, as shown by the dashed line in Fig. 3. In our case the first ~ 11 eigenvalues should be considered significant.

Turning back to the eigenimages shown in Fig. 2, we first note that the plots represent variation from the mean by either red (positive) or blue (negative) colours, and the contribution of each eigenimage to any given image may be either positive or negative. It is evident first that around a fifth of the variance in the data set is explained by emission poleward of the statistical oval on the dawn side (for positive contribution), or a lack of emission in that region relative to the main emission on the dusk side (for negative contribution). Hence, an image with a positive $\mathbf{A}_1$ contribution exhibits emis-

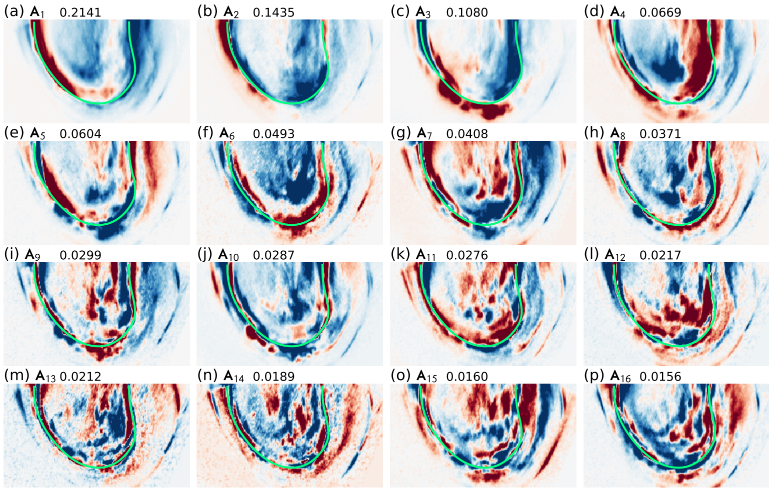

Figure 2. Plot showing the first 16 eigenimages labelled $\mathbf{A}_i$ along with the corresponding proportion of the variance explained. Red and blue colours show positive and negative values, respectively. The statistical oval is shown in green.

sion either moved or expanded poleward on the dawn side, and an image with negative $\mathbf{A}_1$ would present a well-defined dark polar region. To a lesser extent, eigenimage $\mathbf{A}_1$ contributes poleward patchy emission, either in the active region near noon and down the dusk side for positive contribution or further poleward for negative, along with an equatorward arc on the dusk side. Eigenimage $\mathbf{A}_2$ contributes emission localised on the dawn side mostly equatorward of the statistical oval for positive contribution, and more longitudinally extended emission on the poleward side of the statistical oval plus poleward emission on the dusk side for negative. Eigenimage $\mathbf{A}_3$ contributes mostly equatorward patchy emission (positive) or emission on or poleward of the main emission (negative). We finally highlight eigenimage $\mathbf{A}_4$, which contains patchy emission on or poleward of the statistical oval on the dusk side (postive) or a patch of emission significantly poleward toward noon (negative). In all four of these eigenimages there is a difference in the sign or magnitude of the contribution to the dawn and dusk sides of the poleward auroras, indicating independence of behaviour, thus possibly suggesting an asymmetric auroral response to driving mechanisms. Further eigenimages contribute smaller scale features that are increasingly less significant to the overall data set, such as patchy emis-

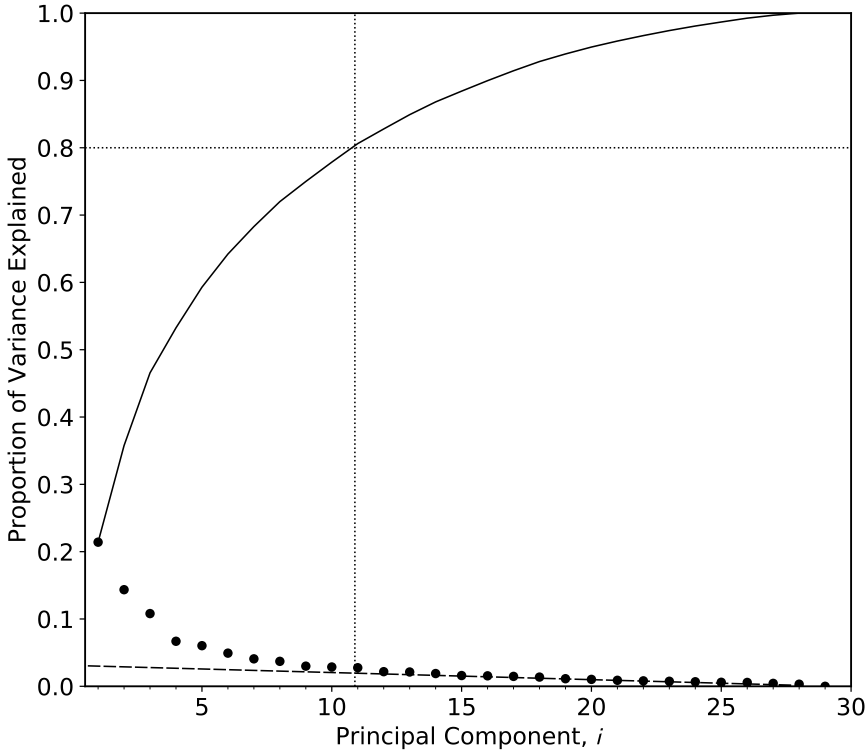

Figure 3. Plot showing the proportion of the variance explained for each component (circles). The cumulative variance explained is shown by the solid line, while the dotted lines indicate the 80% explained level and corresponding component number. The dashed line indicates a Scree test line.

sions in the swirl region, though many contain contributions associated with the main emission and are related to variation of the morphology of this auroral component.

Individual auroral images $\mathbf{I}$ comprise projections α_i along the orthonormal basis set $\mathbf{A}_i$, and hence can be expressed by

$$\mathbf{I} = \sum_{j=1}^{m} \alpha_j \mathbf{A}_j \quad . \tag{1}$$

Each image is then associated with a set of real numbers α_i which can be used to determine how the auroral morphology changes with e.g. interplanetary conditions and to classify the images by clustering in α-subspace. The variation of the first 8 projections with time are shown in Fig. 4a-h. The magnitude of each value of α_i is shown, with positive values plotted as red crosses, and negative values as blue pluses. Connecting lines are shown to guide the eye, and the colours indicate the interplanetary conditions as determined from inbound Juno data by Nichols et al. (2017) using their colour scheme, i.e. yellow indicates a deep solar wind rarefaction, cyan indicates a shallow rarefaction and blue represents solar wind compression with cause (coronal mass ejection, CME, or corotating interaction region, CIR) as labelled. Also indicated at the end of the interval is the period identified as a strong solar wind compression by Hospodarsky et al. (2017) from outbound Juno magnetopause and bow shock crossings, indeed observed roughly one solar rotation after the previous compression. The period is bound by the time of the first outbound magnetopause crossing but the compression event likely started a few hours previous to this owing to the timescale for motion of the magnetopause at a fraction of the speed of the solar wind (Cowley et al., 2007). Grey indicates that Juno was in the magnetosphere, such that the interplanetary conditions were not measured in situ during these intervals. It is first evident that all projections α_i exhibit significant variability with time, and for brevity we will explicitly discuss here the first 4 and go on to consider clustering of all these values below. Considering first α_1 shown in Fig. 4a, during the interval for which there are measurements of the interplanetary conditions, there are 8 positive values indicating main emission expanded poleward of the statistical oval on the dawn side, of which 7 are in compressions. Negative α_1 indicating a coherent dark polar region is evident at varying levels in both compressions and rarefactions. Projection α_2 shown in Fig. 4b, is strongly peaked on days 142 and 154, both days noted by Nichols et al. (2017) as exhibiting dawn storms, such that projection α_2 is evidently a sensor for dawn storms. Projection α_3 is strongly dominated by the presence of patchy

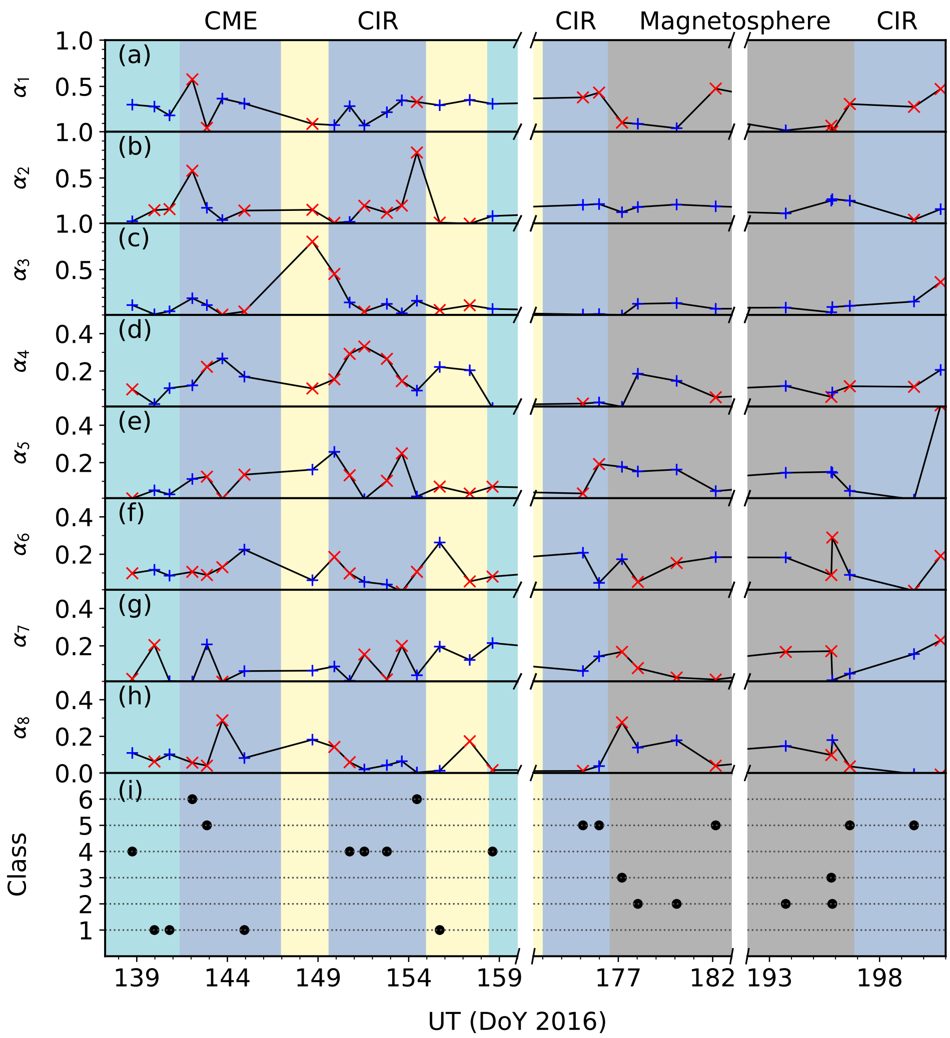

Figure 4. Plot showing (a-h) the projections α_i of the images along the basis set $\mathbf{A}_i$ as a function of Day of Year (DoY). Red crosses indicate positive values and blue pluses show negative values. Panel (i) show the repeatable image class identifications. The coloured background indicate interplanetary conditions as described in the text.

equatorward emission around day 149. Projection α_4 exhibits variation, but a clear feature is a positive peak over the interval of the second compression, indicating enhanced poleward emission on the dusk side during this event. Negative values of α_4 indicating enhanced poleward emission near noon occur in all solar wind conditions. Together, these findings are consistent with the qualitative descriptions of the behaviour the auroral morphological response to the solar wind described by Clarke et al. (2009); Nichols, Clarke, Gérard, Grodent, and Hansen (2009); Nichols et al. (2017).

2.3 Image Classification

The above application of PCA to the auroral data set has reduced the dimensionality of each image from n $(= 25,920)$ to $\sim$11 significant values of α_i. These projections can be used to classify the images. Machine learning classifiers fall into two categories: supervised, meaning the algorithm is trained using a data set labelled by a human, and unsupervised meaning the algorithm is not given a labelled training set. The aim of this study is to provide an objective classification of repeatable auroral morphology, such that we use an unsupervised algorithm, specifically Density-Based Spatial Clustering of Applications with Noise (DBSCAN)(Ester et al., 1996). The DBSCAN algorithm requires two hyperparameters, which are the minimum number of points in a cluster $N_{\min}$ and a (here Euclidian) distance parameter ε, specifying cluster density. Briefly, core points are defined as being surrounded by at least $N_{\min}$ points within distance ε, and reachable points are connected to core points via unbroken paths between points of no longer than ε. Core and reachable points are defined as being within a cluster, and all other points are considered noise. In this exploratory study with a small number of samples, we consider a cluster to contain ≥ 2 images, for which we note the algorithm is then equivalent to a hierarchical clustering algorithm. Parameter ε can be chosen freely, and, as is standard in machine learning applications, the optimal value is obtained via a hyperparameter grid search. We executed the clustering algorithm on the 11 significant projections α_i using values of ε between 0.1 and 1 with 0.05 increments, and have adopted the value which yields the most clusters, i.e. $\varepsilon = 0.40$. This results in 6 clusters, which we identify with image classes, and 5 noise points.

The classes thus defined for each image are shown in Fig. 4i, numbered 1-6, and the normalised images $\mathbf{I}$ grouped by class are shown in Fig. 5. A prototype image $\mathbf{I}$ for each class, obtained using the mean values of α_i for each class in Eq. 1 with $j = 1 \ldots 11$ is shown in Fig. 6. Qualitatively, Class 1 shown in Figs. 5a and 6a exhibits broad, low latitude, dim-to-medium intensity ME on the dawn side, brighter and narrower ME on the dusk side, and modestly active polar emission separated from the main emission by a wide, well-defined dark polar region, along with an equatorward arc on the dusk side. Image Class 2 shown in Figs. 5b and 6b exhibits narrow, brighter ME on the dawn side with bright patches along the post-noon ME, and some polar emission. Class 3 in Figs. 5c and 6c is characterised by narrow, brighter ME on the dawn side, relatively bright polar emission on the dusk side. The principal feature of Class 4 shown in Figs. 5d and 6d

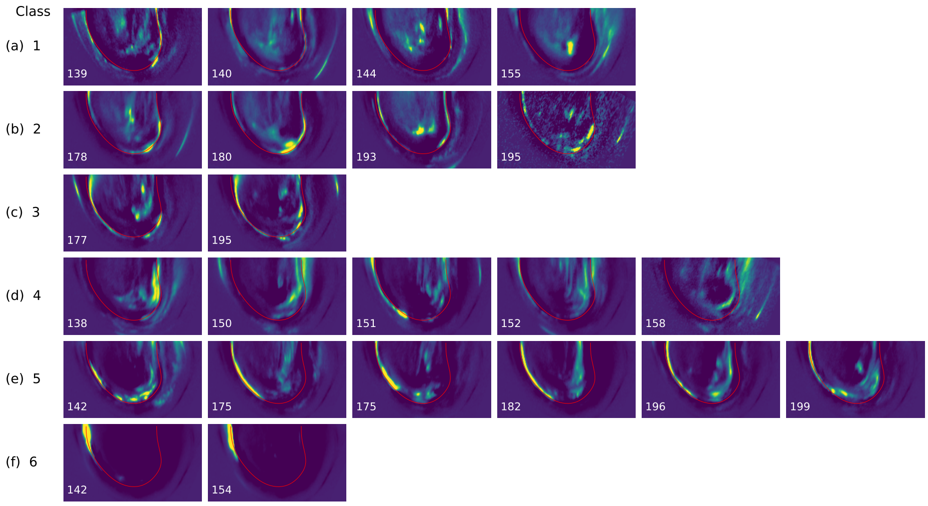

Figure 5. Plot showing the individual images grouped into the identified morphological classes as labelled. The DoY is shown for each image.

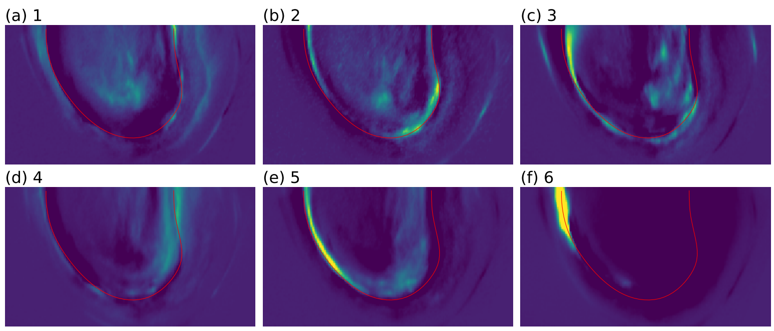

Figure 6. Plot showing class prototype images created using the mean projection values α_i for each cluster as labelled.

is relatively bright emission on and poleward of the ME on the dusk side. Image Class 5 shown in Figs. 5e and 6e exhibits very bright emission along the length of the dawn-side ME, sightly poleward of the statistical oval, and emission throughout the dusk side polar region. Finally, Class 6 shown in Figs. 5f and 6f is a dawn storm exhibiting extremely bright and expanded ME on the dawn side. It is worth noting that a number of these classes can be mapped roughly onto G18's families, i.e. we suggest Class 1 corresponds most closely to G18's family A, our Class 2 to G18's family B, our Class 4 to G18's family F, and our Class 5 to G18's families C or E. The classification system thus successfully identifies a variety of repeatable morphologies.

We consider now the occurrence of these image classes over time and with interplanetary conditions, as shown in Fig. 4i. We note that the classes are broadly distributed over the observing interval, such that even the two classes (3 and 6) comprising only two images are not contiguous in the observing sequence, indicating recurring morphologies. During the intervals in which interplanetary data is available, the classes that occur in both compression and rarefactions are Classes 1 and 4, while the classes that occur solely in compressions are Classes 5 and 6. Dawn storms are known to occur independently of the solar wind, such that the only class which is not a dawn storm and which occurs solely during interplanetary compressions is Class 5. Given the time scale for solar wind compression of the magnetopause, it is likely that the image obtained on DoY 196, also identified as Class 5, was also obtained under solar wind compression conditions. Classes 2 and 3 occur only when Juno was in the magnetosphere, such that their association with solar wind conditions is unknown. However, both occurences of Class 3 were close to observed compression conditions and it is hence possible, though not conclusive, that this class is also associated with solar wind compressions.

To assess this association of Class 5 with interplanetary compressions, we have also examined the only other HST data set that satisfies our CML criteria and for which an extended concurrent interplanetary data is available, i.e. that obtained with the Advanced Camera for Surveys during the New Horizons flyby in 2007 (Clarke et al., 2009; Nichols, Clarke, Gérard, Grodent, & Hansen, 2009). During that interval, New Horizons observed the forward shock of a compression on DoY 53, and entered the magnetosphere shortly after on DoY 56. An MHD model of the projected solar wind (Zieger & Hansen, 2008) predicted a second forward shock to occur between DoY 63 and 66. We projected the HST images obtained in this program onto the basis set $\mathbf{A}_i$ obtained above and then identified the images as belonging to a particular class if the Euclidian distance between the resulting α_i and the class mean was less than ε. Example results are shown in Fig. 7a-f, which are evidently qualitatively similar to the prototypes shown in Fig. 6. We also show in Fig. 7g the classes plotted versus UT as in Fig. 4i for the interval with New Horizons data, along with an indication of the solar wind conditions using a similar colour scheme. The time of the observed solar wind forward shock and magnetopause crossing are also labelled, along with an indication of the uncertainty interval of the arrival of the second forward shock as discussed by Nichols, Clarke, Gérard, Grodent, and Hansen (2009) (lighter blue region labelled FS2). It is worth noting that the only occurrences of Class

5 are in the first compression or not long after New Horizons left the solar wind while it was under compression conditions, or in the interval during which the second forward shock was expected to impinge on the planet. We thus conclude that identification of an auroral image as Class 5 is a satisfactory diagnostic of solar wind compressions. Class 3 also occurs during the second forward shock interval, consistent with the above discussion regarding its possible association with the solar wind, but again not conclusive.

3 Conclusions

We have presented the first application of machine learning techniques to the study of outer planetary auroral emissions, and have examined their response to interplanetary conditions. We used Principal Component Analysis to show that the most recurrent source of variance of Jupiter's auroral emission is aurora (or the lack of it) poleward of the statistical oval on the dawn side. Further significant repeatable sources of variance are localised expansions of the ME on the dawn or dusk sides and poleward emission near noon and along the dusk side. The dawn and dusk sides of the poleward auroral emission evidently vary independently, suggestive of an asymmetric response to driving mechanisms. The individual identified components respond differently to interplanetary conditions, e.g. of the 8 occasions when the most significant component contributes poleward-expanded main emission on the dawn side, and poleward emission in the active region and along the dusk side, 7 are during solar wind compressions. A component contributing significant poleward dusk emission also strongly peaks during a solar wind compression. We then showed using a DBSCAN clustering analysis that, together, the most significant components form 6 repeatable auroral morphological classes, each with a different pattern of auroral intensities. For example, we identified one morphological class (6) with dawn storms, and a further class (5) with solar wind compressions. This class, which presents very bright, modestly poleward emission along the length of the dawnside main emission and emission throughout the dusk side polar region is the only (non-dawn storm) class to be observed solely during interplanetary compressions. This does not preclude the occurence of other classes during compressions, but it does strongly suggest that the occurence of this particular morphological class is indicative of a compressed magnetosphere. We tested this assertion using HST observations obtained during the New Horizons flyby and found this class only occurred either during a compression or within the uncertainty of a compression region onset. This study thus provides

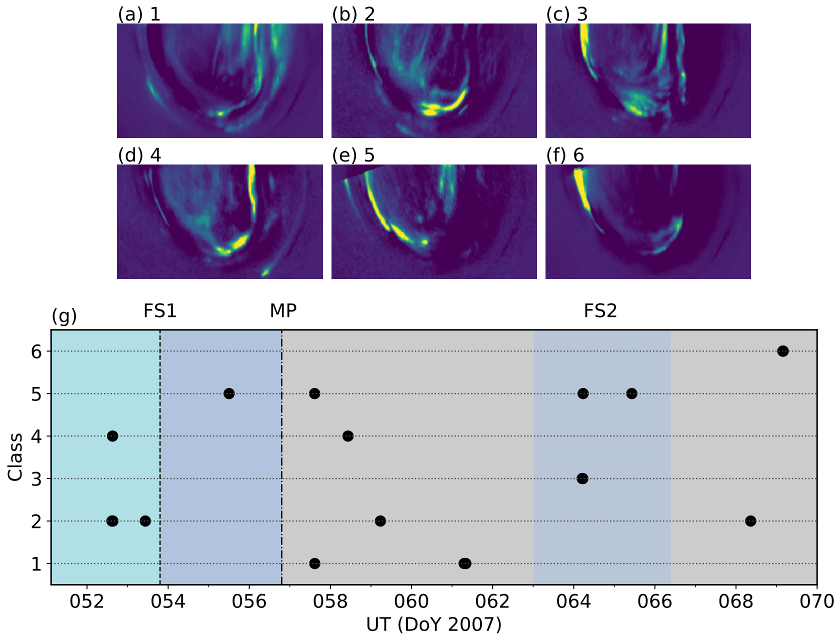

Figure 7. Plot showing (a-f) example images from the 2007 HST program identified for each image class, and (g) the image classes identified versus DoY in 2007 for the New Horizons interval. The solar wind rarefaction interval is indicated in cyan. The vertical dashed line labelled FS1 indicates the time of observation of a forward shock, the dot-dashed line labelled MP indicates where New Horizons entered the magnetosphere, and the blue in between indicates the observed solar wind compression. The grey region indicates where New Horizons was in the magnetosphere, and the lighter blue region labelled FS2 indicates the uncertainty interval for the arrival of a second forward shock.

a proof-of-concept that such machine learning techniques are a useful new tool for diagnosing solar wind conditions at Jupiter, and analysis of the morphology of Jupiter's complex auroras and their response to magnetospheric drivers. In a future study we plan to apply such methods to the much larger data set obtained during the Juno mission.

Acknowledgments

This work is based on observations made with the NASA/ESA Hubble Space Telescope (program GO 14105), obtained at STScI, which is operated by AURA, Inc. for NASA. This work was supported by STFC Consolidated Grant ST/N000749/1. HST data are available at the MAST Archive.

References

Bonfond, B., Grodent, D., Gérard, J.-C., Radioti, A., Saur, J., & Jacobsen, S. (2008, March). UV Io footprint leading spot: A key feature for understanding the UV Io footprint multiplicity? *Geophys. Res. Lett.*, *35*(5), L05107.

Bonfond, B., Saur, J., Grodent, D., Badman, S. V., Bisikalo, D. V., Shematovich, V. I., ... Radioti, A. (2017). The tails of the satellite auroral footprints at Jupiter. *Journal of Geophysical Research: Space Physics*, *59*(1), 836.

Cattell, R. B. (1966, April). The Scree Test For The Number Of Factors. *Multivariate Behavioral Research*, *1*(2), 245–276.

Clarke, J. T., Grodent, D., Cowley, S. W. H., Bunce, E. J., Zarka, P. M., Connerney, J. E. P., & Satoh, T. (2004). Jupiter's aurora. In F. Bagenal & W. B. McKinnon (Eds.), *Jupiter: The planet, satellites and magnetosphere* (pp. 639–670). Cambridge, UK: Cambridge. Univ. Press.

Clarke, J. T., Nichols, J. D., Gérard, J.-C., Grodent, D., Hansen, K. C., Kurth, W. S., ... Cecconi, B. (2009, May). Response of Jupiter's and Saturn's auroral activity to the solar wind. *J. Geophys. Res.*, *114*(A), A05210.

Clausen, L. B. N., & Nickisch, H. (2018, July). Automatic Classification of Auroral Images From the Oslo Auroral THEMIS (OATH) Data Set Using Machine Learning. *Journal of Geophysical Research: Space Physics*, *123*(7), 5640–5647.

Cousins, E., Matsuo, T., AD, R., & J., A. B. (2015). Dominant modes of variability in large-scale Birkeland currents. *J. Geophys. Res.*, *120*, 6722–6735.

Cowley, S. W. H., Nichols, J. D., & Andrews, D. J. (2007, June). Modulation of Jupiter's plasma flow, polar currents, and auroral precipitation by solar wind-induced compressions and expansions of the magnetosphere: a simple theoretical model. *Ann. Geophysicae*, *25*, 1433–1463.

Dumont, M., Grodent, D., Radioti, A., Bonfond, B., & Gérard, J.-C. (2015, January). Jupiter's equatorward auroral features: Possible signatures of magnetospheric injections. *J. Geophys. Res.*, *119*(12), 10,068-10,077.

Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the 2nd international conference on knowledge discovery and data mining* (p. 226-231). Portland, OR: AAAI Press.

Gray, R. L., Badman, S. V., & Bonfond, B. (2016). Auroral evidence of radial transport at Jupiter during January 2014. *J. Geophys. Res.*, *121*(10), 9972–9984.

Grodent, D., Bonfond, B., Yao, Z., Gérard, J.-C., Radioti, A., Dumont, M., ... Valek, P. (2018, May). Jupiter's Aurora Observed With HST During Juno Orbits 3 to 7. *Journal of Geophysical Research: Space Physics*, *123*(5), 3299–3319.

Grodent, D., Clarke, J. T., Kim, J., Waite Jr, J. H., & Cowley, S. W. H. (2003, November). Jupiter's main auroral oval observed with HST-STIS. *J. Geophys. Res.*, *108*(A11), 1389.

Grodent, D., Clarke, J. T., Waite Jr, J. H., Cowley, S. W. H., Gérard, J.-C., & Kim, J. (2003, October). Jupiter's polar auroral emissions. *J. Geophys. Res.*, *108*(A10), 1366.

Hospodarsky, G. B., Kurth, W. S., Bolton, S. J., Allegrini, F., Clark, G. B., Connerney, J. E. P., ... Valek, P. W. (2017, May). Jovian bow shock and magnetopause encounters by the Juno spacecraft. *Geophys. Res. Lett.*, *44*(10), 4506–4512.

Jolliffe, I. T. (2002). *Principal component analysis* (2nd ed.). New York: Springer.

Kim, H. J., Lyons, L. R., Ruohoniemi, J. M., Frissell, N. A., & Baker, J. B. (2012, June). Principal component analysis of polar cap convection. *Geophys. Res. Lett.*, *39*(11), L11105.

Milan, S. E., Carter, J. A., Korth, H., & Anderson, B. J. (2015, December). Principle Component Analysis of Birkeland Currents Determined by the Active Magnetosphere and Planetary Electrodynamics Response Experiment. *American Geophysical Union*, *41*, SA41A–2318.

Nichols, J. D., Badman, S. V., Bagenal, F., Bolton, S. J., Bonfond, B., Bunce, E. J., ... Yoshikawa, I. (2017). Response of Jupiter's auroras to conditions in the interplanetary medium as measured by the Hubble Space Telescope and Juno. *Geophys. Res. Lett.*, *44*(15), 7643-7652.

Nichols, J. D., Clarke, J. T., Gérard, J.-C., & Grodent, D. (2009, April). Observations of Jovian polar auroral filaments. *Geophys. Res. Lett.*, *36*(8), L08101.

Nichols, J. D., Clarke, J. T., Gérard, J.-C., Grodent, D., & Hansen, K. C. (2009). Variation of different components of Jupiter's auroral emission. *J. Geophys. Res.*, *114*(A6), A06210.

Radioti, A., Tomas, A. T., Grodent, D., Gérard, J.-C., Gustin, J., Bonfond, B.,

... Menietti, J. D. (2009, April). Equatorward diffuse auroral emissions at Jupiter: Simultaneous HST and Galileo observations. *Geophys. Res. Lett.*, *36*(7), L07101.

Turk, M., & Pentland, A. (1991). Eigenfaces for Recognition. *Journal of Cognitive Neuroscience*, *3*(1), 71–86.

Yang, Q., Tao, D., Han, D., & Liang, J. (2019, 5). Extracting auroral key local structures from all-sky auroral images by artificial intelligence technique. *Journal of Geophysical Research: Space Physics*, *124*(5), 3512–3521. Retrieved from `https://doi.org/10.1029/2018JA026119` doi: 10.1029/2018JA026119

Zieger, B., & Hansen, K. C. (2008, August). Statistical validation of a solar wind propagation model from 1 to 10 AU. *J. Geophys. Res.*, *113*(A12), 8107.