AGU PUBLICATIONS

Space Weather

RESEARCH ARTICLE
10.1002/2015SW001320

Key Points:
- Citizen science reports can be used to determine extent of auroral visibility
- When compared with observations, an updated SWPC view line is somewhat conservative
- New, more accurate, view lines are determined using citizen science data

Correspondence to:
N. A. Case,
n.case@lancaster.ac.uk

Citation:
Case, N. A., E. A. MacDonald, and R. Viereck (2016), Using citizen science reports to define the equatorial extent of auroral visibility, *Space Weather*, *14*, 198–209, doi:10.1002/2015SW001320.

Received 5 OCT 2015
Accepted 12 FEB 2016
Accepted article online 16 FEB 2016
Published online 3 MAR 2016

Using citizen science reports to define the equatorial extent of auroral visibility

N. A. Case[1,2,3], E. A. MacDonald[1,2], and R. Viereck[4]

[1]New Mexico Consortium, Los Alamos, New Mexico, USA, [2]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA, [3]Department of Physics, Lancaster University, Lancaster, UK, [4]NOAA Space Weather Prediction Center, Boulder, Colorado, USA

Abstract An aurora may often be viewed hundreds of kilometers equatorward of the auroral oval owing to its altitude. As such, the NOAA Space Weather Prediction Center (SWPC) Aurora Forecast product provides a "view line" to demonstrate the equatorial extent of auroral visibility, assuming that it is sufficiently bright and high in altitude. The view line in the SWPC product is based upon the latitude of the brightest aurora, for each hemisphere, as specified by the real-time oval variation, assessment, tracking, intensity, and online nowcasting (OVATION) Prime (2010) aurora precipitation model. In this study, we utilize nearly 500 citizen science auroral reports to compare with the view line provided by an updated SWPC aurora forecast product using auroral precipitation data from OVATION Prime (2013). The citizen science observations were recorded during March and April 2015 using the Aurorasaurus platform and cover one large geomagnetic storm and several smaller events. We find that this updated SWPC view line is conservative in its estimate and that the aurora is often viewable further equatorward than is indicated by the forecast. By using the citizen reports to modify the scaling parameters used to link the OVATION Prime (2013) model to the view line, we produce a new view line estimate that more accurately represents the equatorial extent of visible aurora. An OVATION Prime (2013) energy flux-based equatorial boundary view line is also developed and is found to provide the best overall agreement with the citizen science reports, with an accuracy of 91%.

1. Introduction

Knowing when, and from where, an aurora will be visible is an aspect of space weather science that interests researchers and the general public alike. In fact, one of the only ways the general public can experience space weather firsthand is to witness an aurora and there is a small, but growing, tourism industry catering to people who want to do just that. Several auroral precipitation models, based either upon current solar wind conditions or estimated real-time geomagnetic indices, exist which can aid in this regard by predicting the size and location of the auroral oval [e.g., *Spiro et al.*, 1982; *Hardy et al.*, 1985, 1989; *Roble and Ridley*, 1987; *Zhang and Paxton*, 2008; *Newell et al.*, 2010b, 2014; *Mitchell et al.*, 2013].

Oval variation, assessment, tracking, intensity, and online nowcasting (OVATION) Prime is one such auroral precipitation model [*Newell et al.*, 2010b, 2014]. This particular model is driven by the rate of delivery of interplanetary magnetic flux to Earth's magnetopause as parameterized by the $d\Phi_{MP}/dt$ magnetospheric coupling function [*Newell et al.*, 2007]. This coupling function is in turn dependent upon the solar wind conditions as measured at Earth's first Lagrangian orbital point (L1), located approximately 1 million miles (1.5×10^6 km) upstream on the Sun-Earth line. Since the coupling function is solar wind driven, the model can be run in real time using upstream solar wind data. This real-time run ability makes it especially useful for space weather forecasting as the model typically provides a 30–40 min forecast of the overall size and intensity of the auroral ovals.

When calculating the auroral precipitation, OVATION Prime accounts for several different auroral types (i.e., diffuse aurora, monoenergetic, broadband, and ion), seasonal variations, and the magnetic latitude (MLAT) and magnetic local time (MLT) of each of its modeled bins (which are 0.5° in MLAT by 0.25°, or 1 min, in MLT in size). All of which ensures that it is often more accurate at modeling the auroral oval than other real-time models [*Newell et al.*, 2014; *Lane et al.*, 2015] and can reliably predict when an aurora will be visible [*Machol et al.*, 2012].

The original version of the model, referred to as OVATION Prime (2010) [*Newell et al.*, 2010b], has been further developed to increase its accuracy at larger geomagnetic activity levels (particularly in the *Kp* 5+ to 8+ range) and to reduce the noise between neighboring bins. This latest version of the model is known as OVATION Prime (2013) [*Newell et al.*, 2014]. OVATION Prime (2010) is subsequently referred to as OP10 and OVATION Prime (2013) as OP13 throughout.

NOAA's Space Weather Prediction Center (SWPC) has been running OP10 in real time since 2011. As described in the following Methods section, SWPC uses the model to provide auroral precipitation data for their public-orientated aurora forecast product. Included in this forecast product is an estimate of the most equatorward latitude, for both hemispheres, from which an aurora might be visible, known as a "view line."

While both OP10 and OP13 have been validated and found to provide reasonable estimates of the location and intensity of the aurora [e.g., *Machol et al.*, 2012; *Newell et al.*, 2014; *Lane et al.*, 2015], no extensive testing has yet been performed on the accuracy of the SWPC forecast product or the location of the view line.

In this study, nearly 500 citizen science reports, collected by the Aurorasaurus project [*MacDonald et al.*, 2015], are evaluated and compared against an updated version of the SWPC view line which is based on OP13 data. Furthermore, the reports are then used to create an observationally based Aurorasaurus view line and to test an auroral equatorial boundary-based view line.

We note one previous related effort by *Gartlein and Moore* [1951] showed that a decade long network of dedicated observers could track the southern extent of the aurora over North America. The work contained herein can be seen as an extension to this, in which real-time, globally distributed reports are compared to modern models for auroral extent, something that was not possible for *Gartlein and Moore* [1951].

2. Methods

As previously mentioned, SWPC uses the real-time data output from OP10 to drive its aurora forecast product. The model output is converted from geomagnetic coordinates into geographic and is then resampled into an array of 1024 bins in longitude and 512 in latitude (i.e., each bin is approximately 0.35° in latitude and longitude).

The precipitating energy flux of each auroral type, excluding ion precipitation, is summed and converted into an empirical estimate of the "probability of visible aurora." The ion portion of energy flux is excluded since, generally, ion precipitation does not contribute to the visible aurora in the traditional auroral oval. We note that ion (or proton) precipitation does, however, contribute to subvisual auroral structures, including proton aurora [*Donovan et al.*, 2012] and stable auroral red arcs that form at midlatitudes [*Baumgardner et al.*, 2008].

As shown in equation (1), the purely empirical conversion between the summed precipitation energy flux, Σj (measured in erg cm^{-2} s^{-1}), and the percentage probability of visible aurora, $P(A)$, involves simply scaling the energy flux for each bin and adding an offset.

$$P(A) = 10 + 8\Sigma j \tag{1}$$

The resultant probability values are then smoothed, small values are clipped, and an upper limit of 100% is applied. A text file containing these gridded $P(A)$ values is made available to download in real time from the SWPC website.

Additionally, since an aurora can be viewed, especially during high activity, hundreds of kilometers equatorward of the visible auroral oval, a coarse view line is estimated. The view line indicates the most equatorward latitude, for a range of longitudes, from which an observer might be able to see an aurora (i.e., at latitudes poleward of this line, an aurora should be visible).

The SWPC view line is determined independently for both hemispheres. Each of the 1024 geographic longitudinal arrays (spaced at 0.35° intervals) are split by hemisphere, and the maximum probability of visible aurora ($P(A)_{\max}$) in that longitudinal hemispheric array is determined. The latitude of the most equatorward bin, in that array, containing this maximum probability is then found ($\phi_{P(A)\max}$).

As shown in equation (2), $\phi_{P(A)\max}$ is scaled equatorward, by a factor dependent upon $P(A)_{\max}$, to give the view line latitude ($\phi_{\mathrm{VL}}^{\mathrm{SWPC}}$) for that specific longitude and hemisphere.

$$\phi_{\mathrm{VL}}^{\mathrm{SWPC}} = \phi_{P(A)\max} \pm \left(\frac{P(A)_{\max}}{20} + 3 \right) \tag{2}$$

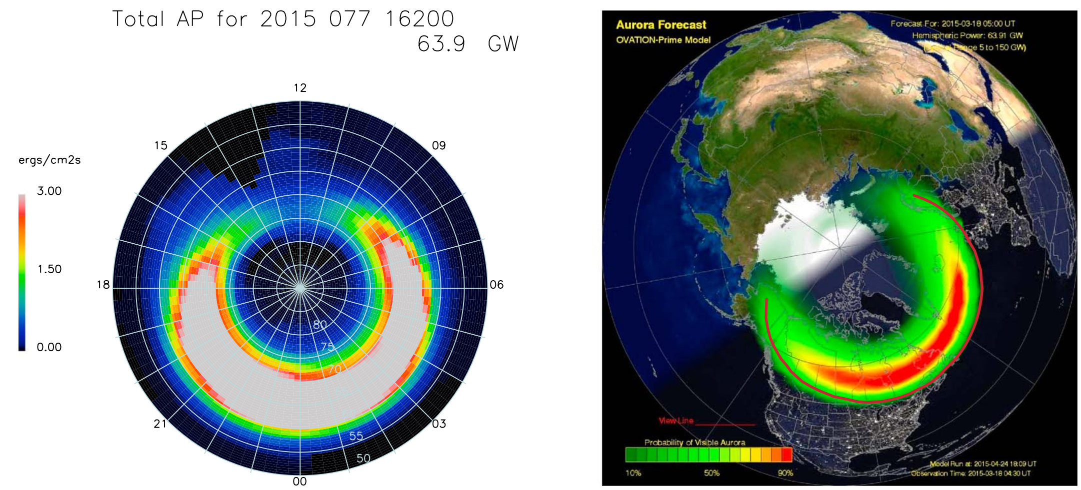

Figure 1. (left) The OVATION Prime (2013) auroral precipitation output, in terms of energy flux, in geomagnetic coordinates (04:30 UT on 18 March 2015). (right) The updated SWPC output, in terms of visible aurora probability, in geographic coordinates using the same OP13 data.

Note that the ± sign is required to scale the latitude equatorward for both hemispheres (when using a ±90° range of latitudes). A five element smoothing function is applied to the view line, and the view line is then clipped at the day/night terminators.

In this study, the official SWPC aurora forecast product has been updated so that it uses OP13 data, rather than OP10. While this may sometimes result in a slightly different view line than is provided by the official product on the SWPC website, it should ensure that the output is more accurate—especially during strong auroral displays. No other changes to the SWPC product have been made. This updated OP13 version of the SWPC aurora forecast product is herein referred to as the "updated SWPC product."

An example of the OP13 output and the corresponding updated SWPC product, including the view line determined using equation (2), is shown in Figure 1.

3. Citizen Science Data

Used in this study are 494 Aurorasaurus citizen science reports which were recorded during March and April 2015. These months were selected because they encompassed several periods of high geomagnetic activity, including a severe G4-level storm on the NOAA geomagnetic storm scale [*Poppe*, 2000]. During such intense geomagnetic activity the number of aurora observations, perhaps unsurprisingly, increases significantly allowing for larger statistical analyses [*Case et al.*, 2015b]. Covering larger storms is also particularly useful since most auroral models lack observational data during periods of high activity (owing to their relative rarity).

Aurorasaurus has been collecting a standardized set of auroral visibility reports, made by the general public, since November 2014. The reports are submitted via the project's website and mobile apps and are also sourced from Twitter. These reports, which act as "ground truths" for auroral visibility, all contain a timestamp and location, and many contain extra details about the sighting such as the color, structure, or activity of the aurora. Some will also include a photograph of the sighting.

The Aurorasaurus reports are grouped into two primary categories, positive and negative, identifying if the observer was able to see an aurora. Positive reports, which make up 85% of the reports in this case study, are composed of reports submitted directly to Aurorasaurus, known as positive sightings, and reports posted on Twitter. Twitter reports, which can also provide useful information about auroral activity [*Case et al.*, 2015a], are found using keyword searching, verified by Aurorasaurus users as real-time aurora sightings and then manually verified by Aurorasaurus team members (see *MacDonald et al.* [2015], for further details). Once verified, these Twitter reports are known as verified tweets.

There are also two types of negative reports: those that did not see the aurora because their view was obstructed (e.g., cloud cover, physical obstacles/terrain, or light pollution) and those whose view was not obstructed; hence, an aurora was simply not visible at that location. It is the latter type of negative report

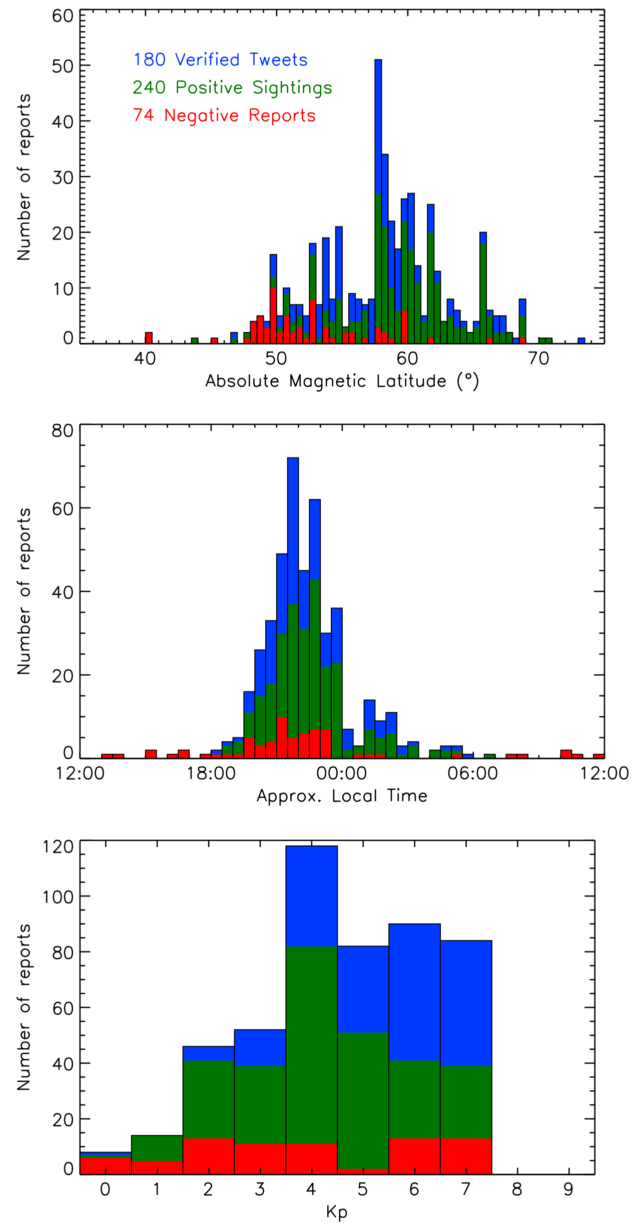

Figure 2. Histograms of the Aurorasaurus reports, submitted during March and April 2015, grouped by (top) absolute magnetic latitude (in 0.5° bins), (middle) approximate local time (30 min bins), and (bottom) *Kp* index at the start of the report. The stacked red bars indicate negative reports, the green bars indicate positive sightings, and the blue bars indicate verified positive reports posted on Twitter, known as verified tweets.

that is of interest in this study, since the view line does not take into account local conditions. Therefore, the negative reports have been filtered to those that indicated a clear, unobstructed view of the night sky.

Furthermore, all reports submitted directly to Aurorasaurus, through either its website or its mobile apps, are checked for obvious data integrity issues. For example, reports spanning more than three hours (usually a result of the user selecting an incorrect end time) are filtered out, as are any reports submitted by the same user within the same time period (i.e., multiple submissions of the same report).

4. Results

Plotted in Figure 2 (top) is the absolute magnetic latitudes of the 494 Aurorasaurus reports submitted during March and April 2015, grouped into 0.5° intervals. The stacked color bars indicate the number of each type

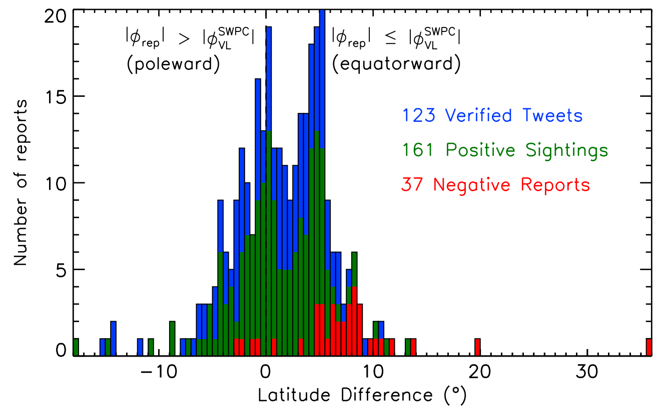

Figure 3. A histogram of the differences in latitude between the Aurorasaurus reports (ϕ_{rep}) and the updated SWPC view line (ϕ_{VL}^{SWPC}). The differences are grouped in 0.5° intervals, and the stacked bars indicate the number of each type of report in each interval. The red bars indicate negative reports, the green bars indicate a positive sightings, and the blue bars indicate verified tweets.

of report in each interval. Negative reports (of which there are 74) are red, positive sightings (240) are green, and verified tweets (180) are blue. The positive sightings and verified tweets collectively span from 43.8° to 73.2° in absolute magnetic latitude, with a median latitude of 58.6°.

In Figure 2 (middle), the reports are grouped by their estimated local time (LT) in 30 min bins. The term "estimated" is used since all reports are actually recorded in UT, and so the local time of the report is determined using this UT value and the report's longitude. Adjustments for daylight savings time are made when appropriate. The positive sightings and verified tweets collectively span from approximately 19:00 to 07:00 LT, with the median start time of 22:45 LT.

In Figure 2 (bottom), the reports are grouped by the corresponding *Kp* value at the start of the report. The *Kp* index [*Bartels et al.*, 1939], is a 3-hourly quasi-logarithmic index describing the global geomagnetic activity level, ranging from 0 to 9, and is provided by NASA's OMNIweb data archive. The corresponding *Kp* values for the Aurorasaurus reports span from 0 to 7, with a median value of 5.

4.1. Updated SWPC View Line

To investigate the accuracy of the updated SWPC view line, the difference in latitude between each report ($|\phi_{rep}|$) and the view line ($|\phi_{VL}^{SWPC}|$) is determined. Reports that are equatorward of the view line result in a positive difference, while those that are poleward result in a negative difference. A histogram of these differences in shown in Figure 3. We note that some of the reports could not be compared with the updated SWPC view line which was generally the result of invalid solar wind data resulting in no valid OP13 output.

The median difference of the positive reports, which includes positive sightings and verified tweets, is +1.26° (approximately 140 km equatorward). Though since the view line is an estimate of the most equatorward latitude from which an observer might see the aurora, in this study we are primarily interested in those positive reports that occur equatorward of the view line (i.e., $|\phi_{rep}| < |\phi_{VL}^{SWPC}|$). When filtering to those reports, which account for 62.0% of the total positive reports, the median difference is +3.70° (or approximately 400 km equatorward).

The overall accuracy (ACC) of the updated SWPC view line can be determined from the ratio of the sum of the true positives (ΣTP) and true negatives (ΣTN) to the total number of reports (ΣR) [*Machol et al.*, 2012]. Specifically,

$$\mathrm{ACC} = \frac{\Sigma\mathrm{TP} + \Sigma\mathrm{TN}}{\Sigma R} \tag{3}$$

where a true positive is a positive report that occurred on, or poleward of, the view line and a true negative is a negative report that occurred equatorward of the view line. For the updated SWPC view line, $\Sigma\mathrm{TP} = 108$, $\Sigma\mathrm{TN} = 33$, and $\Sigma R = 321$; thus, the accuracy is found to be 43.9%.

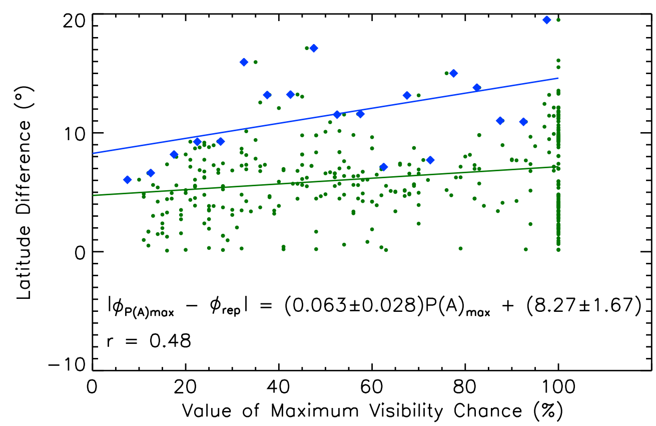

Figure 4. The latitude difference between the positive reports and the location of maximum visibility (i.e., $|\phi_{P(A)\max}| - |\phi_{\mathrm{rep}}|$), plotted as a function of the maximum visibility ($P(A)_{\max}$). The data are filtered to positive reports with positive differences (i.e., $|\phi_{P(A)\max}| > |\phi_{\mathrm{rep}}|$). The solid green line is the least squares fit through all of the data and the solid blue line is the least squares fit through the maximum difference in each 5° bin (represented by blue diamonds). The fit equation shown is for the blue line.

4.2. Aurorasaurus View Line

The validity of the SWPC view line coefficients (equation (2)) can be tested by replacing $\phi_{\mathrm{VL}}^{\mathrm{SWPC}}$ with the latitudes of the positive reports (ϕ_{rep}). Then, by rearranging equation (2) and plotting the difference between the positive reports and the location of maximum probability of visible aurora (i.e., $|\phi_{P(A)}\max| - |\phi_{\mathrm{rep}}|$) as a function of $P(A)_{\max}$, the coefficients of the fit can be determined.

Again, we are interested in the most equatorward location of the visible aurora, rather than just an average. As such, a least squares fit through all the data is not the most appropriate fit to make. Instead, a least squares fit through the maximum difference (i.e., the largest value of $|\phi_{P(A)\max}| - |\phi_{\mathrm{rep}}|$) in 5° intervals is computed. As shown in Figure 4, the line of best fit takes the form: $|\phi_{P(A)\max}| - |\phi_{\mathrm{rep}}| = (0.063 \pm 0.028)\,P(A)_{\max} + (8.27 \pm 1.67)$. A linear fit is assumed owing to the linear relationship in equation (2). The Pearson's correlation coefficient for this linear relation is $r = 0.48$.

A view line determined using the Aurorasaurus positive reports and $P(A)_{\max}$ can now be created and is given in equation (4):

$$\phi_{\mathrm{VL}}^{\mathrm{AS}} = \phi_{P(A)\max} \pm \left(\frac{P(A)_{\max}}{16} + 8\right) \tag{4}$$

In Figure 5, the difference between the reports and this new Aurorasaurus view line ($\phi_{\mathrm{VL}}^{\mathrm{AS}}$) is shown in a similar form to Figure 3. The median difference of the positive reports is now −4.55° (500 km poleward) and 92.7%

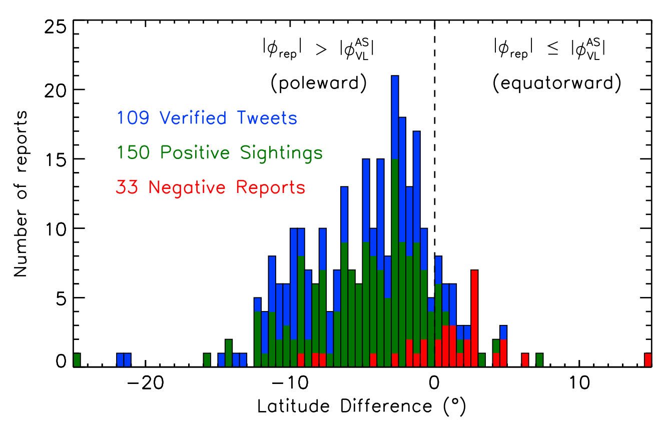

Figure 5. Of the same form as Figure 3. A histogram of the differences in latitude between the Aurorasaurus reports (ϕ_{rep}) and the Aurorasaurus view line ($\phi_{\mathrm{VL}}^{\mathrm{AS}}$).

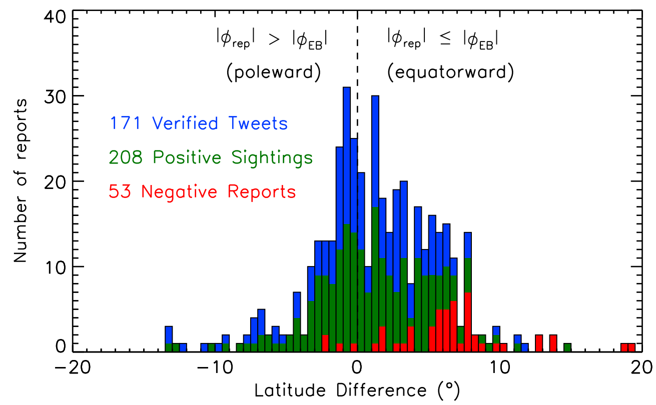

Figure 6. Of the same form as Figure 3. A histogram of the differences in latitude between the Aurorasaurus reports (ϕ_{rep}) and the auroral oval equatorial boundary (ϕ_{EB}).

of the positive reports are poleward of the view line. The accuracy of the view line, as defined in equation (3), is 90.1% ($\Sigma\mathrm{TP}=240$, $\Sigma\mathrm{TN}=23$, and $\Sigma R=292$).

4.3. Equatorial Boundary View Line

The updated SWPC view line ($\phi_{\mathrm{VL}}^{\mathrm{SWPC}}$) and the Aurorasaurus one based upon it ($\phi_{\mathrm{VL}}^{\mathrm{AS}}$) are determined using the latitude of the peak intensity of the aurora ($\phi_{P(A)\mathrm{max}}$). This leads, at times, to unrealistic situations wherein the view-line lies within a wide auroral oval (see Figure 1). A view line based upon the latitude of the equatorial boundary of the aurora, rather than the latitude of the peak intensity, is therefore investigated. In the following, the equatorial boundary is defined as the most equatorward latitude at which $P(A) \geq 18\%$, which equates to $\Sigma j \geq 1$ erg cm^{-2} s^{-1} (cf. *Machol et al.* [2012], who cite this threshold as approximately corresponding to visible aurora).

In Figure 6, the difference between the latitude of the Aurorasaurus reports (ϕ_{rep}) and the equatorial boundary of the OP13 modeled auroral oval (ϕ_{EB}) is investigated. The median difference for all positive reports is 0.62° (approximately 70 km equatorward) and when filtering to reports where $|\phi_{\mathrm{rep}}| < |\phi_{\mathrm{EB}}|$ the median difference is +3.06° (or approximately 350 km equatorward). The accuracy of using just the equatorial boundary as the view line is 49.7%.

The difference in latitude between the positive reports and the equatorial boundary, as a function of $P(A)_{\mathrm{max}}$, is shown in Figure 7. The figure takes a similar form to Figure 4, and the line of best fit through the maximums is found to be $|\phi_{\mathrm{EB}}| - |\phi_{\mathrm{rep}}| = (0.00 \pm 0.03)P(A)_{\mathrm{max}} + (7.65 \pm 2.06)$. Using this fit, a view line based upon the relationship between the positive reports and the equatorial boundary can be determined and is given in equation (5).

$$\phi_{\mathrm{VL}}^{\mathrm{EB}} = \phi_{\mathrm{EB}} \pm 8 \tag{5}$$

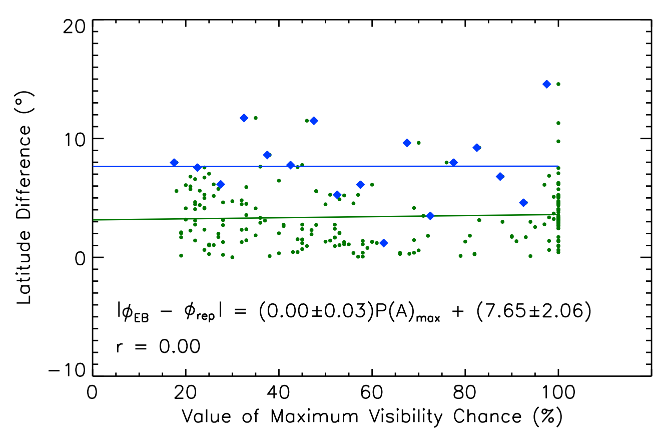

Figure 7. Of the same form as Figure 4. The latitude difference between the positive reports and the equatorial boundary ($|\phi_{\mathrm{EB}}| - |\phi_{\mathrm{rep}}|$) is plotted as a function of the maximum visibility ($P(A)_{\mathrm{max}}$).

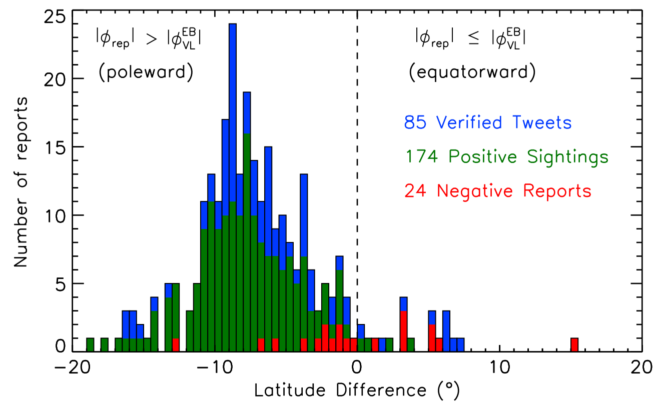

Figure 8. Of the same form as Figure 3. A histogram of the differences in latitude between the Aurorasaurus reports (ϕ_{rep}) and the equatorial boundary-based view line (ϕ_{VL}^{EB}).

In Figure 8, the difference between the reports and this new equatorial boundary-based view line (ϕ_{VL}^{EB}) are shown in a similar form to Figure 3. The median difference of the positive reports is 7.74° (850 km poleward) and 95.0% of the positive reports are poleward of the view line. The accuracy of this view line is 91.2% (ΣTP = 246, ΣTN = 12, and ΣR = 283).

The results of comparing the Aurorasaurus citizen science reports with each of the view lines previously discussed (i.e., the updated SWPC view line, the Aurorasaurus view line, and the equatorial boundary-based view line) are summarized in Table 1.

5. Discussion

An aurora "view line" estimates the most equatorial latitude from which an observer might see an aurora based upon the current (or predicted) auroral oval size and strength. Since an aurora can reach approximately 400 km in altitude [*Kataoka et al.*, 2013], a simple estimate places this view line in the region of 9–10° in latitude from the auroral oval.

Of course, such a basic approach neglects several factors including the width of the auroral oval (which can span several degrees in latitude), the total aurora precipitation flux (which can affect its luminosity), and the type of aurora. In the early evening, for example, the aurora typically consists of quiet arcs that do not extend far in latitude. Therefore, the viewing range will likely be reduced in those early evening sector longitudes. Similarly, the patchy or pulsating aurora in the dawn sector may not have the large vertical extent required to be observed from large distances away. It is usually the premidnight/midnight sector that has bright aurora, with large vertical rays, and the spread in altitude required to observe aurora from large distances.

As shown in Figure 2 (middle), the majority of the Aurorasaurus reports take place in the premidnight local time sector. Thus, perhaps unsurprisingly, the majority of positive reports occur most often when the aurora is likely to be at its brightest and visible from large distances away. Of course, the premidnight hours are also the most sociable for citizen scientist observers to be out "aurora hunting." Future work should attempt to account for local time when determining the location of the view line, but such work is beyond the scope of this initial case study.

As shown in equation (2), to try to account for some of these complicating factors, SWPC provides an equatorward view line estimate that is based upon both the maximum intensity of the aurora (or, rather, the maximum

Table 1. A Summary of the View Lines and Their Accuracies

View Line	Equation	Accuracy (%)
Updated SWPC	$\phi_{VL}^{SWPC} = \phi_{P(A)\max} \pm \left(\frac{P(A)_{\max}}{20} + 3\right)$	43.9
Aurorasaurus	$\phi_{VL}^{AS} = \phi_{P(A)\max} \pm \left(\frac{P(A)_{\max}}{16} + 8\right)$	90.1
Equatorial boundary	$\phi_{VL}^{EB} = \phi_{EB} \pm 8$	91.2

probability of visible aurora) and the location at which this occurs, as determined by OP10. Though in this study, SWPC's estimate was updated by using OP13 as the auroral precipitation data source.

It should be expected that if this estimate is performing well, almost all of the Aurorasaurus positive reports (i.e., positive sightings and verified tweets) would be poleward of this view line. Of course, we might expect that some positive reports would be equatorward, though, owing to factors not accounted for in OP13, such as a sudden brightening of the aurora (e.g., the result of a substorm), the observer's altitude, or their camera sensitivity.

Plotted in Figure 3 are the differences between the Aurorasaurus reports and the updated SWPC view line. The distribution of the histogram showed that 62% of the positive reports were equatorward of the view line with a median difference of +3.70° or approximately 400 km equatorward. These results suggest that the updated SWPC view line is somewhat conservative. Additionally, the accuracy of the view line (as determined using equation (3)) was poor at 43.9%.

It is important to note, however, that the large majority (85%) of the Aurorasaurus reports are positive (i.e., positive sightings or verified tweets). This is perhaps to be expected since citizen scientists are more likely to be motivated to report their observations when the outcome is favorable (i.e., they saw an aurora) [cf. *Sequeira et al.*, 2014]. Unfortunately, however, this positive reporting skew may be considered as a form of sampling bias that may affect the determined view line accuracy since we are, in particular, lacking in "true negatives" (i.e., as predicted, an observer equatorward of the view line was unable to see an aurora) and "false positives" (i.e., an observer poleward of the view line, with clear skies and an unobstructed view, was unable to see an aurora).

Additionally, reports made by citizen scientists are generally made near areas of fairly high population. As such, and as shown in Figure 2 (top), the majority of the Aurorasaurus reports are most likely from the equatorial edge of a visible aurora. While this is actually quite useful when determining the equatorial extent of a visible aurora, it does mean that there is a bias toward lower latitude values when determining the overall accuracy of a view line (i.e., there are less "true positives" at latitudes greater than the view line latitude — even though an aurora would, indeed, be visible from there).

To further compare the view line with the positive reports, the difference in latitude between them was plotted against the maximum probability of visible aurora in Figure 4. We note that the large grouping at $P(A)_{max} =$ 100%, which contained around 30% of the positive reports, is simply an artifact of the conversion of energy flux into a percentage (which, clearly, has an upper bound). Since we were specifically interested in comparing the observations with the updated SWPC view line, the visibility percentage was chosen in this case study. However, further work could investigate the relationship between the OP13 modeled energy flux, rather than the SWPC percentage values, and the latitude from which an aurora was visible.

Additionally, when determining the distance between the reports and the view line, it is simply the distance along the same longitude that is calculated. In reality, though, an aurora may be visible from a location with a different longitude (i.e., an aurora may be visible to the north-east rather than due north). The calculated distance between the reports and the view line may, therefore, not accurately represent the actual distance between the reporter and the closest location of the view line.

A line of best fit through these data provides observationally derived coefficients that can replace those in equation (2). Since a view line is the most equatorward location from which the aurora might be seen, the line of best fit was plotted through the maximum differences. The original view line coefficients were then replaced with these fitting parameters, to produce an observationally based Aurorasaurus view line, in equation (4). We note that there is some considerable spread in the data, which results in large uncertainties in the fitting parameters. Indeed, the correlation coefficient for this fit is moderate poor, $r = 0.48$, which further demonstrates the uncertainty in this relationship. Using additional data to compute this fit may help to reduce such uncertainties.

The reports were then compared with this Aurorasaurus view line (ϕ_{VL}^{AS}) in Figure 5. This comparison demonstrated that 92.7% of the positive reports were poleward of the view line and the median difference was −4.55° (500 km poleward). The accuracy of this modified view line is 90.1% (more than double the updated SWPC view line accuracy).

Both the updated SWPC view line and the Aurorasaurus view line scale the location of the maximum auroral visibility equatorward by some factor, since this is the most likely location of aurora visible directly overhead. However, it is quite possible that the aurora may also be visible overhead at latitudes further equatorward than this (for example, the maximum visibility may be 100% at one latitude but still 90% several degrees equatorward of this). As such, the location of the reports was also compared with the location of the auroral equatorial boundary in Figure 6. The location of the equatorial boundary was defined as the most equatorward latitude at which $P(A) \geq 18\%$.

The median difference for the positive reports was 0.62° (approximately 70 km equatorward). When filtering to only those reports equatorward of the equatorial boundary, the median difference was +3.06° (approximately 350 km equatorward). The accuracy of using just the equatorial boundary as a view line was found to be 49.7%.

Of course, it is expected that there should be some scaling involved, just as there is with the updated SWPC and Aurorasaurus view lines. To determine the appropriate scaling factors, the latitude difference between the positive reports and the equatorial boundary was plotted against the maximum probability of visible aurora in Figure 7. The coefficients of the fit then provided the scaling needed to compute an equatorial boundary-based view line (ϕ_{VL}^{EB})—as shown in equation (5).

Interestingly, there is found to be no dependence on $P(A)_{max}$, rather ϕ_{VL}^{EB} is just the location of the equatorial boundary (ϕ_{EB}) scaled equatorward by 8°. This suggests that location of the equatorial boundary may itself scale based upon $P(A)_{max}$ (i.e., ϕ_{EB} moves equatorward as $P(A)_{max}$ increases) which is, perhaps, unsurprising.

The reports were then compared to the equatorial boundary-based view line. This comparison shows that 95.0% of the positive reports were poleward of the view line and that the median difference was −7.74° (850 km poleward). The accuracy was found to be slightly higher than the Aurorasaurus view line (91.2%) and considerably higher than the updated SWPC view line (43.9%).

During the aforementioned comparisons, the number of reports used varies from the total number shown in Figure 2. This is due to some reports being recorded at longitudes with no associated view line, e.g., the observation was at dusk/dawn and the view line had been trimmed from that longitude due to the presence of the day/night terminator, or in the case of the equatorial boundary-based view line, $P(A)_{max}$ was below the threshold of 18%. While this is unfortunate, the number of reports remaining for each comparison is still significant.

6. Conclusion

More than 60 years ago, *Gartlein and Moore* [1951] showed that dedicated amateur observers could make critical contributions to some of the earliest auroral models, and now, as shown in this study, observers armed with modern mobile crowdsourcing technologies are making demonstrable improvements to the latest models too.

Specifically, in this study, citizen science reports of auroral visibility (or lack thereof), provided by the Aurorasaurus project, were compared to an updated version of NOAA's SWPC OVATION Prime-based view line. The reports consist of positive sightings and negative reports submitted directly to Aurorasaurus, along with verified tweets, which are positive sightings reported on Twitter and verified by Aurorasaurus users. The reports were collected during March and April 2015 and covered a range of latitudes, local time, and geomagnetic activity.

The reports demonstrated that during these two months, the updated SWPC view line underestimated the distance from which an aurora could be observed. Over 60% of the positive reports (which includes positive sightings and verified tweets) occurred at latitudes equatorward of the view line. The accuracy (equation (3)) was found to be poor at just under 44% (though, as previously discussed, there are several caveats to this accuracy value). These results suggested that further investigation into the scaling parameters used in the SWPC view line calculation (equation (2)) was warranted.

New scaling parameters for the view line equation were determined from the relationship of the differences in latitude between the positive reports and the updated SWPC view line and the maximum probability of visible aurora. This modified Aurorasaurus view line takes a very similar form to the updated SWPC view line but scales the line further equatorward. With these modified parameters, nearly 93% of the positive reports occurred at latitudes poleward of the view line. The accuracy drastically improved and, in fact, more than doubled to 91%.

The updated SWPC view line and the modified Aurorasaurus version of it scale the latitude of the maximum probability of visible aurora. As discussed, this may not always be the most appropriate location to scale. Therefore, a view line based upon the location of the auroral equatorial boundary was also constructed. This view line also performed well, and slightly better than the Aurorasaurus view line, with 97% of the positive reports occurring poleward of the view line and the overall accuracy of 95%. Though the restriction placed on when the view line should be drawn (i.e., $P(A)_{max}$) resulted in comparisons with fewer reports.

Of course, it should be expected that view lines created using a set of observations should perform well when then compared to those observations. It is therefore sensible to test these view lines (equations (2), (4), and (5)) further using other data and adapt them as necessary. For example, the Aurorasaurus data set has aurora observations spanning from November 2014 to present. Further work to incorporate those observations seems a worthwhile endeavor. We note, however, that the equatorial-based view line has been running on the Aurorasaurus website since November 2015 and, anecdotally, has matched well with the reports from citizen scientists that have been submitted since then (including those submitted during a few large auroral events).

We note that in this case study, there was a limited number of negative reports which resulted in the accuracy of the view lines being predominantly determined by how well they were able to predict true positives. Additionally, reports were generally provided by observers located on the equatorward edge of a visible aurora. As a result, the view lines created in this study may sometimes overestimate the distance from which the aurora can be seen and so should be treated as the most optimistic values. Ideally, the number of useful negative reports (whereby a user has a clear sky but is unable to view the aurora) should roughly equal the number of positive reports and the reports should span right across the auroral oval. Future studies should aim to address these issues.

Additionally, it is important to note that the view lines used in this case study are based upon the output of the OP13 model. Although this model has shown to be accurate at modeling the extent of the auroral oval, at least statistically, there may be times when it does not perform quite so well (e.g., during substorms [*Newell et al.*, 2010a; *Machol et al.*, 2012]) or the real-time aurora is not quite as expansive as suggested. It is also unable to make any estimates as to the height or color of the aurora which are both factors that may significantly affect where an aurora can be seen from [*Machol et al.*, 2012]. Improving the modeling of the aurora, and the conversion of energy flux to SWPC's "probability of visible aurora" (by taking local time into account, for example), will therefore improve the accuracy of any view line. Using the Aurorasaurus reports to help account for such factors may improve the accuracy of OVATION Prime and, subsequently, any view lines based upon it.

Lastly, in this case study, the three view lines discussed were all estimates of the equatorial extent of auroral visibility. In principle, the poleward extent of visible aurora could also be determined. The poleward view line could use either the same scaling parameters discussed here or, more ideally, new scaling parameters could be determined. In practice, such determination might prove more difficult due to a lack of citizen science reports at extremely high latitudes. Further work might attempt to mitigate this issue, perhaps using automated camera observations, to provide a view line for even the most poleward of observers.

Acknowledgments
This material is based upon work supported, in part, by the National Science Foundation (NSF) under grant 1344296. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of NSF. The OVATION Prime output and associated view line were kindly supplied by the Space Weather Prediction Center, Boulder, CO, National Oceanic and Atmospheric Administration (NOAA), U.S. Department of Commerce. The output can be freely downloaded from the NOAA SWPC product pages (http://www.swpc.noaa.gov/products/aurora-30-minute-forecast).

References

Bartels, J., N. H. Heck, and H. F. Johnston (1939), The three-hour-range index measuring geomagnetic activity, *Terr. Magn. Atmos. Electr.*, *44*(4), 411–454, doi:10.1029/TE044i004p00411.

Baumgardner, J., J. Wroten, J. Semeter, J. Kozyra, M. Buonsanto, P. Erickson, and M. Mendillo (2008), A very bright SAR arc: Implications for extreme magnetosphere-ionosphere coupling, *Ann. Geophys.*, *25*(12), 2593–2608, doi:10.5194/angeo-25-2593-2007.

Case, N. A., E. A. MacDonald, M. Heavner, A. H. Tapia, and N. Lalone (2015a), Mapping auroral activity with Twitter, *Geophys. Res. Lett.*, *42*, 3668–3676, doi:10.1002/2015GL063709.

Case, N. A., E. A. MacDonald, and K. G. Patel (2015b), Aurorasaurus and the St Patrick's Day storm, *Astron. Geophys.*, *56*(3), 3.13–3.14, doi:10.1093/astrogeo/atv089.

Donovan, E., E. Spanswick, J. Liang, J. Grant, B. Jackel, and M. Greffen (2012), Magnetospheric dynamics and the proton aurora, in *Auroral Phenomenology and Magnetospheric Processes: Earth And Other Planets*, edited by A. Keiling et al., AGU, Washington, D. C., doi:10.1029/2012GM001241.

Hardy, D. A., M. S. Gussenhoven, and E. Holeman (1985), A statistical model of auroral electron precipitation, *J. Geophys. Res.*, *90*(A5), 4229–4248, doi:10.1029/JA090iA05p04229.

Hardy, D. A., M. S. Gussenhoven, and D. Brautigam (1989), A statistical model of auroral ion precipitation, *J. Geophys. Res.*, *94*(A1), 370–392, doi:10.1029/JA094iA01p00370.

Kataoka, R., et al. (2013), Stereoscopic determination of all-sky altitude map of aurora using two ground-based Nikon DSLR cameras, *Ann. Geophys.*, *31*(9), 1543–1548, doi:10.5194/angeo-31-1543-2013.

Gartlein, C. W., and R. K. Moore (1951), Southern extent of Aurora Borealis in North America, *J. Geophys. Res.*, *56*(1), 85–96, doi:10.1029/JZ056i001p00085.

Lane, C., A. Acebal, and Y. Zheng (2015), Assessing predictive ability of three auroral precipitation models using DMSP energy flux, *Space Weather*, *13*, 61–71, doi:10.1002/2014SW001085.

MacDonald, E. A., N. A. Case, J. H. Clayton, M. K. Hall, M. Heavner, N. Lalone, K. G. Patel, and A. Tapia (2015), Aurorasaurus: A citizen science platform for viewing and reporting the aurora, *Space Weather*, *13*, 548–559, doi:10.1002/2015SW001214.

Machol, J. L., J. C. Green, R. J. Redmon, R. A. Viereck, and P. T. Newell (2012), Evaluation of OVATION Prime as a forecast model for visible aurorae, *Space Weather*, *10*, S03005, doi:10.1029/2011SW000746.

Mitchell, E. J., P. T. Newell, J. W. Gjerloev, and K. Liou (2013), OVATION-SM: A model of auroral precipitation based on SuperMAG generalized auroral electrojet and substorm onset times, *J. Geophys. Res. Space Physics*, *118*, 3747–3759, doi:10.1002/jgra.50343.

Newell, P. T., T. Sotirelis, K. Liou, C.-I. Meng, and F. J. Rich (2007), A nearly universal solar wind-magnetosphere coupling function inferred from 10 magnetospheric state variables, *J. Geophys. Res.*, *112*, A01206, doi:10.1029/2006JA012015.

Newell, P. T., T. Sotirelis, K. Liou, A. R. Lee, S. Wing, J. Green, and R. Redmon (2010a), Predictive ability of four auroral precipitation models as evaluated using Polar UVI global images, *Space Weather*, *8*, S12004, doi:10.1029/2010SW000604.

Newell, P. T., T. Sotirelis, and S. Wing (2010b), Seasonal variations in diffuse, monoenergetic, and broadband aurora, *J. Geophys. Res.*, *115*, A03216, doi:10.1029/2009JA014805.

Newell, P. T., K. Liou, Y. Zhang, T. Sotirelis, L. J. Paxton, and E. J. Mitchell (2014), OVATION Prime-2013: Extension of auroral precipitation model to higher disturbance levels, *Space Weather*, *12*, 368–379, doi:10.1002/2014SW001056.

Poppe, B. B. (2000), New scales help public, technicians understand space weather, *Eos Trans. AGU*, *81*(29), 322–328, doi:10.1029/00EO00247.

Roble, R. G., and E. C. Ridley (1987), An auroral model for the NCAR Thermospheric General Circulation Model (TGCM), *Ann. Geophys.*, *5*, 369–382.

Sequeira, A. M., P. E. Roetman, C. B. Daniels, A. K. Baker, and C. J. Bradshaw (2014), Distribution models for koalas in South Australia using citizen science collected data, *Ecol. Evol.*, *4*(11), 2103–2114, doi:10.1002/ece3.1094.

Spiro, R. W., P. H. Reiff, and L. J. Maher Jr. (1982), Precipitating electron energy flux and auroral zone conductances—An empirical model, *J. Geophys. Res.*, *87*(A10), 8215–8227, doi:10.1029/JA087iA10p08215.

Zhang, Y., and L. J. Paxton (2008), An empirical Kp-dependent global auroral model based on TIMED/GUVI FUV data, *J. Atmos. Sol. Terr. Phys.*, *70*(8–9), 1231–1242, doi:10.1016/j.jastp.2008.03.008.