Final Report on:

Measurement of Precipitation Induced FUV Emission and Geocoronal Lyman Alpha from the IMI Mission

NAS8-40213

Submitted to NASA Marshall Space Flight Center

Submitted by:
Space Sciences Laboratory
Lockheed Martin Palo Alto Research Laboratory
Palo Alto, California 94304

Principal Investigator: Dr. S. B. Mende
Co-Investigators: Dr. S. A. Fuselier, Dr. R. L. Rairden
Department 91-20, Building 252
3251 Hanover Street
Palo Alto, CA 94304

TITLE OF INVESTIGATION: **Measurement of precipitation induced FUV emission and geocoronal Lyman alpha from the IMI mission**

PRINCIPAL INVESTIGATOR: Dr. Stephen B. Mende
INSTITUTION: Lockheed Martin Palo Alto Research Laboratory

Abstract: This final report describes the activities of the Lockheed Martin Palo Alto Research Laboratory in studying the measurement of ion and electron precipitation induced Far Ultra-Violet (FUV) emissions and Geocoronal Lyman Alpha for the NASA Inner Magnetospheric Imager (IMI) mission. This study examined promising techniques that may allow combining several FUV instruments that would separately measure proton aurora, electron aurora, and geocoronal Lyman alpha into a single instrument operated on a spinning spacecraft. The study consisted of two parts. First, the geocoronal Lyman alpha, proton aurora, and electron aurora emissions were modeled to determine instrument requirements. Second, several promising techniques were investigated to determine if they were suitable for use in an IMI-type mission. Among the techniques investigated were the Hydrogen gas cell for eliminating cold geocoronal Lyman alpha emissions, and a coded aperture spectrometer with sufficient resolution to separate Doppler shifted Lyman alpha components.

Table of Contents

1. Introduction

One of the most promising techniques for remote sensing, imaging, measurement of the magnetospheric particle population is measuring the directionality of the arriving fluxes, or imaging of the charged exchanged neutral particles. The flux of neutral particles are produced by the interaction of a magnetospheric fast ions with the ambient neutral "population". The goal of such imaging is to describe the distribution of the fast parent ions in magnetospheric regions. In order to derive the parent ion fluxes it is necessary to also measure the neutral densities which produce the charged exchanged neutrals. The measurement of the ambient neutral population can be accomplished in a remote sensing manner by measuring the solar Lyman alpha scattering by the neutrals. Knowing solar Lyman alpha fluxes would allow the computation of the density of the neutral hydrogen available for charge exchange.

Previously flown satellite imaging experiments have demonstrated the suitability of the vacuum ultraviolet region for remote sensing observations of auroral particle precipitation. In the wavelength region 120 - 145 nm, a downward viewing imager is uncontaminated by the earth albedo and the intensity of the auroral emissions in most cases is competitive with the re-scattered light even during daylit conditions. These features permit the quantitative imaging of the auroral regions during day and night conditions. An instrument suitable for such observation should also have adequate wavelength resolution to separate key spectral features and simultaneously observe the Doppler profile of the auroral Lyman alpha line. The auroral Lyman alpha line provides a measure of the auroral ion precipitation which is a highly valued foot print generated by precipitating auroral protons.

In the development of the Inner Magnetospheric Imager (IMI) program it has become clear that any near term spacecraft program will have to utilize relatively modest size instruments. In order to economize in cost and instrument resource needs it was suggested that a single multi purpose instrument capable of (a) measurement of the scattered geocoronal Lyman alpha, (b) imaging the aurora day and night in the UV and (c) the observation of proton precipitation would be highly desirable. The purpose of this report to investigate the approaches leading towards such multi purpose a single instrument with such multiple capabilities.

2. Auroral Emissions in the Far Ultra-Violet

The FUV wavelength range (120 to 150 nm) is demonstrated to be one of the most effective wavelength ranges for making remote sensing optical observations of auroral precipitation [Chakrabarti et al., 1983; Paresce et al., 1983]. The advantages of this wavelength range over the more familiar visible wavelength range are:

1) The aurora can be brighter than the daylit atmospheric background, permitting monitoring of the auroral oval during both day and night [Frank and Craven, 1988; Anger et al., 1987].

2) The atmosphere is not transparent in this wavelength region permitting interpretation of the signal without requiring correction for albedo variations of the underlying Earth.

Two additional properties of the FUV wavelength range define three major measurement objectives of the FUV imager.

1) Because the atmosphere is not transparent (due to O_2 photoabsorption), the ratio of emission lines can be used to obtain the penetration depth of electrons into the atmosphere and therefore the initial energy of the parent electron distributions that create the electron aurora [Strickland et al., 1983].

2) This wavelength region contains the Lyman alpha emission line (121.6 nm). The Doppler shift of this line can be used to determine the incident energy of the parent proton distributions that create proton aurora [Edgar et al., 1973; Ishimoto et al., 1989; Strickland et al., 1993].

3) The Lyman alpha emission line at 121.6 nm is produced by scattering of solar Lyman alpha by the Earth's geocorona. Measurement of its intensity with distance yields the geocorona density [e.g., Rairden et al., 1986]. This density measurement is important for determining the magnetospheric energetic ion density from energetic neutral measurements.

2.1 Proton Precipitation and the Proton Aurora

Lyman Alpha is produced by energetic magnetospheric protons that charge exchange in the upper atmosphere (i.e., proton aurora [Paresce et al., 1983]) and by excitation of geocoronal Hydrogen surrounding the Earth [Chamberlain, 1963]. Thus, we have in the charge exchange case:

$$H^+(\text{energetic}) + H\ (O, N, O_2, N_2, \text{etc}) \rightarrow H^*\ (\text{ENA}) + H^+$$

Followed by a decay to the ground state by emitting a photon. In the case of the geocoronal H, we have the decay to the ground state of an initial excited H atom (excited by solar radiation)

$$h\nu + H \rightarrow H^* \rightarrow H + h\nu$$

In both cases, the wavelength of the emission can be calculated by the simple (Rydberg) formula:

$$\kappa = \frac{1}{\lambda} = R_H\left(\frac{1}{1^2} - \frac{1}{n^2}\right) \quad n = 2, 3, 4,... \qquad (1)$$

Where R_H is the Rydberg Constant (R_H = 10967757.6 m^{-1}). From Equation 1 it is easy to calculate that the 2 to 1 electron transition in the H atom will produce emission at 121.6 nm (Ly α) and 3 to 1 electron transition will produce emission at 102.6 nm (Ly β).

Note however that the emissions produced by the initially energetic magnetospheric protons will be Doppler shifted by several tenths of nm because the energetic neutral atoms are moving downward into the atmosphere at a considerable speed. Thus, the Lyman alpha line observed on a spacecraft with an apogee well above the Earth will be the combination of geocoronal emissions at 121.6 nm and a (red) shifted component at a few tenths of nm higher [Ishimoto et al., 1989]. Separation of these two components would allow the determination of the energy of the incident magnetospheric protons and the cold geocoronal H density.

Van Zyl et al. [1984] discuss the steps for determining the Lyman alpha emission from the incident proton distribution. Although we would like to reverse the process, it is instructive to follow their steps to see how the Lyman alpha emission is produced. There are three basic steps to determining the Lyman alpha emission intensity and line profile.

1) determine the number of Lyman alpha photons produced per incident proton and the change in energy of the proton as it collides with the upper atmosphere.

2) determine the number of incident protons and their energy.

We first consider step 1.

A precipitating proton in the upper atmosphere will suffer collisions with the gases. Some collisions will yield neutral H in excited states and others will maintain the proton charge.

Collisions of the type shown below are most important [Basu et al., 1987].

$$H^+ + M \rightarrow H + M^+$$
$$H + M \rightarrow H^+ + M + e^-$$
$$H + M \rightarrow H + M^+ + e^-$$

where M is any constituent of the upper atmosphere.

In these collisions, the H products that are created in an excited state are the important reactions for the Lyman alpha production while the excited atmospheric constituents and the secondary electrons created in the latter two reactions are important for determining the incident proton contributions to other molecular and atomic emissions in the aurora.

An equilibrium will be established so that a flux of H^+ and H will propagate through the atmosphere. The fraction of ionized and neutral H in equilibrium will depend on the initial energy of the precipitating protons. This charge state equilibrated flux as a function of proton energy is shown by Van Zyl et al. [1984]. Basically, most of the initial protons will be charge exchanged as long as the energy is below about 10 keV. However, by 100 keV, there will be very few charge exchanged protons.

Next, we must investigate the energy loss of the primary precipitating proton in the production of excited H (or H*). Again, this has been investigated by Van Zyl et al. [1984] for a pure N_2 atmosphere. In addition, the Lyman Alpha yield in Rayleighs/(erg cm^2 s) for H atom emissions was modeled by Strickland et al. [1993]. The ratio of the Lyman alpha photon production relative to the H alpha photon production is 4.00.

The following table, from Van Zyl et al. [1984] gives the H_α photon yields as a function of the primary H^+ incident energy on the atmosphere. In the right most column, the Lyman alpha yield in photons per proton is shown by multiplying the H_α yield by 4.00 after Strickland et al. [1993].

Energy (keV)	H_α photon yield	L_α photon yield = 4 x H_α yield
1	4.6	18.4
3	9.4	37.6
10	15.0	60.0
30	22.9	91.6
100	35.8	143.2
300	43.1	172.4

Table 1

The above table gives the total photon yield for an incident proton of a given energy. In reality, a proton with incident energy of e.g. 10 keV will decay in energy and the photons created by it will be Doppler shifted at a range of wavelengths corresponding to proton energies from 10 keV to zero. Thus, to determine the Lyman-Alpha production, we need to know the number of photons produced as an incident proton changes energy from its initial energy to its final energy of ~0 keV. To do this, we fit the data for Lyman-alpha photon production in the above table with a 3rd order polynomial and then determine the number of photons produced as the proton energy decreases by simply taking the difference in the total number of photons produced at two different energies. Table 2 shows the values of the total photons per proton and the differential values in transitioning from one energy step to the next. For example, in the table, there are 6.71 photons per proton produced by protons that transition from 2 keV to 1 keV in energy.

Energy (keV)	Photons/ proton	Delta photons/ proton	Energy (keV)	Photons/ proton	Delta photons/ proton
0.1	1.80	1.80	16	75.84	4.09
0.5	9.40	7.60	18	79.52	3.68
1	20.02	10.62	20	82.87	3.35
2	26.73	6.71	25	90.11	7.24
3	33.29	6.56	30	96.16	6.05
4	38.98	5.69	40	105.88	9.73
5	43.92	4.95	50	113.54	7.66
6	48.28	4.36	60	119.85	6.31
7	52.18	3.90	70	125.20	5.35
8	55.70	3.52	80	129.84	4.64
9	58.91	3.21	90	133.93	4.09
10	61.86	2.95	100	137.58	3.65
12	67.13	5.27	200	161.15	23.57
14	71.74	4.61	300	174.30	13.15

Table 2

Now that we know the number of photons per proton from the above table, we now need to know what wavelength these photons will be Doppler shifted to when they are produced by the

incident protons. This is a relatively easy calculation because the wavelength shift is proportional to the incident energy through the Doppler shift formula:

$$\frac{\Delta\lambda}{\lambda} = \frac{v}{c}$$

Where $\Delta\lambda$ is the wavelength shift, λ is the initial wavelength (121.6 nm), v is the velocity of the proton, and c is the speed of light. Relating the velocity to the energy and using the initial wavelength, the above equation reduces to:

$$\Delta\lambda = 0.1773\sqrt{E}$$
(E in keV, $\Delta\lambda$ in nm)

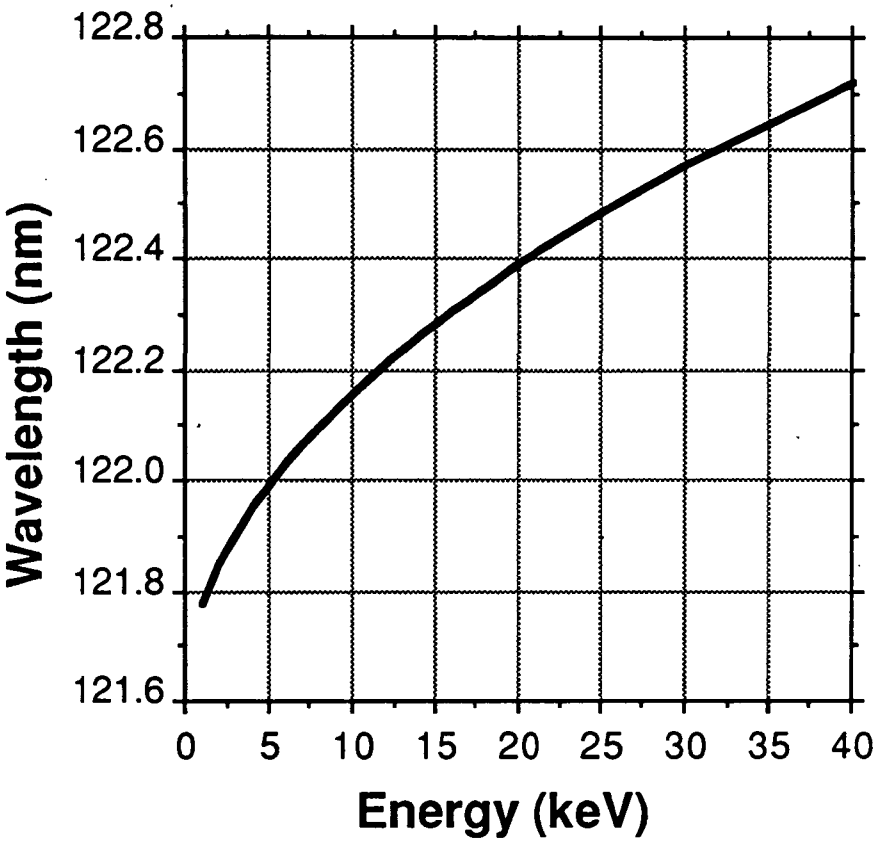


Figure 2.1 Wavelength of proton auroral emissions as a function of the incident proton energy.

Therefore, a 1 keV incident proton will produce approximately 20 photons (from Table 2) that are Doppler shifted by nearly 2 Å from the Lyman-Alpha wavelength (from the above equation).

A plot of the Doppler shift as a function of energy from the last equation (above) is shown in Figure 2.1. The vertical lines separate the Doppler shifts into 2 Å bins. For example, 0-1 keV protons produce Lyman Alpha emissions within 2 Å of the center line, ~1-5 keV protons produce Lyman Alpha emissions from 2-4 Å of the center line, etc..

Now we must determine the incident fluxes of protons into the upper atmosphere and their energies. To do this, we need to know the global distribution of these parameters. Therefore, we consider previous experimental studies of injected particle fluxes.

The azimuthal extent of the injection region is difficult to determine from in situ satellite measurements since orbital constraints result in either poor time resolution with good latitude coverage (low-altitude polar satellites) or poor latitude resolution and good time resolution (synchronous satellites); time resolution in this context refers to the ability to follow substorm development, the characteristic timescale of magnetospheric processes.

Until global imaging of the precipitating particles is achieved, we must rely on statistical studies of the auroral oval obtained from low-altitude polar satellites to determine a statistically averaged auroral ion precipitation map and a corresponding statistically averaged proton aurora intensity. Here, we use the results from Hardy et al. [1989] for a statistical model of auroral ion precipitation. Hardy et al. binned DMSP ion observations as a function of local time, invariant latitude, and Kp index. In each bin, the number flux, energy flux, and average energy of the precipitating distribution were determined. For illustration here, we choose the Kp=4 binning, (that is, moderate magnetospheric activity). Plots of the average energy of the particles and the incident number flux as a function of local time (one half hour bins) and invariant latitude (0.25° latitude bins) are shown in Figures 2.2 and 2.3, respectively. One can see the very high number flux and relatively soft (~few keV) precipitation in the cusp (85° latitude and ~1230 LT). Also evident is the very hard (>20 keV) precipitation in the dusk sector and the fairly intense (~$6x10^6$ ions/(cm^2 s sr)) precipitation at midnight.

To determine the proton aurora intensity using the average energy and number flux in Figure 2.2 and 2.3, we must assume something about the shape of the precipitating distribution. Ishmoto et al. [1989], in their fit to Lyman Alpha observations, assumed that the precipitating proton distribution was a monoenergetic beam. Some observations indicate that the precipitating proton distribution is close to being Maxwellian [e.g., Sharber et al., 1981; Basu et al., 1987].

The procedure we follow here is to assume that the ions are precipitating as a monoenergetic beam with energy equal to the average energy in Figure 2.2 and integral number flux also equal to that in Figure 2.3. Later, we will consider the effects of using a maxwellian precipitating ion flux instead of a monoenergetic beam. With the simplifying assumption of a monoenergetic beam, the calculation of the total, height integrated Lyman alpha flux at a given location is given by:

Total Yield (Rayleighs) = π * integral number flux * #photons/proton(E)

That is, the intensity is simply the number of protons precipitating with energy E times the number of photons/proton produced by a precipitating proton of energy E. The factor of π is present because the flux is integrated over one half of a hemisphere. The integral number flux is in units of 10^6 $cm^{-2}s^{-1}$. Figure 2.4 shows the results of applying this equation to the Kp=4 data from Hardy et al.. For simplicity, we have removed the geocoronal background from the emissions. The peak fluxes of the proton aurora are ~1kR. Figure 2.5 shows how the addition of the ~10 kR geocorna Lyman alpha flux would affect the observations. Note the change in color scale. This figure clearly shows that the geocorona will be as much as 10 times more intense than the proton aurora when observed from a high altitude orbit such as that proposed for IMI.

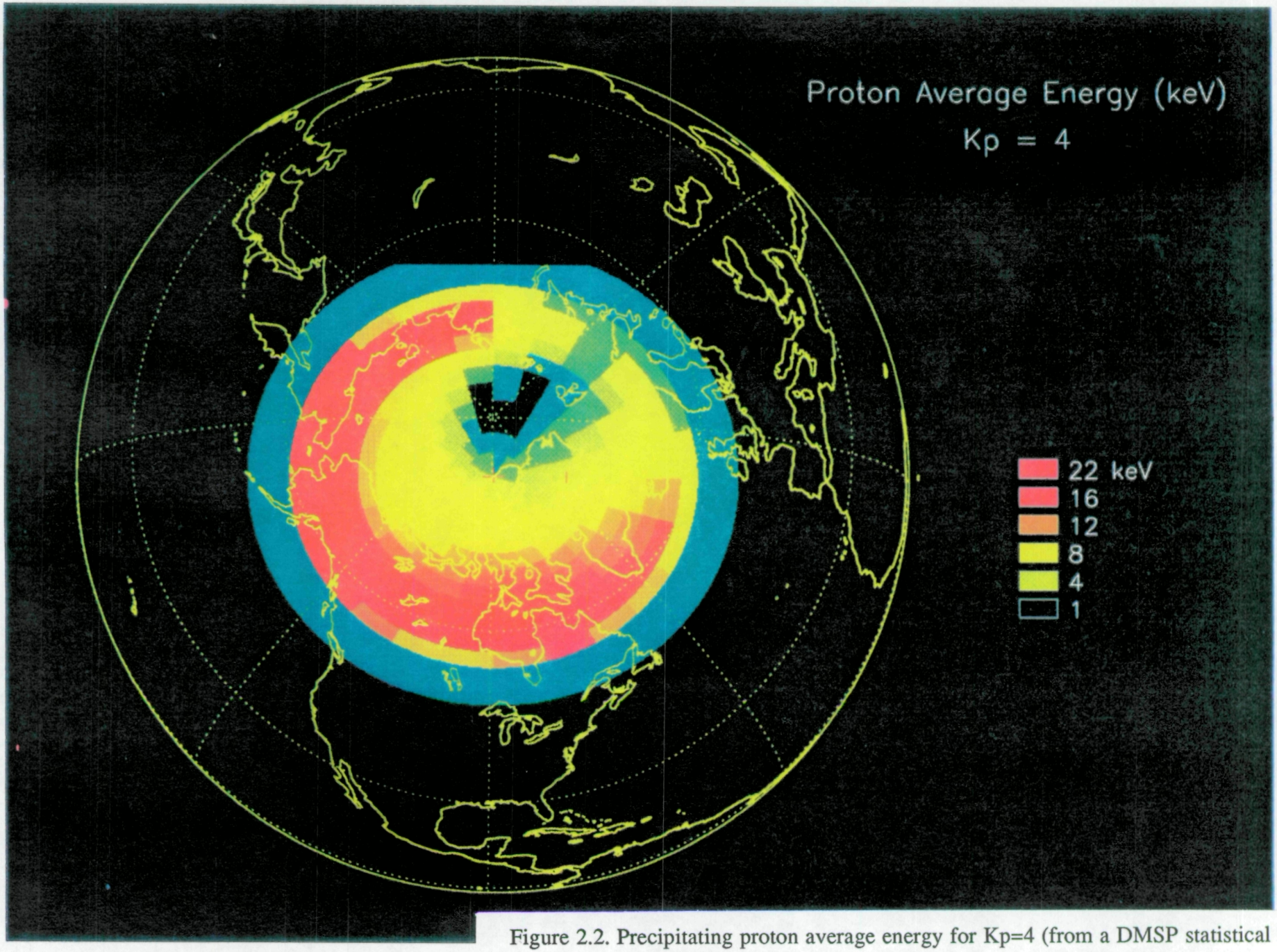


Figure 2.2. Precipitating proton average energy for Kp=4 (from a DMSP statistical study by Hardy et al. [1989]). Local noon is at the top. Relatively soft precipitation occurs in the cusp region while very hard precipitation occurs at midnight and into the dusk sector.

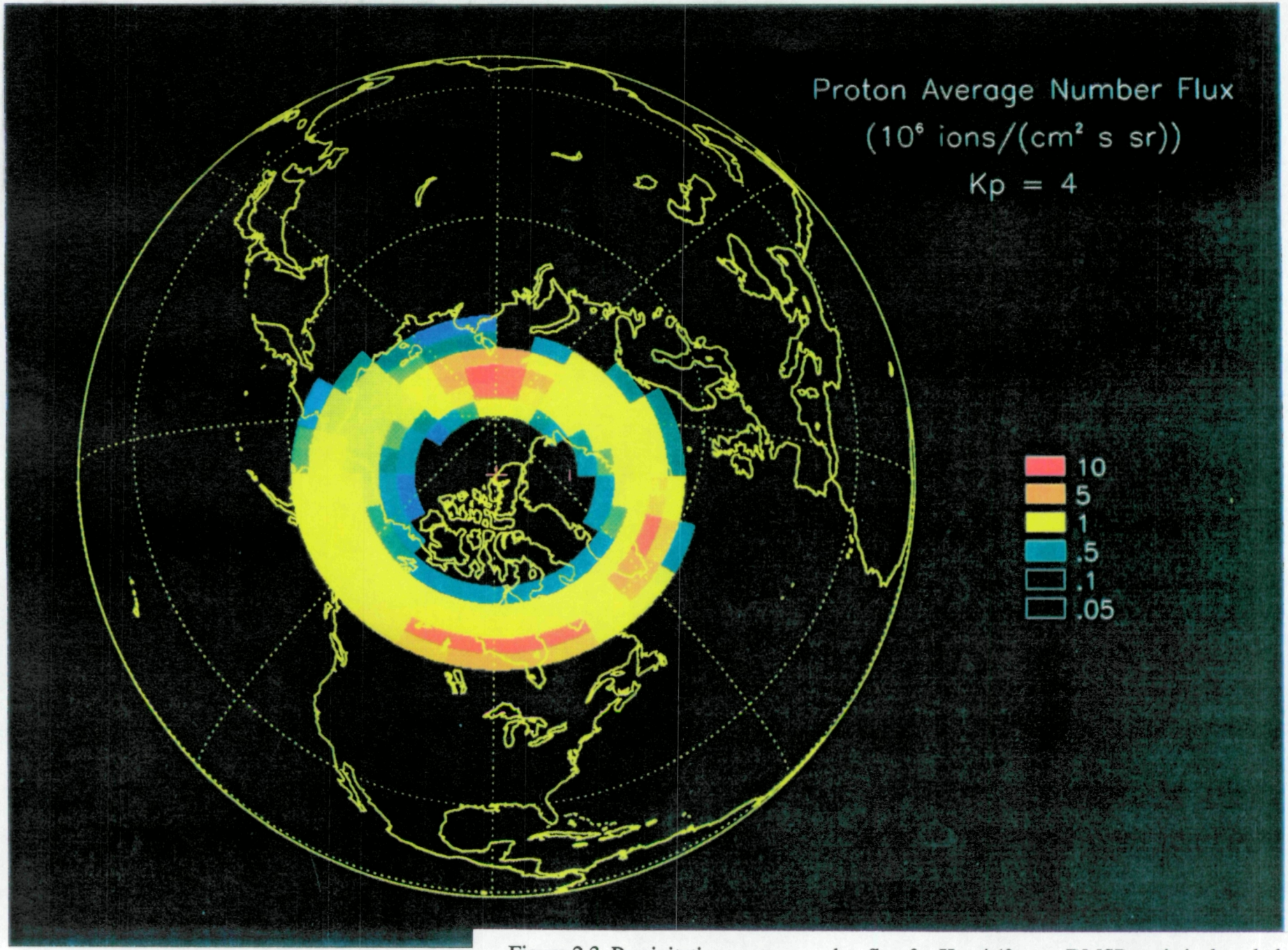


Figure 2.3. Precipitating proton number flux for Kp=4 (from a DMSP statistical study by Hardy et al. [1989]). Intense precipitation occurs in the cusp region and near midnight .

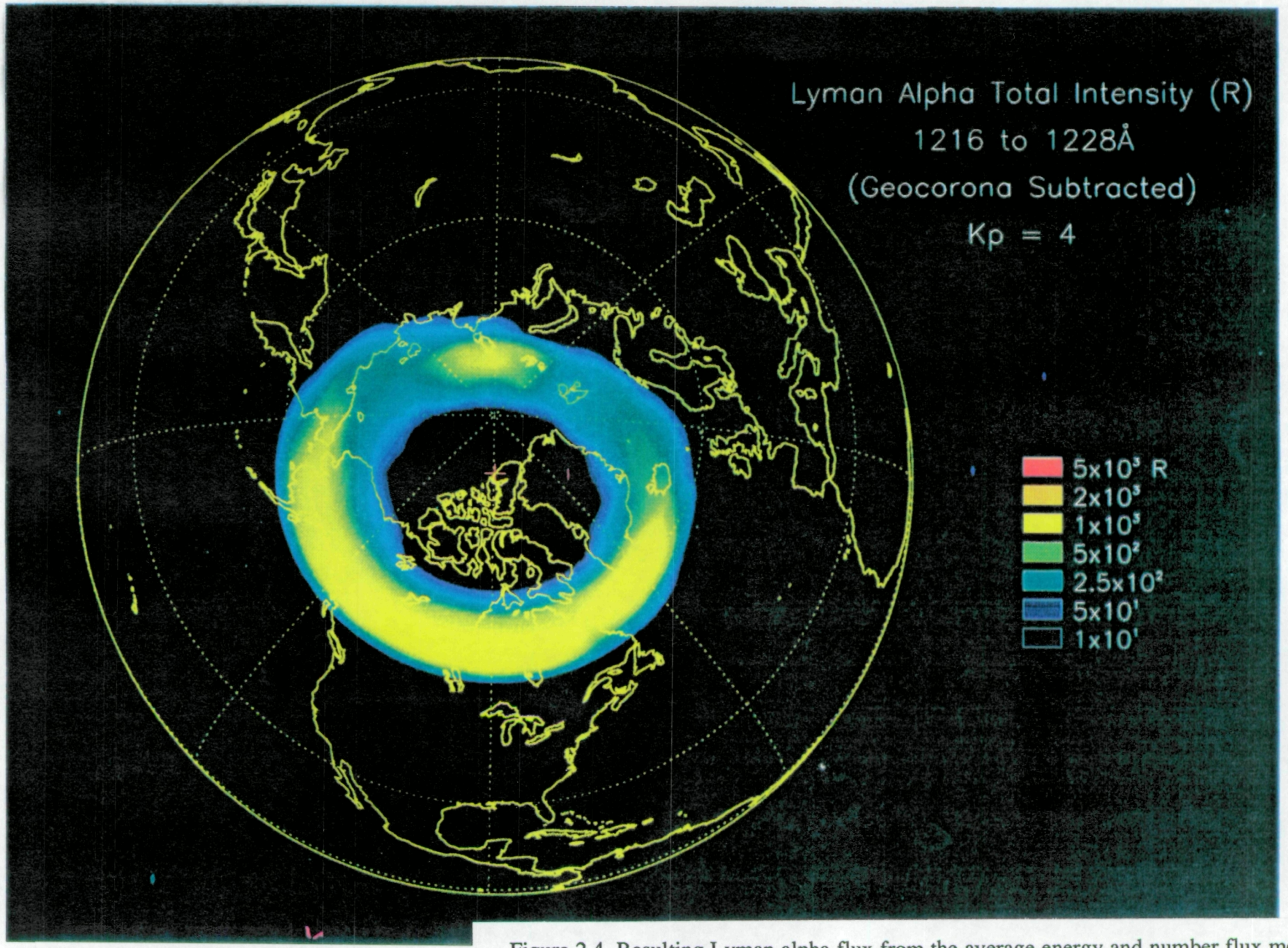


Figure 2.4. Resulting Lyman alpha flux from the average energy and number flux models in Figures 2.2 and 2.3. The geocoronal Lyman alpha has been subtracted to show the proton auroral oval features. A clear cusp region is seen near local noon.

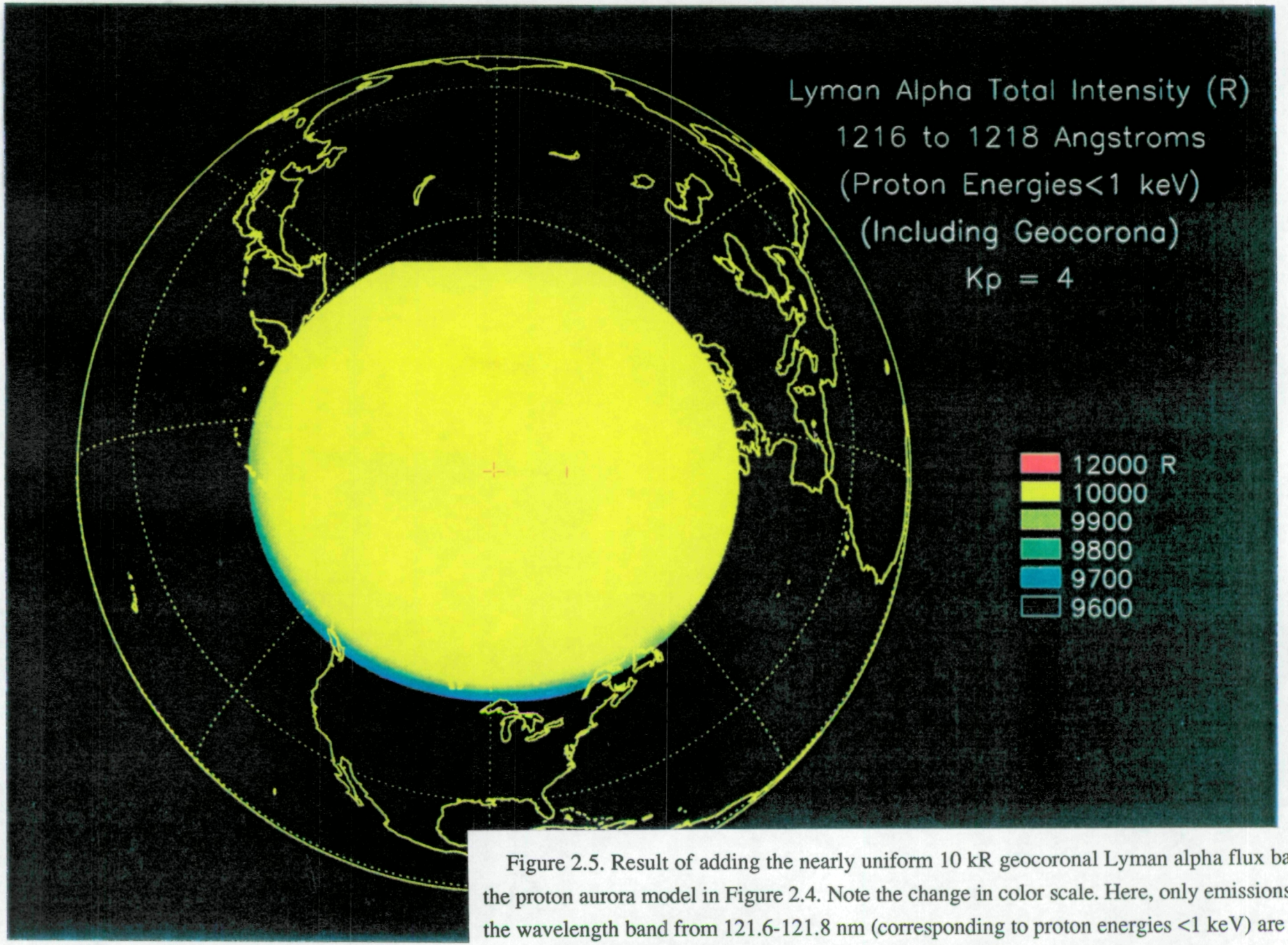


Figure 2.5. Result of adding the nearly uniform 10 kR geocoronal Lyman alpha flux back into the proton aurora model in Figure 2.4. Note the change in color scale. Here, only emissions within the wavelength band from 121.6-121.8 nm (corresponding to proton energies <1 keV) are shown. Because the proton aurora is more than 10 times less intense than the geocorona emissions, it is nearly lost in the intense glow. This is a graphic illustration that either good spectral resolution and very good dynamic range are required for Lyman alpha imaging.

One way of separating the intense geocorna emissions from the less intense proton aurora is to have a spectrometer with sufficient wavelength resolution to separate the geocorona emissions at 121.6 nm from the Doppler shifted proton aurora emissions at higher wavelengths. The resulting separation is shown in Figures 2.6 and 2.7 in 0.2 nm wavelength bins. Above 121.6 nm, there are no geocorna emissions so the subtraction of these emissions in Figures 2.6 and 2.7 was not necessary. Comparing Figures 2.6 and 2.7, one can see the loss of the emissions from the relatively soft precipitation in the cusp region because these emissions are not Doppler shifted to wavelengths above 122.0 nm. The precipitation from the harder precipitation at midnight and in the dusk sector persists out to larger wavelengths.

As Kp increases, the number flux and the energy of the precipitating protons increases. Hardy et al. [1989] generated average energies and number fluxes for Kp ranging from 1 to 6, or from low to very high magnetospheric activity. These fluxes and energies were binned as in Figures 2.2 and 2.3 for each Kp level and the resulting proton auroral fluxes were computed. The proton auroral fluxes with geocorona emissions subtracted for Kp 1 through 6 are shown in Figures 2.8 through 2.13, respectively.

Comparing the proton aurora emissions from low to high magnetospheric activity, it is clear that the emission region grows significantly as the oval expands and that the emissions become somewhat brighter.

Finally, Figures 2.14 through 2.16 depict the effects of a localized injection of 35 keV protons at midnight for an initial Kp=4 auroral oval (Figure 2.14) and the subsequent propagation of this injection front to the dusk side (Figures 2.15 and 2.16). The time sequence from the initial near midnight injection in Figure 2.14 to the injection front propagation to near noon in Figure 2.16 is approximately one half hour. During this time, a new belt of intense precipitation forms equatorward of the original Kp=4 emissions. Intensities for this very strong injection are quite high, over 5 kR initially near midnight at t=0.

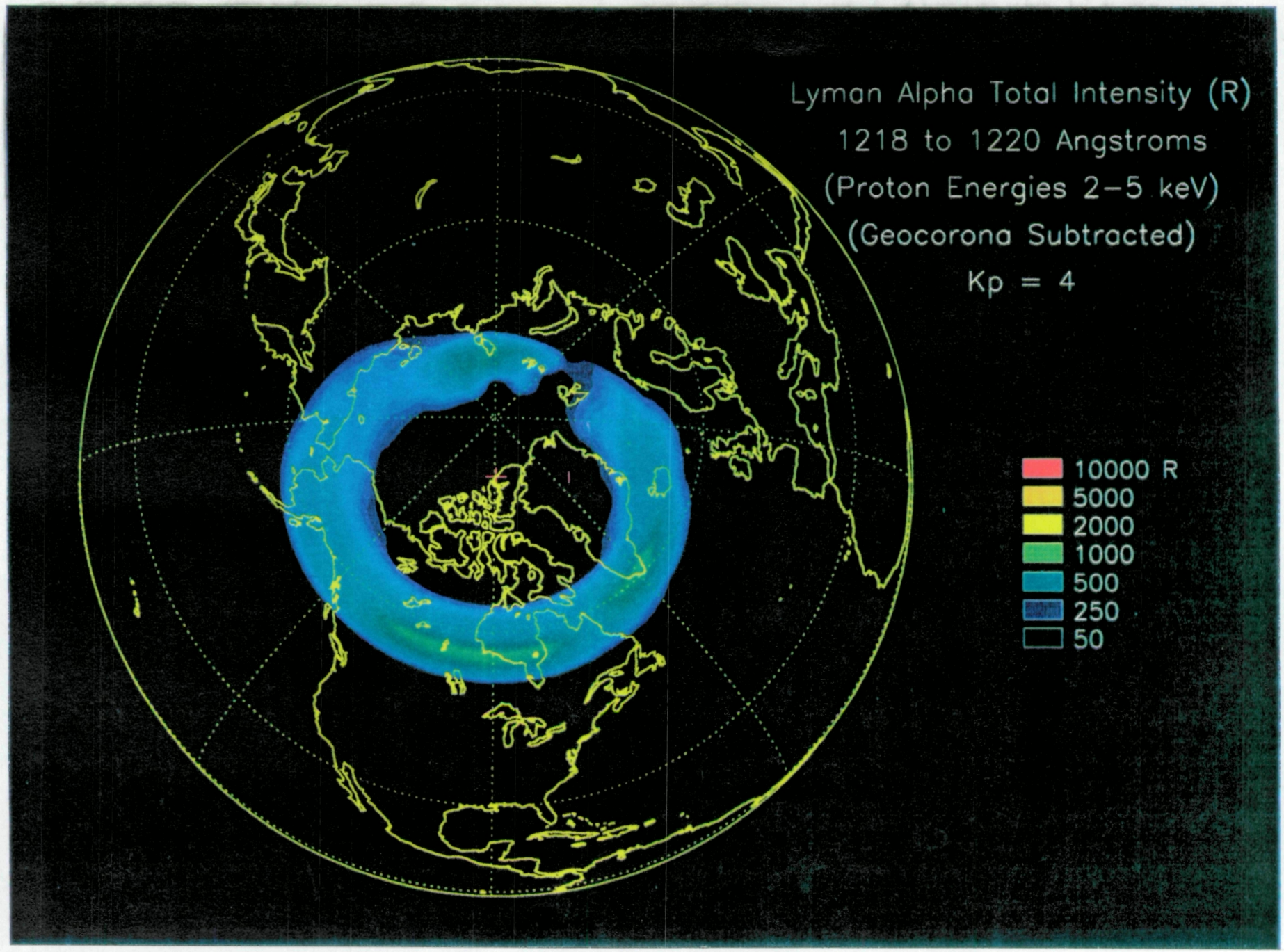


Figure 2.6. Kp=4 proton auroral emissions from 121.8 to 122.0 nm.

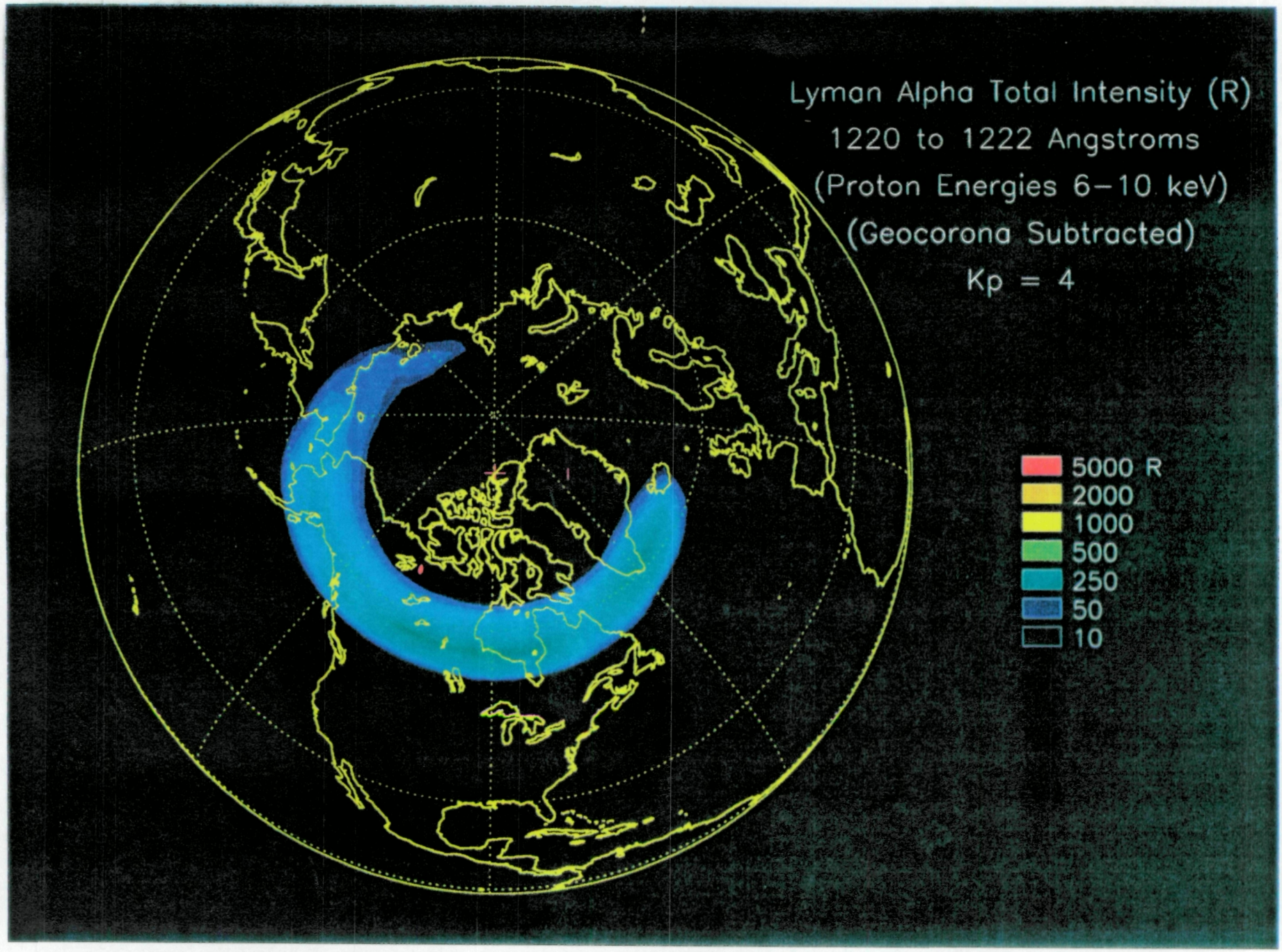


Figure 2.7. Kp=4 proton auroral emissions from 122.0 to 122.2 nm.

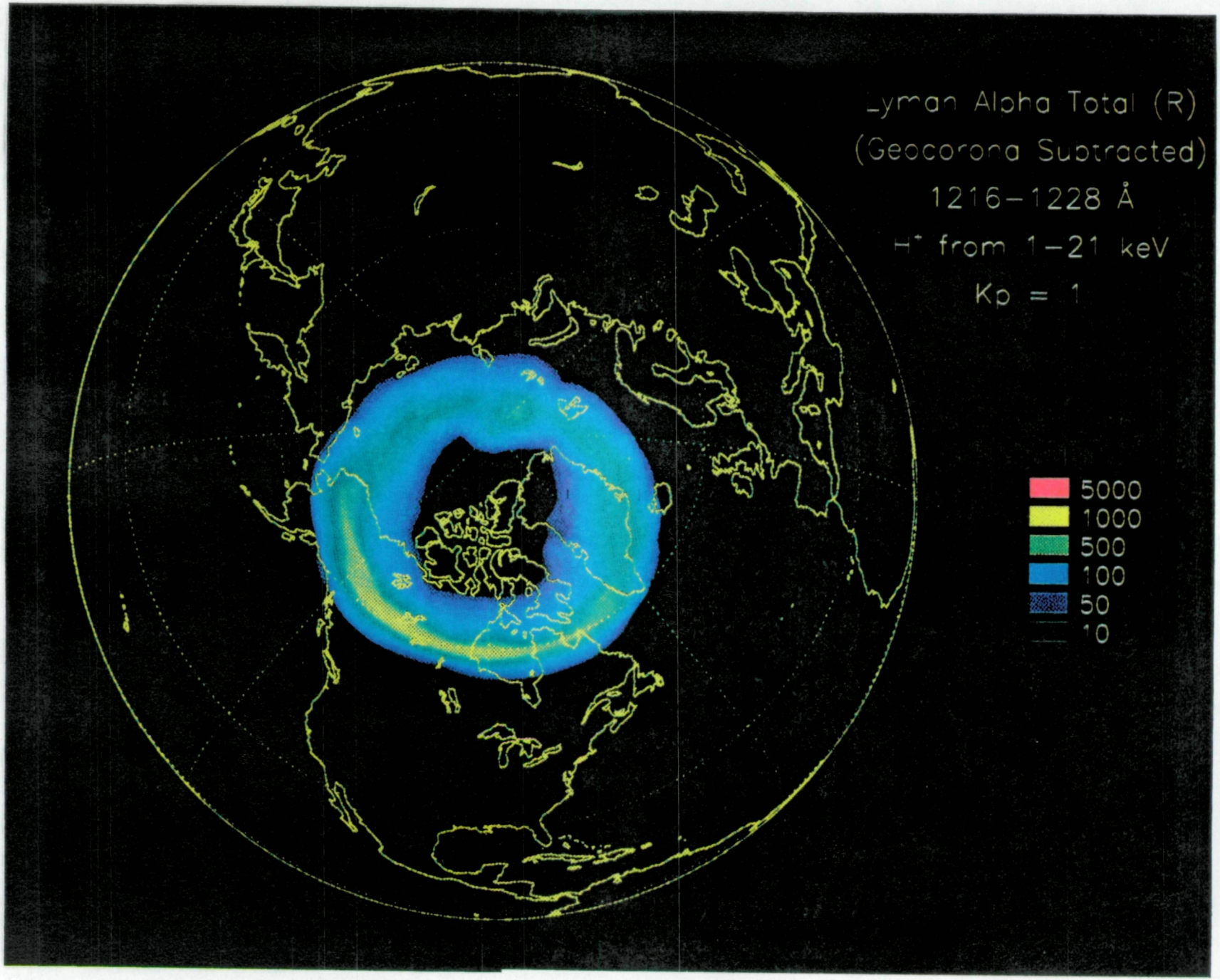


Figure 2.8. Kp-1 proton auroral emissions.

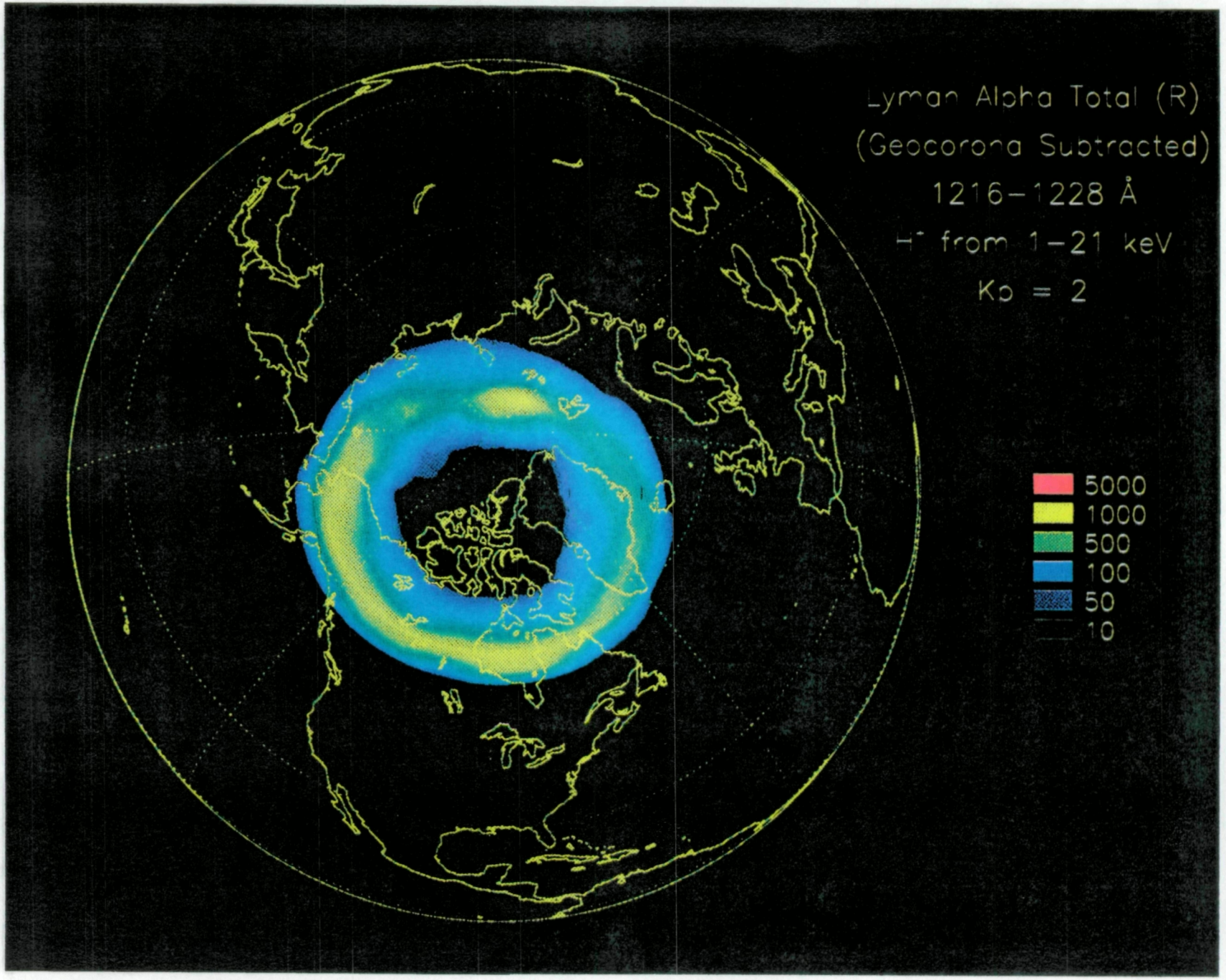


Figure 2.9. Kp=2 proton auroral emissions.

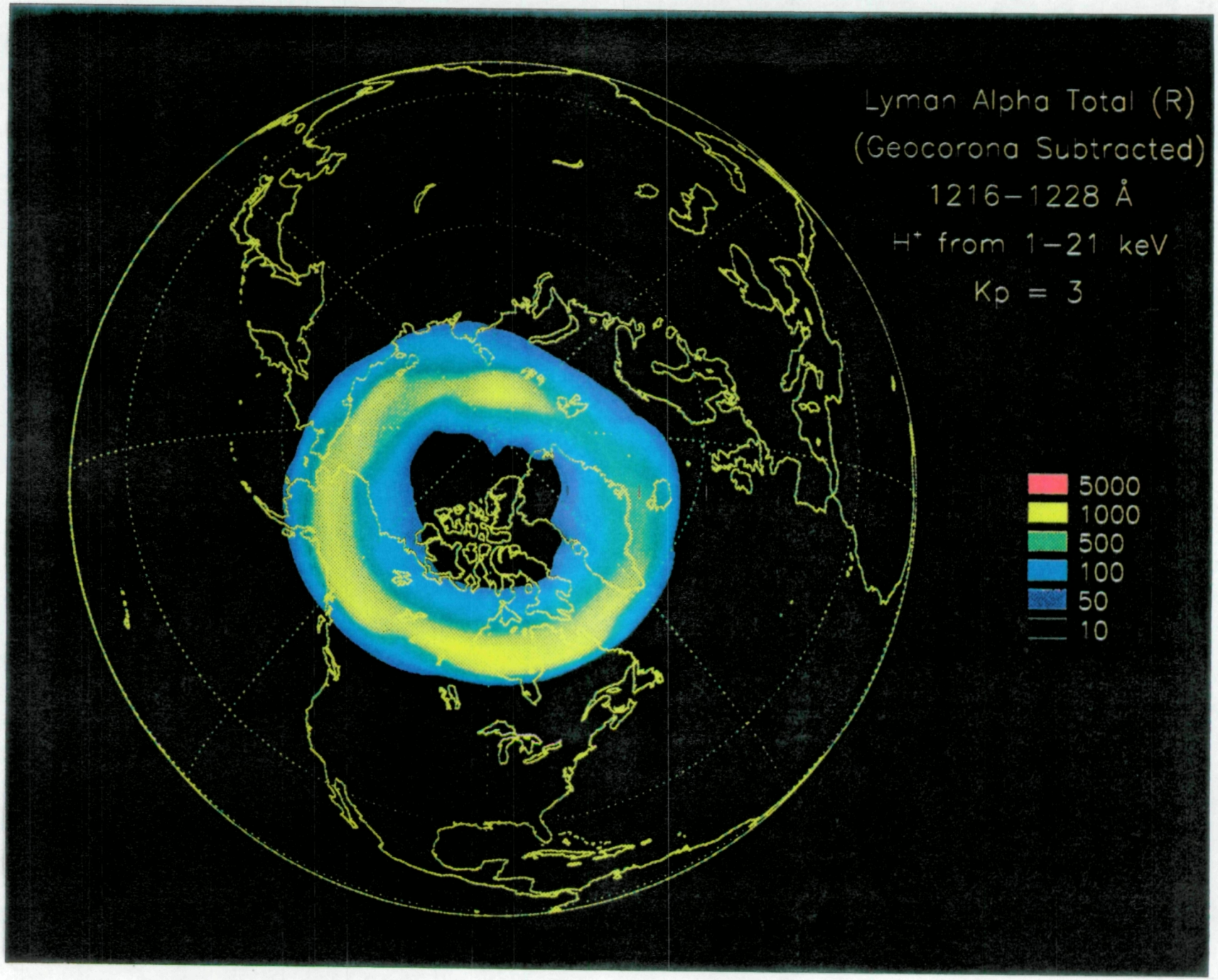


Figure 2.10. Kp=3 proton auroral emissions.

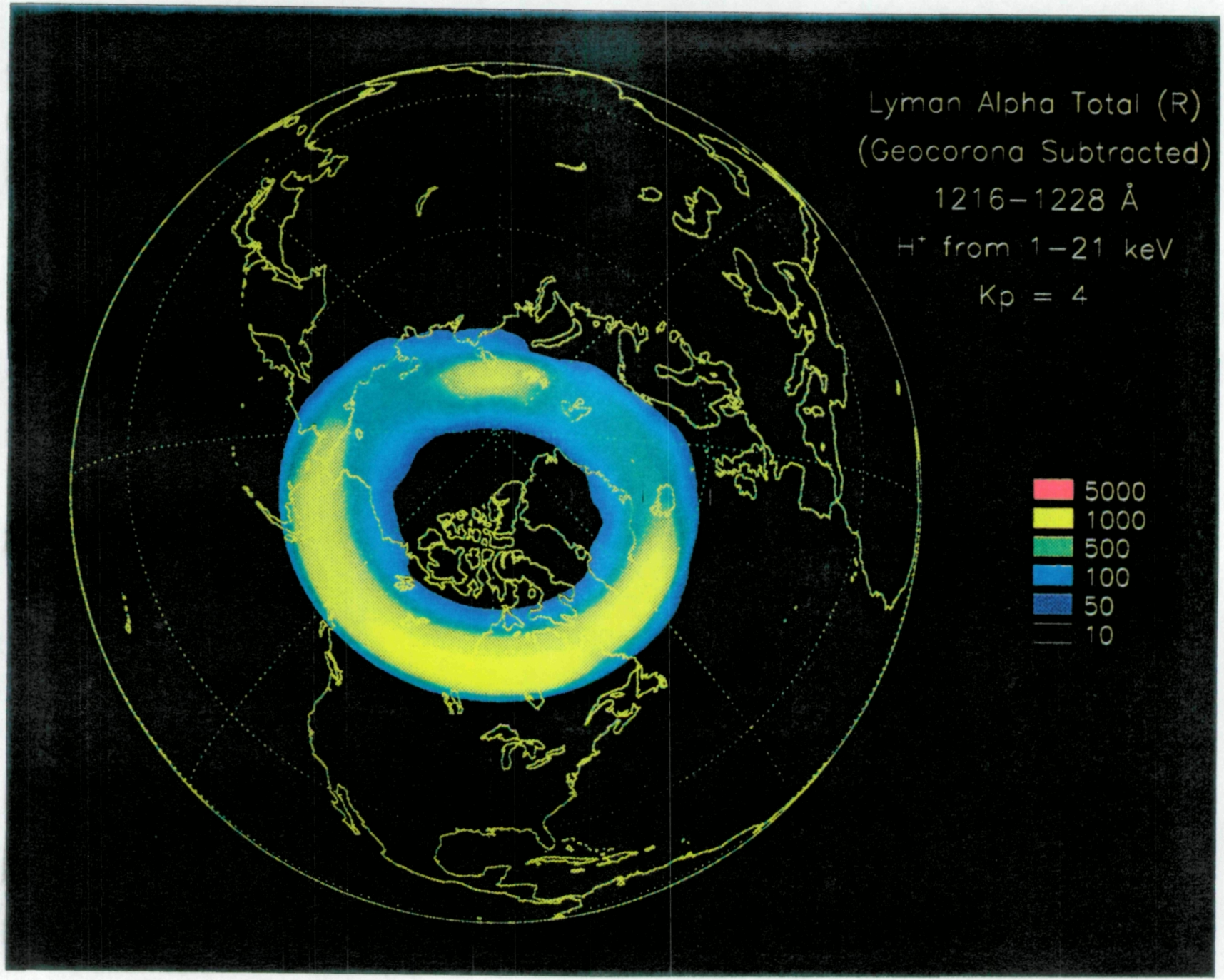


Figure 2.11. Kp-4 proton auroral emissions (same as Figure 2.4)

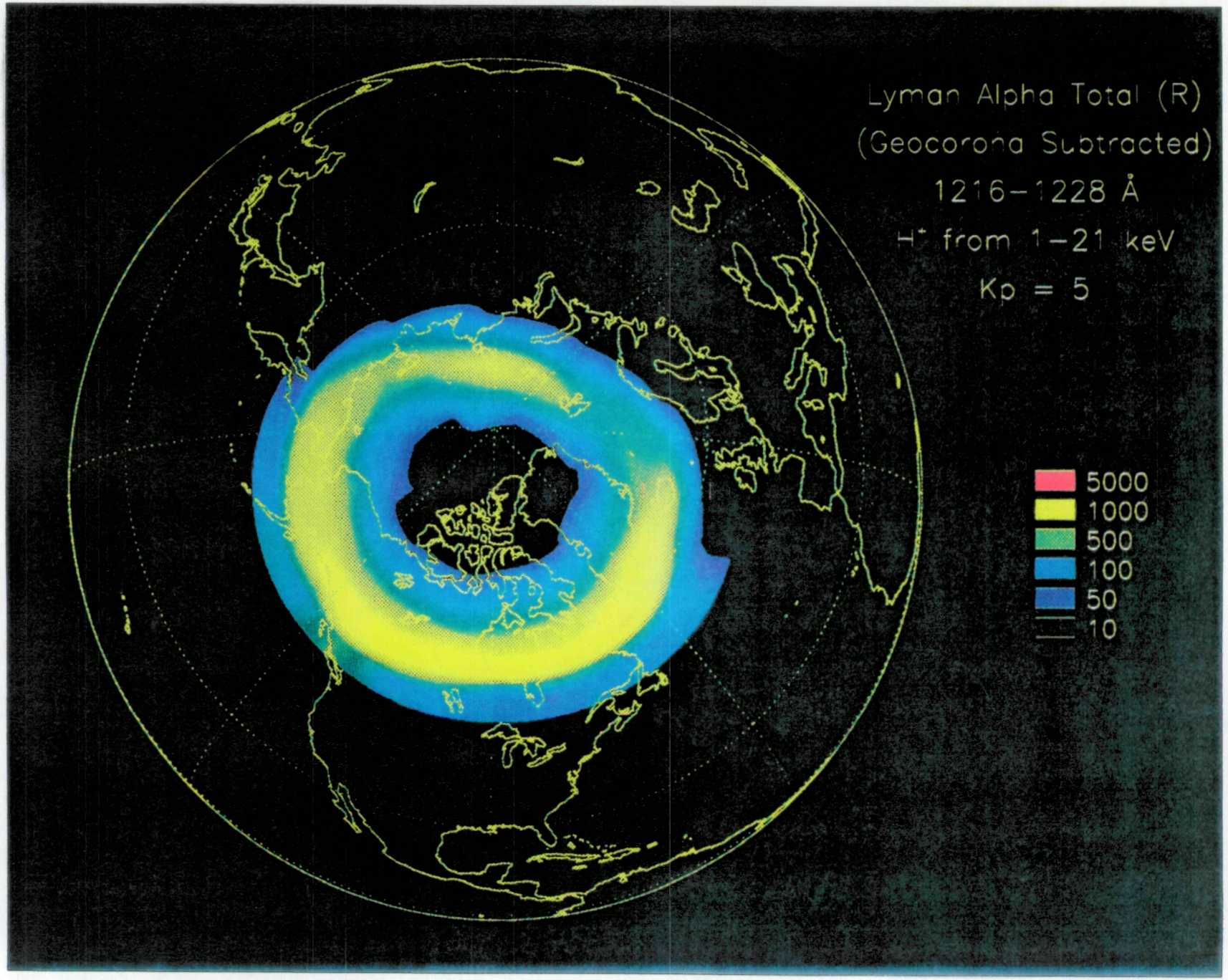


Figure 2.12. Kp=5 proton auroral emissions.

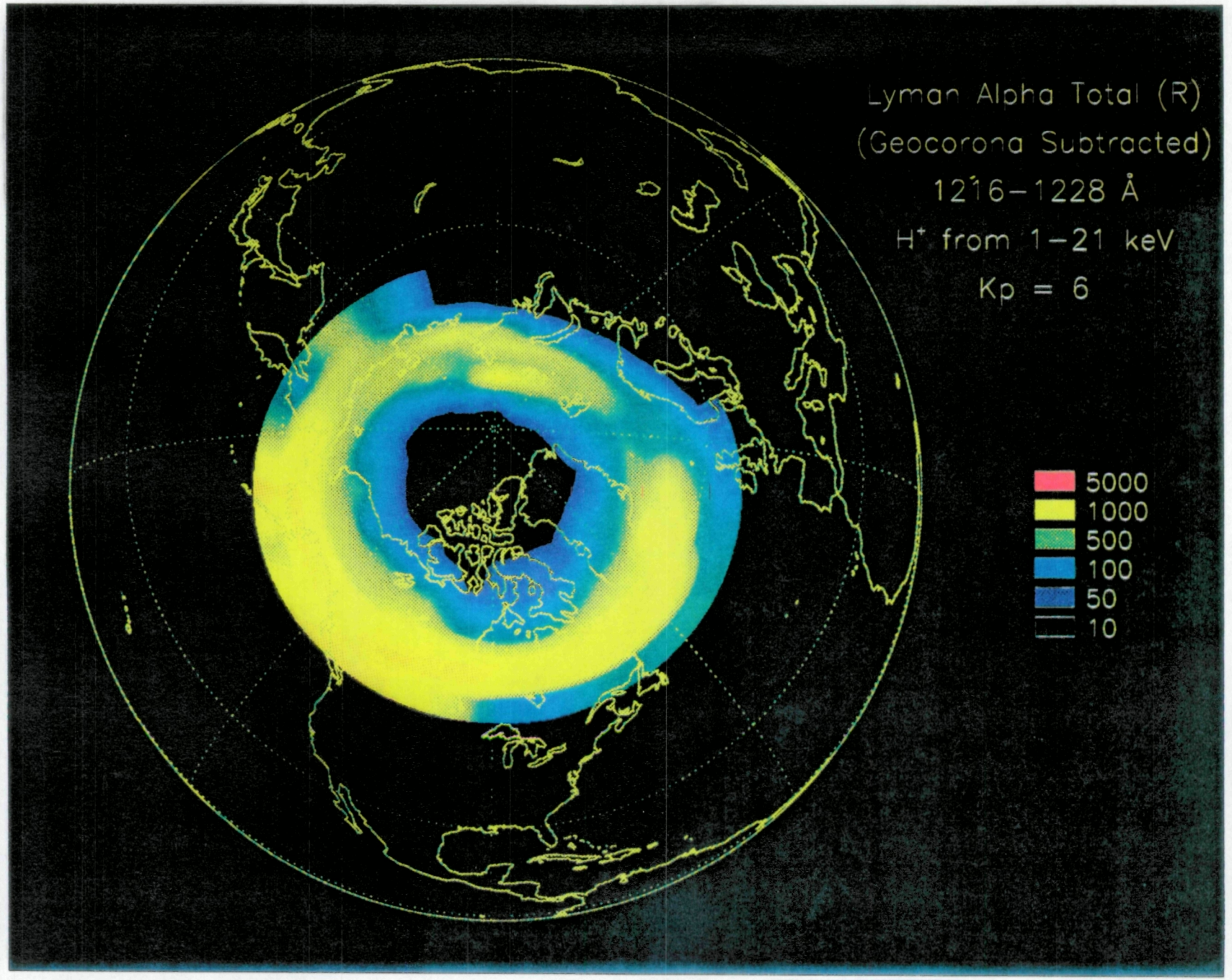


Figure 2.13. Kp=6 proton auroral emissions.

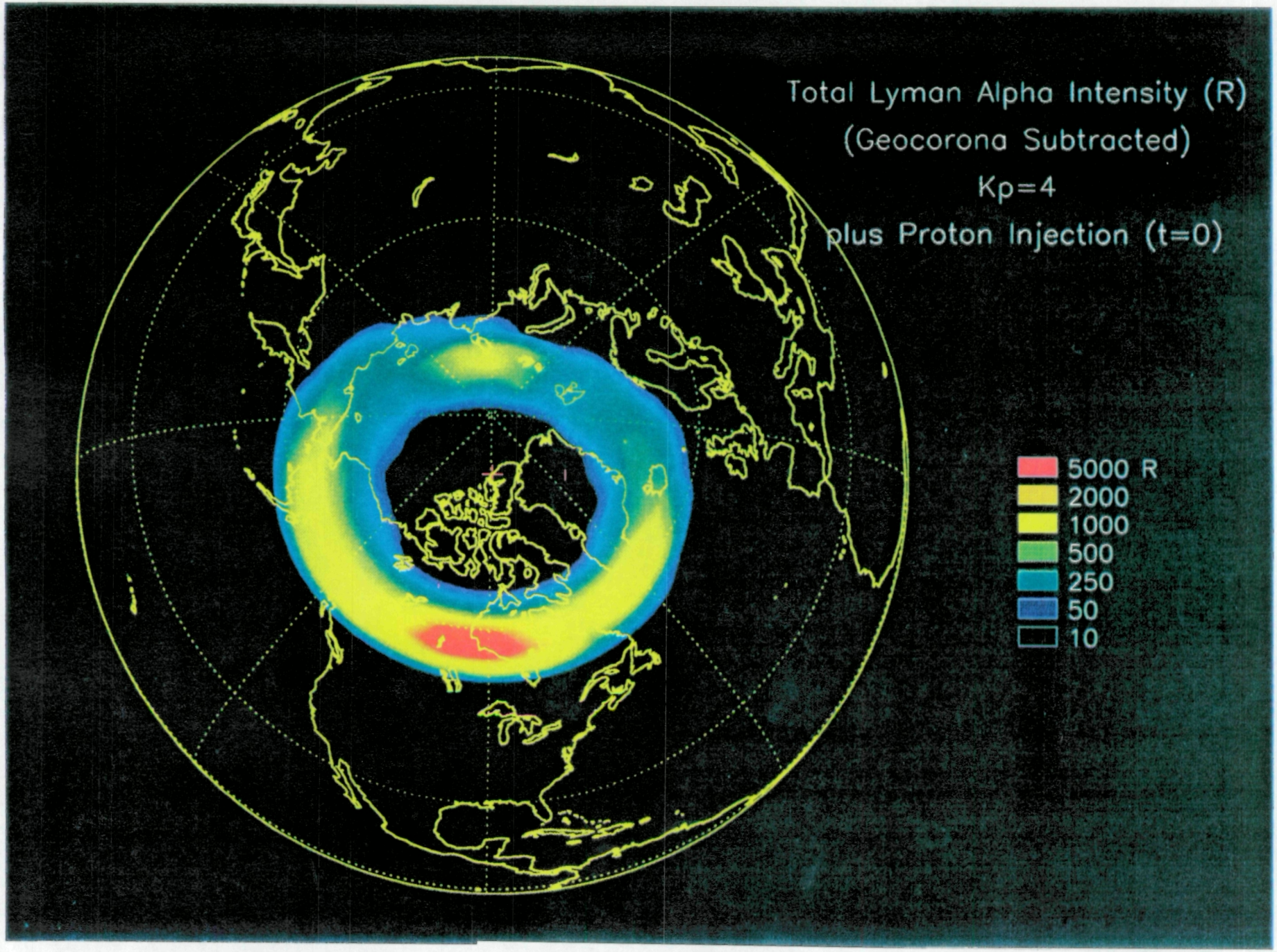


Figure 2.14. Kp-4 proton auroral emissions with a 35 keV, 1 x 10^7 $cm^{-2}s^{-1}$ mono-energetic proton beam injected at midnight.

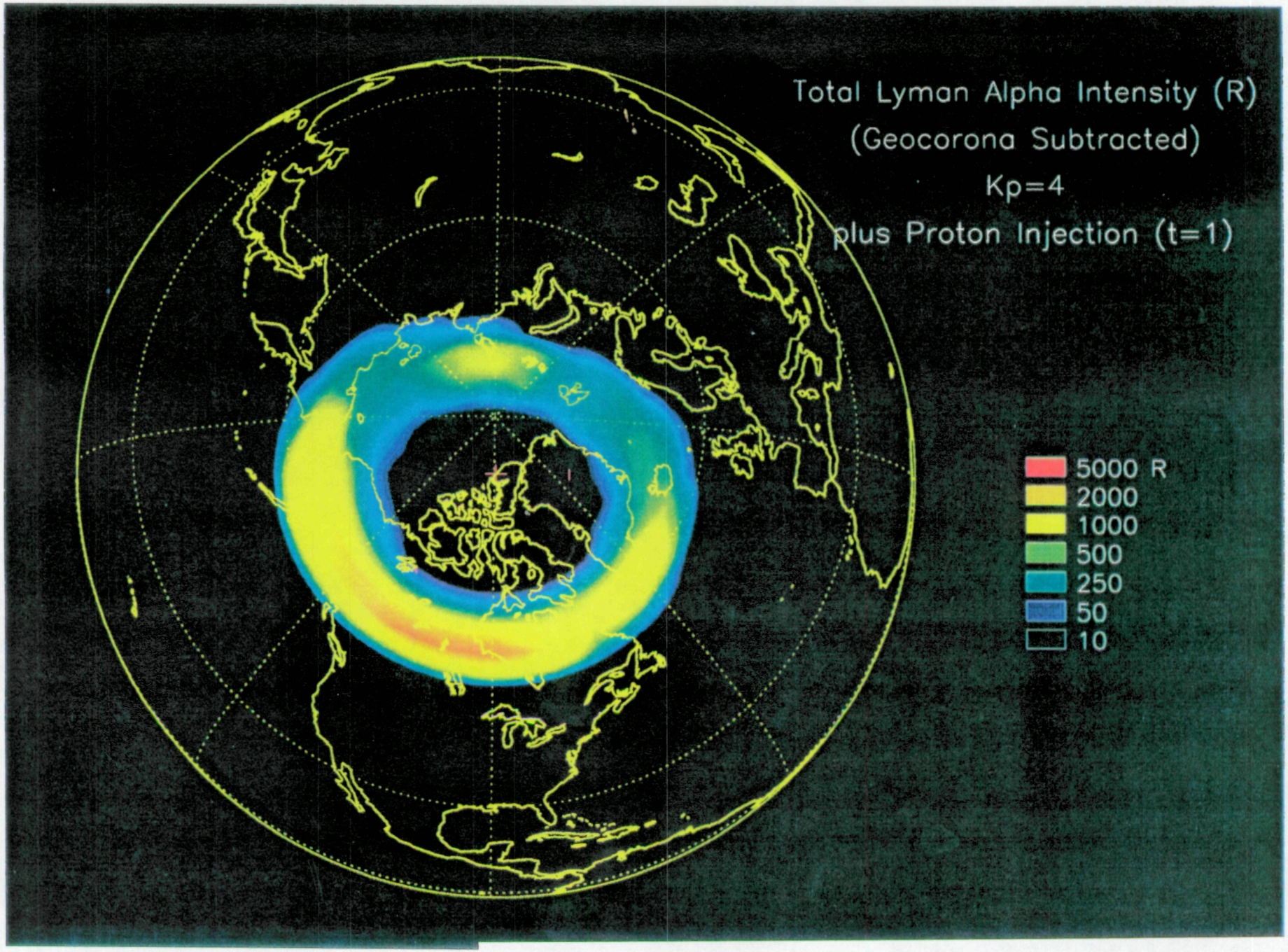


Figure 2.15. Subsequent evolution of the injection front from the mono-energetic beam injection in Figure 2.14.

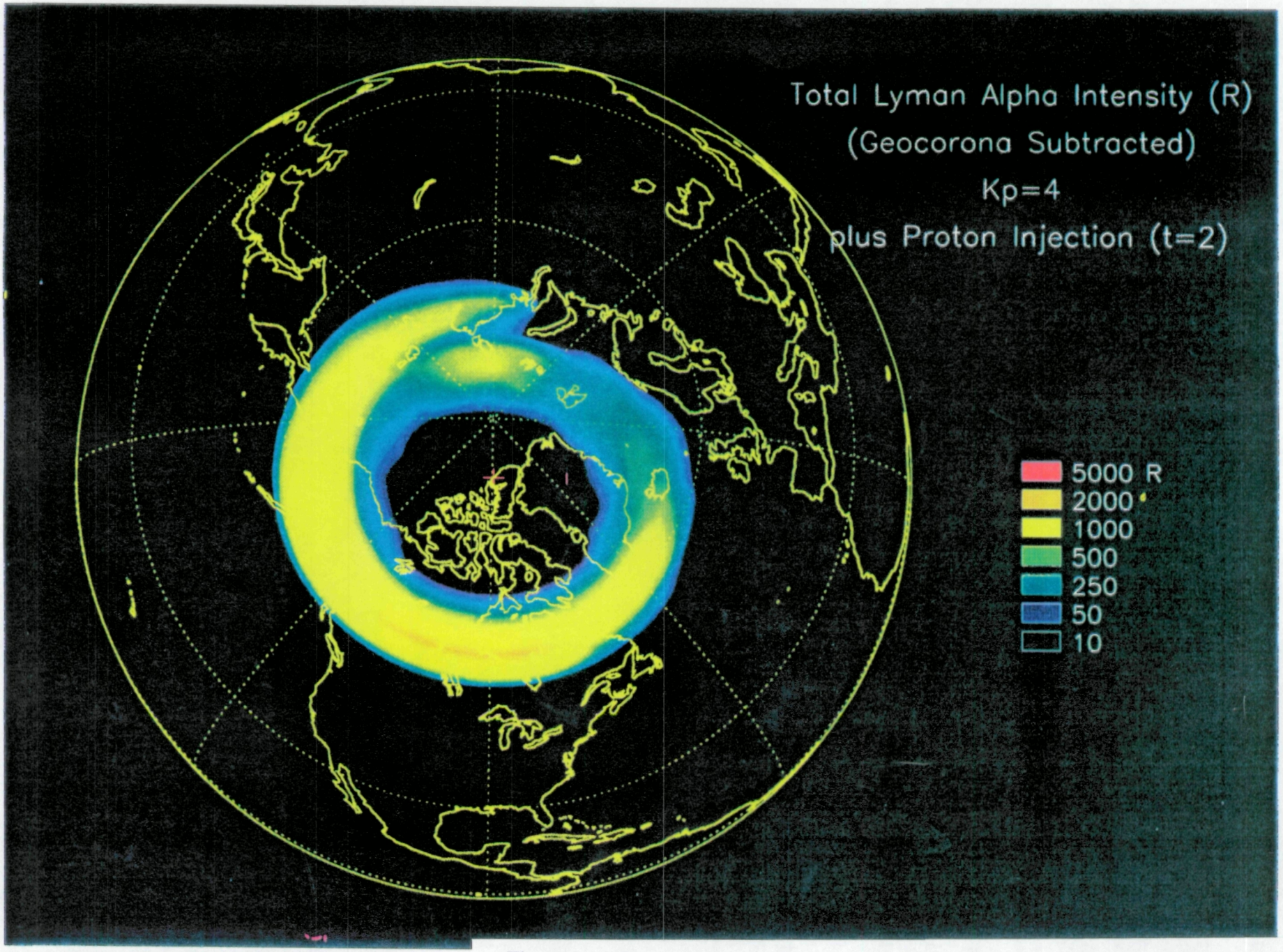


Figure 2.16. Subsequent evolution of the injection front from the mono-energetic beam injection in Figure 2.14.

Because we assume that the precipitating protons are a monoenergetic beam of energy equal to the average energy from Hardy et al. [1989], the calculation of the Doppler spectrum in Figures 2.5, 2.6, and 2.7 is simple. We simply multiply the number of photons in each energy bin by π*integral number flux. A more complicated assumption is that the precipitating protons are in the form of a maxwellian with average energy and total number flux equal to that given by Hardy et al. [1989]. Since we will want to compare the energy spectrum produced by the two different distributions, we want to have a maxwellian that has total integral number flux and average energy equal to that for a monoenergetic beam. From Strickland et al. [1993], we have an equation for the energy flux of a maxwellian which depends on the average energy:

$$J_E = \left[\frac{Q_o}{2\pi E_o^3}\right] E \exp(-E/E_o)$$

where E_o is the average energy from Hardy et al.[1989], E is the energy, J_E is the energy flux $(cm^{-2}\ s\ sr\ eV)^{-1}$ and Q_o is the incident energy flux. The trick is to adjust Q_o such that:

$$\int_{?}^{?} J_E dE = J_{Ntot}$$

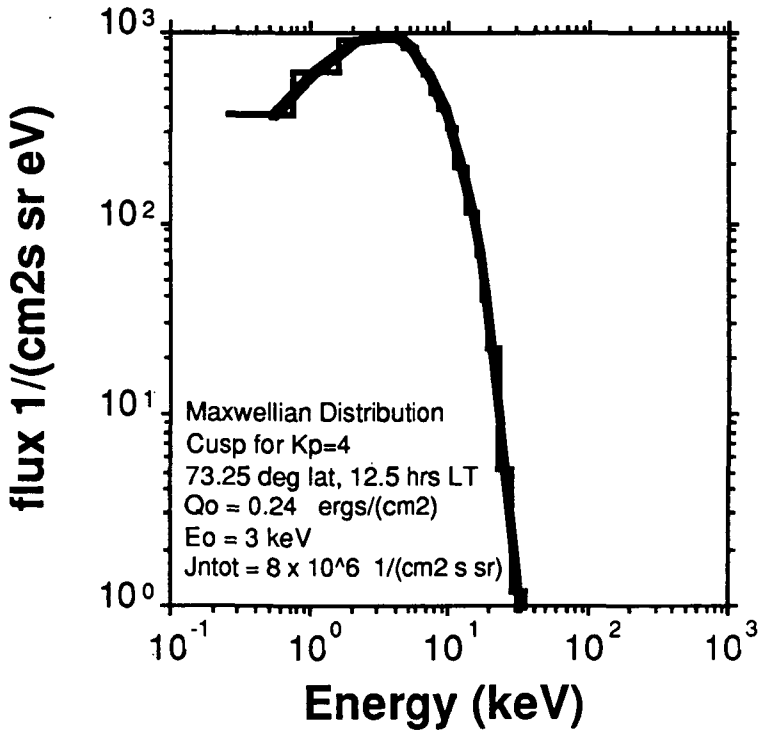


Figure 2.17. Maxwellian distribution in the cusp for Kp=4 The histogram shows the energy bins used to calculate the proton aurora spectrum.

where J_{Ntot} is the total number flux from Hardy et al. [1989]. For example, the cusp in Figure 15 is at 73.75° invariant latitude and 12.5 hours local time. The flux and average energy are:

$J_{Ntot} = 8 \times 10^6\ (cm^2\ s\ sr)^{-1}$

$E_o = 3$ keV

By adjusting Q_o so that $J_{Ntot} = 8 \times 10^6\ (cm^2\ s\ sr)^{-1}$, we obtain the maxwellian distribution that has the same average energy and the same J_{Ntot} as a monoenergetic beam. This distribution for the cusp is shown in Figure 2.17. The total number of photons produced is the sum over the number of photons produced by each individual bin. For the monoenergetic beam, there is only one energy bin centered at 3 keV. However, just like the protons in each energy bin for the maxwellian, as the 3 keV beam decays in energy, it produces photons with smaller Doppler shift.

The differences between maxwellian and monoenergetic beams are considered for four regions: The cusp (near noon), Dawn, Dusk, and Midnight. These regions have the following attributes:

Kp = 4 Hardy et al.	Location Invariant Lat °	Local Time Hr	J_{Ntot} $(cm^2\ s\ sr)^{-1}$	E_o keV	Q_o ergs/(cm^2 s)
Cusp (Noon)	73.75	12.5	8×10^6	3	0.24
Dawn	68.75	5.5	3×10^6	8	0.24
Dusk	63.75	17.5	2×10^6	21	0.44
Midnight	63.75	23.5	6×10^6	17	1.03

Table 3

Figures 2.18 and 2.19 show the resulting spectra from monoenergetic beams and maxwellian distributions, respectively. Also added to this plot is the ~10 kR geocorona line with a line width of ~0.1 Å [Rairden et al., 1986]. Comparing these figures, it can be seen that main difference between the emissions produced by a monoenergetic beam and a maxwellian is the addition of a tail to the spectrum at higher wavelengths. This is due to the higher energy ions in the maxwellian (see Figure 2.17, for example).

An instrument with ~2 Å resolution would be able to resolve the cusp precipitation from the midnight precipitation, even though the precipitation is likely due to the maxwellian distributions. Such wavelength resolution would be necessary to separate the very intense geocoronal Lyman Alpha from the significantly less intense proton aurora.

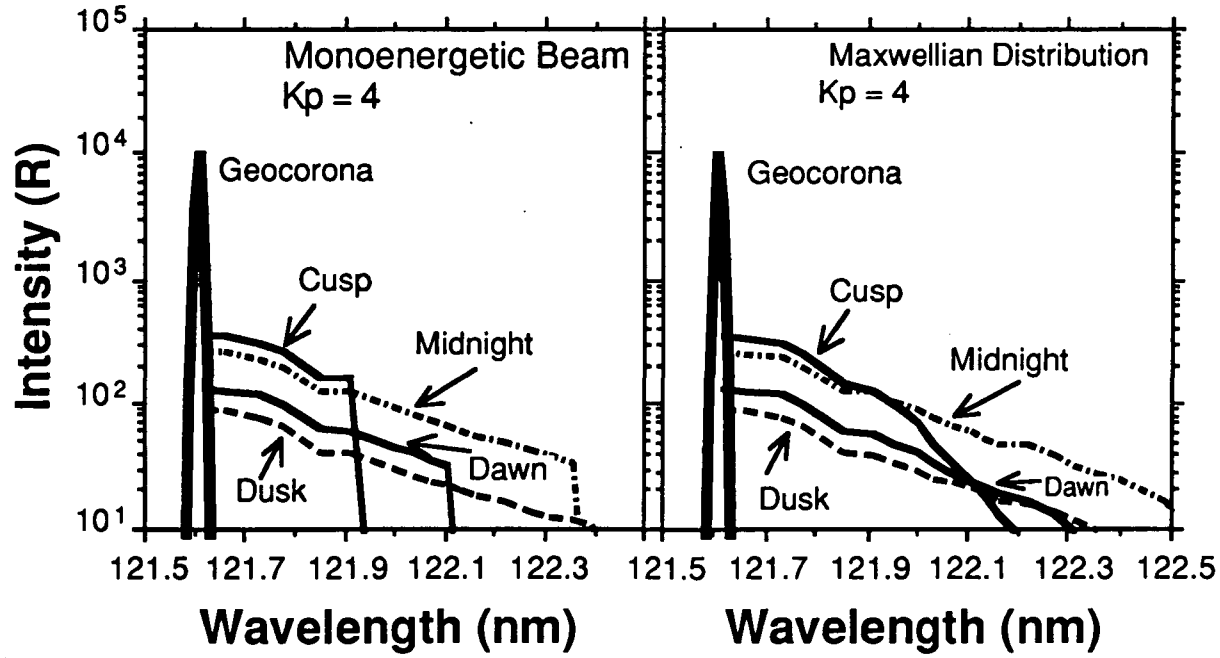


Figure 2.18 and 2.19. Spectra of the geocorona and proton aurora from proton preciptation maps of Hardy et al. [1989]. Comparing the maxwellian to the monoenergetic beam, the main feature is the addition of a tail to the spectrum at higher wavelengths (larger Doppler shifts corresponding to higher energy protons.

2.2 Geocoronal Lyman Alpha

Geocoronal Lyman Alpha emission is generated by resonant scattering of solar Lyman alpha radiation in the Earth's exosphere. The exosphere is a tenuous cloud of neutral H that surrounds the Earth. By measuring the Lyman Alpha intensity as a function of distance from the Earth and modeling the radiative transfer process, the exospheric density profile and temperature can be determined [Chamberlain, 1963; Rairden et al., 1986]. We first summarize the Chamberlain model as discussed by Rairden et al. [1986]

Lyman alpha imagers detect the number of solar Lyman alpha photons scattered into the detector from Hydrogen atoms along the instrument line of sight. Well below the exobase, the geocorona is optically thick, requiring application of radiative transfer theory. Even above this exobase where the geocorona is optically thin, multiple scattering of photons occurs because the higher regions are illuminated above by the sun and below by the scattered photons below the exobase.

The Chamberlain model is a spherically symmetric atomic hydrogen model comprised of three exospheric particle populations. Ballistic particle are those which rise from the exobase with less than escape velocity and then fall back. Escaping particles leave the exobase with greater than escape velocity. Finally, satellite particles are those which are scattered into closed orbits above the exobase by rare collisions in the lower exosphere. The model has 4 parameters:

r_c = exobase critical level

r_{cs} = the degree to which satellite particles are present. This is an upper bound on the perigees of closed orbits

T = Temperature of the exobase

$N(r_c)$ = Density at the exobase

Deviations from spherical symmetry that the model does not account for are 1) diurnal variations in the temperature and density below the exobase, 2) differences in solar illumination between the summer and winter polar regions. These effects will be largest within a few Earth radii. However, it is expected that they will smooth beyond this. One significant exception is an enhancement of H (and therefore Ly α) in the antisunward direction due to radiation pressure scattering of the long flight time H atoms that are greater than a few Earth radii away from the exobase. Evidence for this enhancement was seen in the DE data.

To determine the four parameters in the Chamberlain model, the source function for the Lyman alpha emission is integrated along the line of sight. The source function is defined in terms of the excited H density, the unexcited H density, the lifetime of the emission, the center line absorption cross section and Doppler shift of the emission, and finally the solar Lyman alpha flux.

One of the important uncertainties in the model is the solar Lyman alpha flux. A second is the ad hoc addition of the satellite particle upper bound on the perigees of closed orbits r_{cs}. However,

observations are found to agree reasonably well with this relatively simple model [Rairden et al., 1986]. Deviations from the model represent deviations from spherical symmetry.

Observations of geocoronal H from the DE spacecraft were fit to the Chamberlain model. From these observations, a flux of ~10 kR with a line width of approximately 0.01 nm was determined for the geocorona along the line of sight from the IMI apogee to the atmosphere. Figures 2.4 and 2.5 show how this intense geocorona will nearly overwhelm the weaker proton aurora emissions in the 121.6 nm range. However, a spectrometer with ~0.2 nm resolution could eliminate the geocorona emissions by considering wavelengths longer than 121.6 nm (see Figures 2.18 and 2.19).

2.3. Electron Precipitation and Electron Aurora

Another advantage of the FUV wavelength range over the more familiar visible wavelength range is that it contains several emission lines due to electron precipitation. Global mapping of the energy deposition of precipitating particles and the resulting pattern of conductivity enhancements is a crucial diagnostic in the investigation of the dynamics of the earth's magnetosphere ionosphere system. It is generally agreed that systematic quantitative imaging of the auroral emissions would provide remote sensing measurements of the precipitation energy inputs. Improvement in spectroscopic techniques will enhance the measurements and allow the discrimination of particle species, for example, by separating electron and proton precipitation. Additionally, such techniques could be employed to determine the energy spectrum of the precipitating particles. Previous imaging instruments were generally limited in their spectral resolution and were not able to take advantage of these proposed techniques. In this part of the report, we will examine techniques that can make use of the spectroscopic differences in auroral types and produce remote sensing imaging measurements to permit mapping of the total electron energy deposition, the regions of electron precipitation, and regions with different energies in the precipitating particles. Energy deposition can be determined because the atmosphere is not transparent (due to O_2 photoabsorption) and the ratio of some of these emission lines can be used to obtain the penetration depth of electrons into the atmosphere and therefore the initial energy of the parent electron distributions that create the electron aurora [Strickland et al., 1983].

In addition to some very weak N_2 LBH bands shown, the strongest emission lines in the wavelength range from 120 to 150 nm, the primary emission lines are:

Emission	Wavelength (nm)
N I	120.0
H I (mainly geocoronal + some aurora?)	121.6
N_2 LBH (6,0)	127.5
O I	130.4
N_2 LBH (4,0)	132.5
O I (+ N_2 LBH (3,0))	135.6 (+ 135.4)
N_2 LBH (2,0) (5,2)	138.3
N_2 LBH (1,0) (4,2)	141.5
N_2 LBH (2,1) (5,3)	143.0
N_2 LBH (3,2)	144.0
N_2 LBH (1,1)	146.4
N I (+ N_2 LBH (6,5) (3,3))	149.3

Table 4

The strongest of these lines is the 130.4 nm O I emission line. However, due to multiple scattering in the upper atmosphere, this emission line is not suitable for studying auroral morphology. In fact, the separation of this line from its next strongest line at 135.6 nm places an important wavelength resolution requirement on an FUV imager for the IMI mission.

The N I and O I emissions in Table 4, being atomic emissions, are much more complex than the simple H emissions in the Ly α. They are also produced by electron excitation of the atoms and not proton interaction. The basic process for these emissions is electron bombardment of neutral N_2 and O which leaves a atomic N and O in an excited state [e.g., Meier et al., 1982].

The basic reaction for the N I emissions is:

$$e + N_2 \rightarrow 2e + N^+ + N^*$$
$$N^* \rightarrow N \text{ (or } N^*) + h\nu$$

The basic reaction for the O I emissions is:

$$e + O \rightarrow e + O^*$$
$$O^* \rightarrow O + h\nu$$

The energy of the emission depends on the energy of the initial incident electron. Most of these electrons are not the primary electrons precipitating into the atmosphere from the magnetosphere. These electrons are at too high an energy. Instead, the electron spectrum develops a broad low energy tail due to collisions in the upper atmosphere so that by the time the electron distribution arrives at ~110 km altitude (where most of the energy is deposited), it has a very intense low energy tail extending to less than 10 eV [Meier et al., 1982]. The N and O bands are produced by ~30 eV secondary electrons in the atmosphere. The longer wavelength LBH bands (discussed next) are produced by ~10 eV electrons.

The ground states of O and N are:

O: $1s^2 2s^2 2p^4$

N: $1s^2 2s^2 2p^3$

However, the electron bombardment leaves the atoms in highly excited states. The following table shows the transition between the excited states (not necessarily to the ground state) for the major lines in the 120 - 150 nm wavelength range in Figure 1 and Table 1.

Emission	Transition
N I 120.0 nm	$1s^1 2s^0 2p^2 3s^4\ P \rightarrow 2s^2 2p^3\ ^4S^0$
N I 149.3 nm	$2s^1 2p^2 3s^2\ P \rightarrow 2s^2 2p^3\ ^2D^0$
O I 130.4 nm	$2s^0 2p^3 3s^3\ S \rightarrow 2p^4\ ^3P$
O I 135.6 nm	$1s^0 2s^0 2p^3 3s^5\ S^0 \rightarrow 2p^4\ ^3P$

Table 4

It is apparent that the short wavelength emissions occur in atoms that are very nearly ionized. In fact, the maximum electron energies approach 100 eV, which is the energies where the ionized component created by the bombardment is also excited (creating, for example, N II emission).

The LBH (Lyman-Birge-Hopfield) bands also arise from electron bombardment of N_2. However, these are vibrational/rotational states of the molecule so there is no dissociation. Because the molecule must stay intact, the electron energies must be very low, less than 13 eV [Meier et al., 1982]. The excitation process is:

$$e + N_2 \rightarrow e + N_2^*$$

$$N_2^* \rightarrow N_2 + h\nu \qquad\qquad a^1\Pi_g \rightarrow X^1\Sigma_g^+$$

The wide variety of lines in Table 3 arises because of the large number of vibrational states of the molecule. Each emission "line" actually has a complex structure that is not resolved. This complex structure is due to the large number of rotational levels [Vallance Jones 1991].

While it was not the perview of this proposed research to model the electron aurora, we include some analogous modeling to the proton aurora results in Figures 2.4 to 2.6 for completeness.

As with the protons, the electron aurora were estimated from the statistical results of the electron precipitation measured by low altitude polar spacecraft [see Hardy et al., 1985]. Figures 2.20 and 2.21 show the mean electron energy bins and mean electron number flux bins used to model the electron auroral emissions. These figures are analogous to the proton energy and number flux bins in Figures 2.4 and 2.5. The bins are one half hour in local time by 0.25° in latitude. North is to the top and the high energy dawnside electron precipitation is clearly evident in Figure 2.20.

Similar to the proton emission estimates, we need the photon yield per precipitating electron to convert the energies and number fluxes in Figures 2.20 and 2.21 into electron aurora emission intensity. These were obtained from the recent detailed modeling of the electron and proton auroral precipitation by Strickland et al. [1993]. The results for two wavelengths (135.6 OI and 146.4 nm LBH) for moderate geomagnetic activity (Kp=4) are shown in Figures 2.22 and 2.23, respectively. These emissions should be compared to the proton emissions in Figure 2.5. Since the electron emissions in Figures 2.22 and 2.23 include atmospheric absorption, the ratio of the intensities can be used to invert the emissions and obtain again the precipitating electron energies in

Figure 2.20. An example of this process is shown in Strickland et al. [1993]. In addition, these observations can be used to obtain ionospheric conductivities.

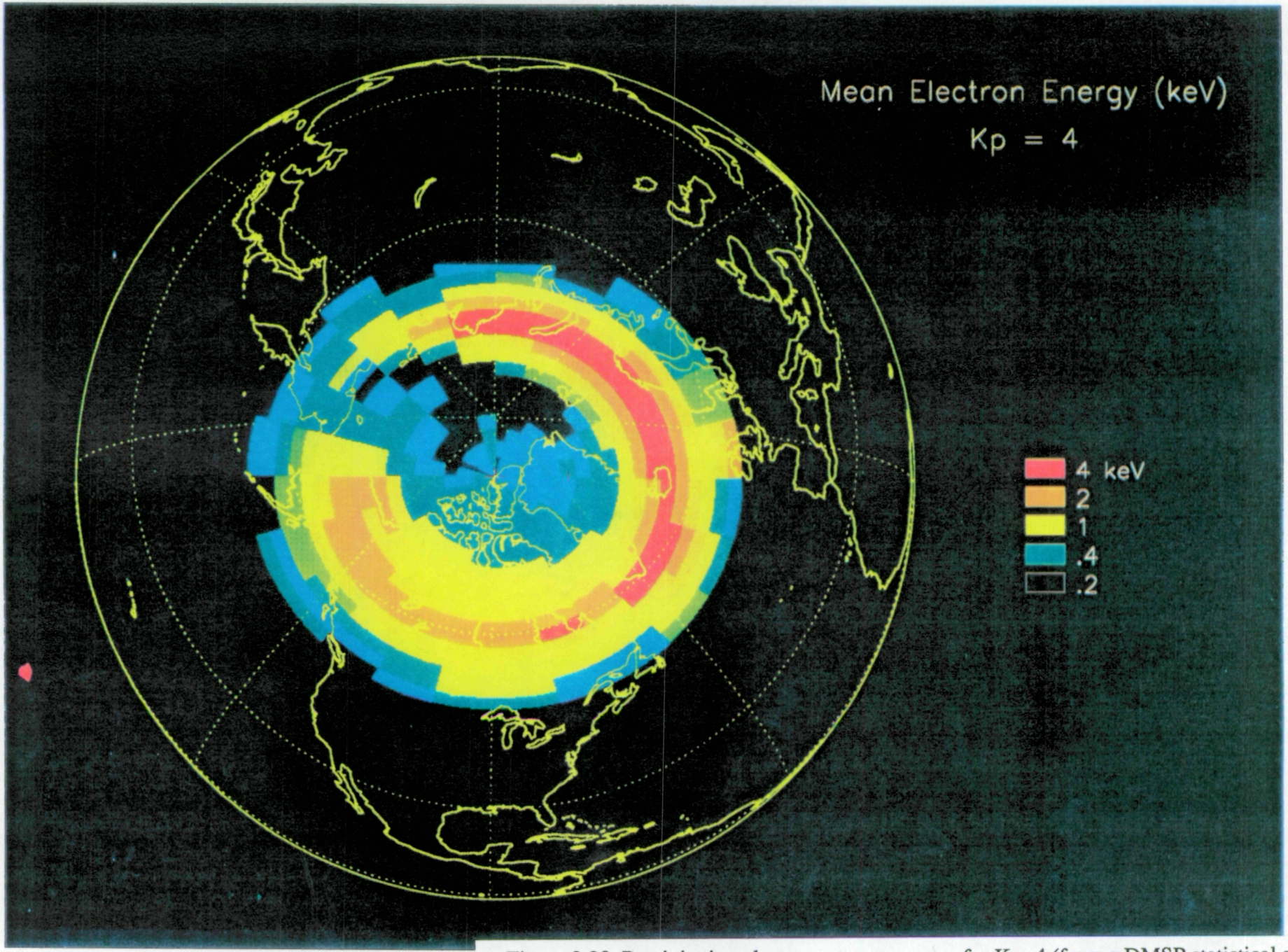


Figure 2.20. Precipitating electron average energy for Kp=4 (from a DMSP statistical study by Hardy et al. [1985]). Local noon is at the top. Very hard precipitation occurs in the dawn sector.

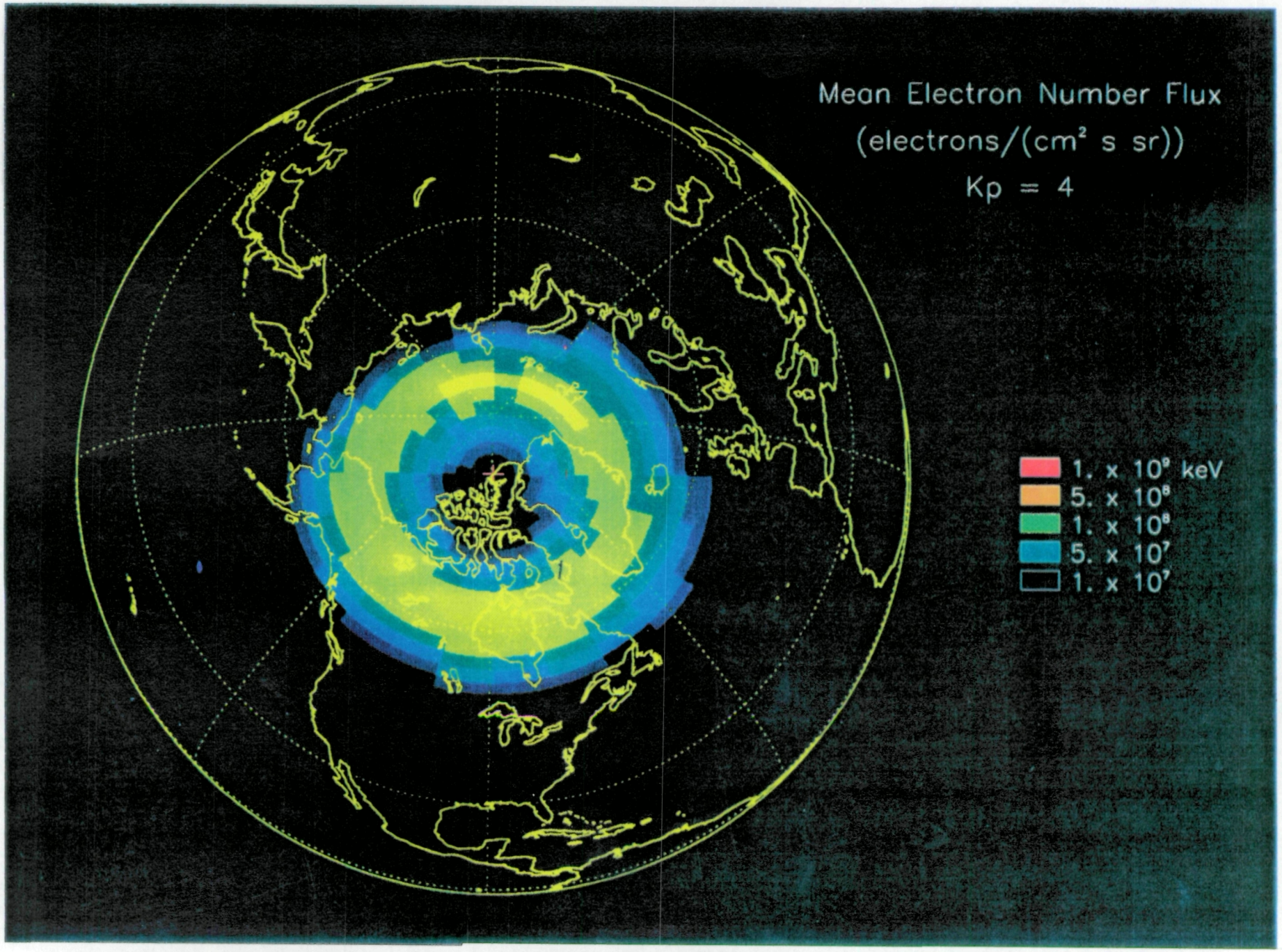


Figure 2.21. Precipitating electron number flux for Kp=4 (from a DMSP statistical study by Hardy et al. [1985]). Intense precipitation occurs in the cusp region .

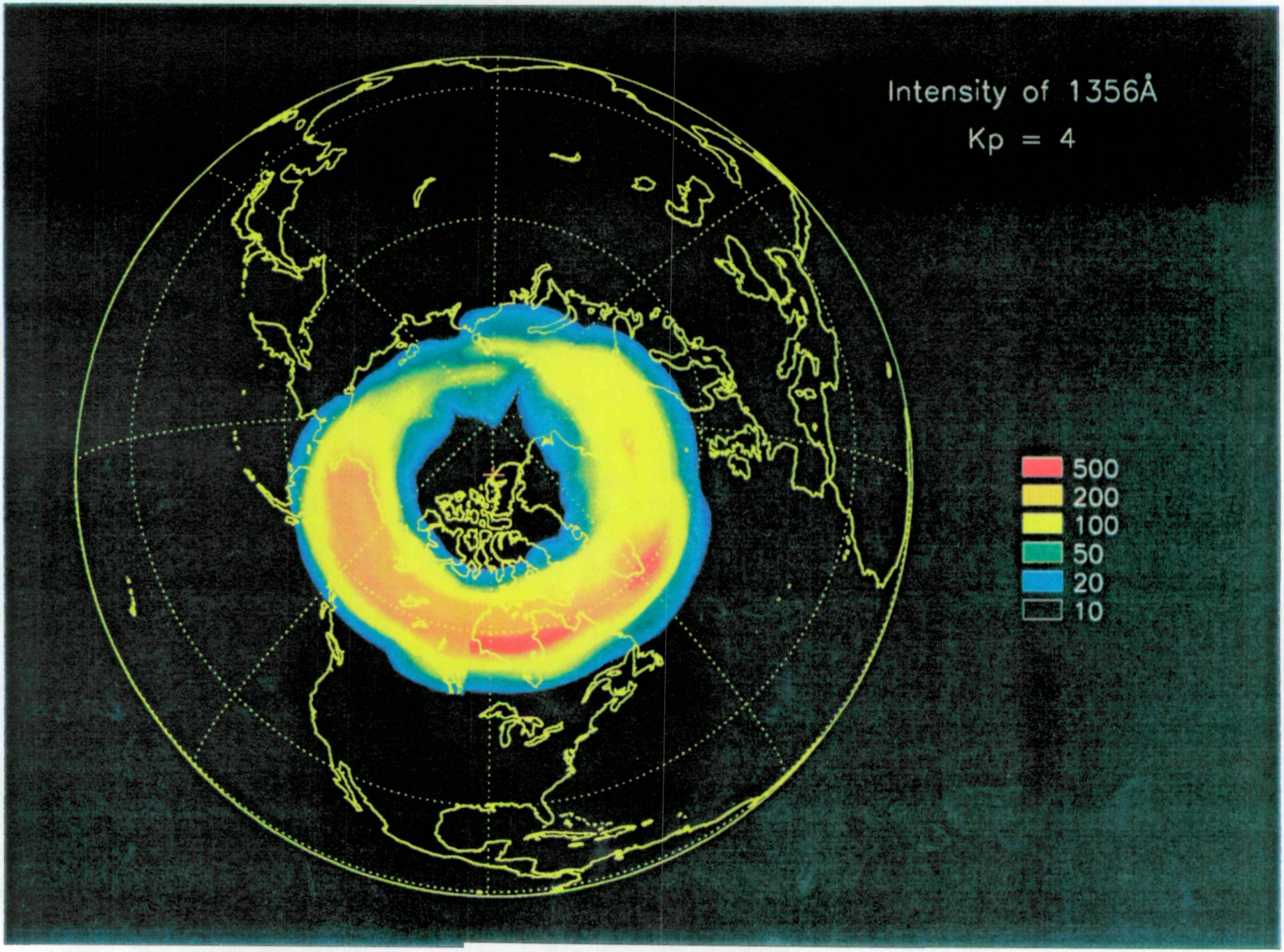


Figure 2.22. Resulting 135.6 nm flux from the average energy and number flux models in Figures 2.20 and 2.21. A clear cusp region is seen near local noon.

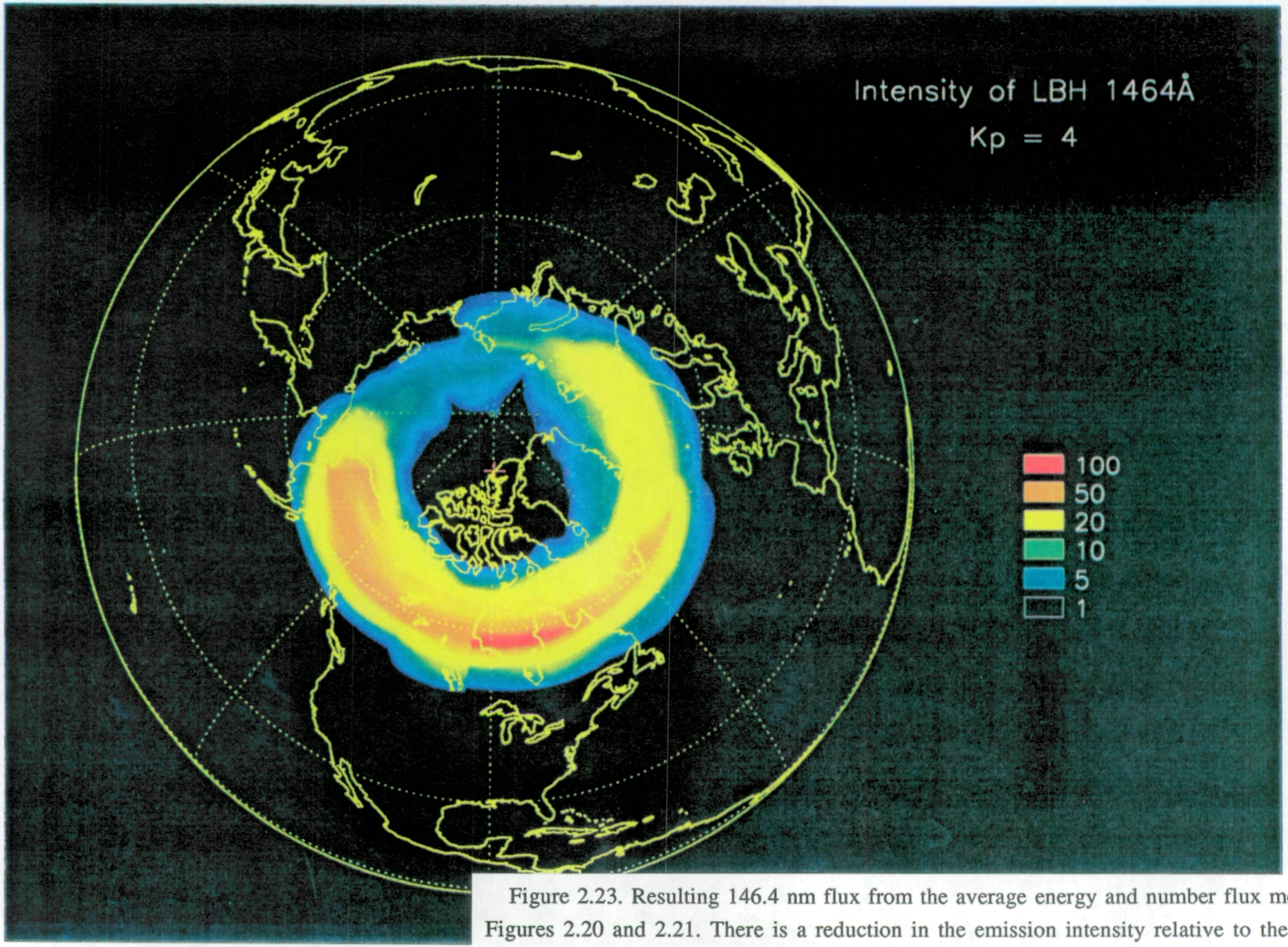


Figure 2.23. Resulting 146.4 nm flux from the average energy and number flux models in Figures 2.20 and 2.21. There is a reduction in the emission intensity relative to the shorter wavelength emissions in Figure 2.22 due to atmospheric oxygen absorption. The magnitude of this absorption is proportional to the energy of the precipitating electrons so that the ratio of the shorter and longer wavelength emissions can be used to determine the precipitating electron energy.

3. Imaging Techniques

3.1 Introduction and Basic Concepts

It is our task to develop a concept for an instrument capable of satisfying the requirements in this report on magnetospheric imaging in the FUV wavelength region while maximizing the science for a minimum resource including cost utilization. In almost every form of imaging including magnetospheric imaging the information collection is limited by the signal to noise ratio of the detected image. It is therefore necessary to discuss the general baseline considerations for imaging requirements.

From the baseline IMI study it appears that that the instruments have to operate on a spinning platform in an eccentric orbit of apogee of 6-9 Re and perigee of a few hundred km or more. For each imager there are several broad scientific tasks can be calculated by simply otaining the number P of "equivalent quanta" collected on each "pixel". In this context each pixel defines the range of spatial region in which during an "exposure time" the intensity of the image can be regarded as a constant. The magnitude of the fluctuation in the number P from one exposure to the next defines the intensity detection threshold because if we wanted to determine which was brighter between two pixels P' and P" of approximately equal intensity then limit is set by the mean fluctuations in the two incident fluxes. Thus the fluctuations set the limit to imaging because in general imaging is the detection of contrast differences between adjacent pixels. The information content i.e the number pixels and the number of detectable shades in an image are described by the signal to noise ratio which is calculated from fluctuations and the mean intensity level of the pixels . The signal to noise ratio in each pixel is ultimately limited by the number of input quanta and the effciency of processing these quanta. It is therefore important to realize that the theoretically optimized signal to noise ratio depends on the size of the detector and the number of photons which are gathered during the exposure time. This is the best possible signal to noise ratio. After that the instrument regardless of its internal technology can only worsen the signal to noise ratio and by losing information associated with the collected photons. For example the image amplification properties of the instrument do not improve the intrinsic signal to noise ratio.

In the discussion of imagers processing quantized information such as photons or neutral particles the number of quanta collected defines the highest possible signal to noise ratio. In comparing images it is necessary discuss what is the size of the equivalent aperture that is how large a photon bucket the instrument represents. Theoretically a weak source can always be made detectable by simply scaling up the geometric size of the detector. To calculate the effectiveness of a certain size imager we need to establish a common basis for all types of imaging detectors. We will proceed by considering a hypothetical fully multiplexed staring imager which simultaneously collects quanta into all the pixels through the entire exposure period. Let us also introduce the

concept of a super pixel, the image area which corresponds to the spatial region which has to be resolved by the imager to satisfy a particular scientific spatial detection requirement. We shall show that the signal the number of counts in a super pixel is the product of photon flux at the object (photons/sec/cm^2/str), the solid angle subtended by the super pixel (str) at the imager, the exposure time (s), the area of the instrument aperture (cm^2) and the efficiency of conversion from input particles to recordable numerical information.

Note that the first three factors are to do with the nature of the phenomena and the scientific requirements which places their limits on the spatial resolution and the exposure time. Only the last two factors are effected by the instrument. In fact the instrument factors could be combined into an equivalent aperture which would be the product of the aperture and the detector efficiency.

The linear size of the super pixel on Figure 3.1.1 is denoted by d_o which when projected on the focal plane of the imager becomes d_i. The photons arriving into the super pixel on the focal plane can be calculated by first principles. If the brightness of the image is J photons sec^{-1} ster^{-1} cm^{-2}, (Note the convertion from photon flux to Rayleighs is one Rayleigh is $10^6 / (4\pi) \sim 80{,}000$ photons sec^{-1} ster^{-1} cm^{-2}) the solid angle subtended by the aperture of the imager is $d\Omega_o$ and the area of a super pixel is d_o^2 at the region to be imaged then the number of photons arriving at the aperture of the imager is:

$$P = J\, d_o^2\ d\Omega_o \text{ photons sec}^{-1}.$$

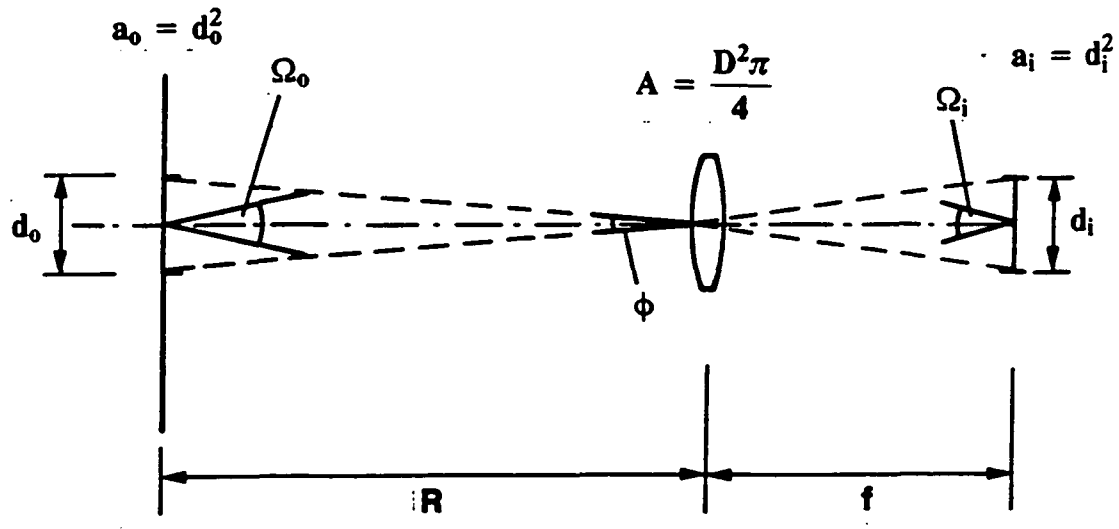
As shown on Figure 3.1.1 this is the same as

$$P = J\, A\, \phi^2 = J\, A\, d\Omega_p,$$

where A is the area of the open aperture (lens or pinhole) of the imager and $d\Omega_p$ is the solid angle or field of view corresponding the super pixel. In order to calculate the number of counts per second it is necessary to include the efficiency, ε of converting the photons into counts. Thus the count rate

$$dC/dt = \varepsilon\, P \text{ counts per second.}$$

The photons are collected on each super pixel at this rate. The exposure time, dt has to be limited to a time shorter than the temporal scale associated with the variability of the phenomena or the motion of the platform. Accordingly the number of recordable numerical events or counts, C which can be collected to produce a useful measurement is:



d_o = PIXEL AT OBJECT

d_i = PIXEL AT IMAGE

R = RANGE

f = FOCAL LENGTH

Ω_o = ANGLE SUBTENDED BY DETECTOR AT OBJECT

ϕ = FIELD OF VIEW ANGLE

Ω_i = ANGLE SUBTENDED BY OPTICS AT IMAGE (CONE ANGLE DEFINED BY F/NUMBER)

FLUX AT OBJECT PLANE;

$$F = d_o^2 \, \Omega_o^2 \, \frac{\pi}{4}$$

BUT

$$d_o = \phi R \qquad \Omega_o = \frac{D}{R}$$

$$F = \phi^2 \, \frac{D^2\pi}{4} = \phi^2 A$$

Figure 3.1.1. General baseline considerations and definitions for any imager.

$$C = J\,A\,d\Omega_p\,\varepsilon\,dt$$

In summary for a stationary staring imager the counts in the exposure time are the product of the photon flux at the source in photons sec^{-1} cm^{-2} sr^{-1}, the area of the imager aperture A, the solid angle represented by the field of view, the conversion efficiency and the exposure time.

This holds for staring imagers filtered or unfiltered. For a rotating platform the collection efficiency is further impaired because the time the imager is staring in the correct direction is further restricted. A spectrometer may be further limited by the fact that the geometric factor represented by $A\,d\Omega_p$ cannot be squeezed through the slit by the grating.

For an optimally designed spectrometer therefore we have to obey an additional restriction. For simplicity let us assume that the imaging spectrometer is viewing square pixels through a square entrance aperture and that $A\,d\Omega_p$ is equal to $L^2\varpi^2$ where L is the linear dimension of the entrance aperture and ω is the field ov view subtended by each pixel. In order to process the full etendu of $A\,d\Omega_p$ the spectrometer has to satisfy the condition that $D\,d\alpha > L\varpi$ in the direction perpendicular to the grating rulings where D is the dimension of the grating and where:

$$d\alpha = d\lambda x(\lambda^2 + l^2)^{-1/2}$$

In this latter expression l is the spatial distance between consecutive grating rulings. High spectral resulution i.e small $d\lambda$ could require a large grating D to compensate for the small $d\alpha$. The expression also shows that small values of l that is dense grating rulings are highly desirable. But even at very high ruling densities there is a limit where $d\alpha$ becomes $d\lambda/\lambda$. This highligths one disadvantage of a spectrometer which for high spectral resolution restricts the etendu unless the grating can be made very large.

Another serious disadvantage of the spectrometer is that in general a slit only covers a small part of the image at any one time. The higher the spectral resolution requires narrower slits and the scanning of the slit across the image becomes more and more time conusuming. These disadvantages drove us to consider filters and especially the hydrogen cell imager.

In the discussion below we have listed a number of tasks for each instrument. Then we have associated a spatial scale with each phenomena. The requirement is that each instrument has to have resolution which is at least as good as the most stringent of the requirements. At the same time the imager has to be able to operate at full efficiency in other applications requiring coarser resolutions. It is assumed therefore that when imaging the coarser resolution object the imager is capable of utilizing full efficiency and that the signal to noise ratio is preserved/enhanced when observing the coarser resolution image.

The above considerations apply to an imager which is simultaneously staring at all pixels from a stable platform. Imagers on rotating platforms, scanning line imagers (e.g. spectrometers), scanning photometers all have commutate the pixels during the observation period and they have a fundamental inefficiency associated with the process which can be expressed in terms of a duty cycle, D_{cyc} which is the fraction of the time during which photons from a superpixel are actively recorded during the nominal exposure time.

Definition of quantities:

R = range the linear distance of the object from the imager (km).

dR = Size of region of interest or super pixel i.e. the region which is to be resolved (km).

A = Area of collecting aperture (cm^2).

dt = Time scale of phenomena, which is the maximum permissible exposure time (sec).

ε = Conversion efficiency from input quanta to recordable events (fraction no units).

D_{cyc} = Duty cycle (fraction no units).

3.2 Hydrogen Cell for geocorona Lyman alpha detection

In prior chapters we have discussed the desirability of making simultaneous measurements of Lyman alpha from the aurora and the geocorona. As we have discussed that the geocoroana and the aurora needs to be separated by techniques other thatn the subtraction a modelled geocoronal profile. This is because the quantum fluctuations in the geocoronal intensity of over 10 kR can easily overwhelm the signal generated by the energetic protons.

For effective separations of the two Lyman alpha components spectral resolution of the order of no more than 2 Angstroms are needed. The technology for filters in the Lyman alpha wavelength region are not available to make such wavelength discrimination. As we have seen a spectrometer type imager can do the separation but at the price of a somwhat inceased instrument size. Another technique is available which uses a Hydrogen gas cell.

Lyman alpha emission is from the ground state of Hydrogen and as such it is strongly scattered by atomic hydrogen gas. This property provides the technique of using a filter consising of a cell field with Hydrogen gas. The cell contains a filament which is heated to high temperature to dissociate the molecular Hydrogen and also to heat the gas. A 10 to 12 cm cell can contain sufficient hydrogen to absorb all the Lyman alpha. Since the geocoronal hydrogen is cold it is being absorbed preferentially by the moderate temperature gas in the cell it is feasible to produce a cell which absorbed all the Hydrogen emission emitted by zero velocity Hydrogen atoms in the geocorona. Hydrogen in the aurora has very strong Doppler shift, several angstroms and therefore the Hydrogen cell passes that throuth unattenuated.

A typical hydrogen cell filter performance is shown on Figure 3.1. As the filament is heated the absorption profile broadens as shown on the figure. This broadening is responsible for increased absorption in the gas.

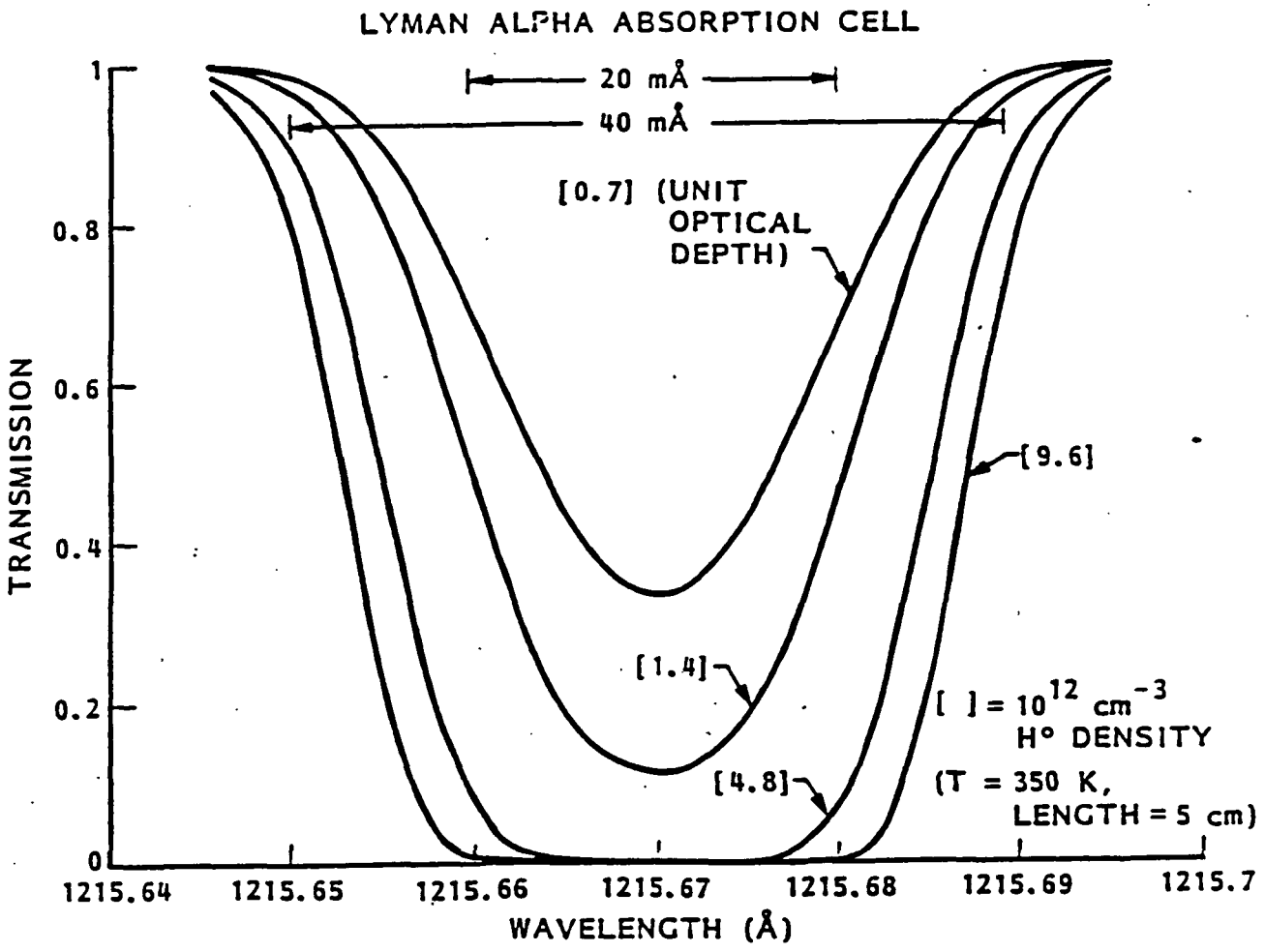


Figure 3.2.1 A typical hydrogen cell filter performance.

Starting with OGO-5 several experiments have been flown which use the Hydrogen cell for wavelength discrimination. They were all single photometer types and none of them did multiplexed imaging. On Spacelab 1 a Hydrogen cell imager was flown by Bertaux et al., 1984. This instrument had a course resoloution spectrometer in series with the absorption cell to eliminate other components such as 1304 OI etc. It also had a deuterium absorption cell to make determinations of geocoronal deuterium concentrations. From the altitude of Spacelab 1 (250 km)

the observation methodology was of course quite different from that envisaged for IMI. During the Spacelab 1 mission they were making upward measurements.

Using the Lyman alpha absorption cell Bertaux et al. [1988] found that there were three components to the observed emissions. The most intense component was the geocoronal emission which when integrated along the line of sight produced about 9-15 kR of emission. They combined the hydrogen cell aborption technique with the orbital motion of the spacecraft to make comprehensive determinations of the exospheric density and temperature in planes parallel with the orbit. The second component was the interplanetary medium emission which was between 100 and 600 Rayleighs. The third component the (350 Rayleighs) of hot emission was apparantly of magnetospheric origin. Chiu et al., [1986] associated these emissions with outflowing Hydrogen. However further analysis of the dat suggested Bertaux et al., 1988 that these emissions were the result of downward flowing ions near the cleft region and they were a type of hydrogen aurora produced by the soft protons precipitating near the cleft.

Theoretically it appears that using a hydrogen cell might be very applicable for separating the cold geocoronal Hydrogen form the hot emissions of auroral origigin. Unfortunately the required gas temperatures have to be extremely high to make a significant absorption of the Lyman alpha emissions. To illustrate the point we have reproduced a table here from Bertaux et al. [1988] in this report as Table 3.2.2

Label of H-cell level	Electrical Power (Watts)	T filament	T cell (K)	τ_c	W (mA)	Wd (km/s)	T maxwell
H24	6.41	1705	390	18.4	18.7	4.6	1279.33
H29	9.27	1800	412	40	21.4	5.28	1685.53
H36	11.59	1935	443	93	24.5	6.05	2212.99
H46	16.99	2110	485	300	29.0	7.16	3099.52
H59	19.38	2170	497	575	30.4	7.50	3400.87

Table 3.2.2. Absorption characteristics of the hydrogen cell.

This table provides technical data on the gas cell. The data is collected from the Spacelab 1 flight during which the fiulament temperature was varied and the gas cell properties were measured (Bertaux et al., 1988). The gas cell filament power is listed in column 2 and the corresponding filament temperature in coulumn 3. The cell temperature is listed on column 4. The resulting optical thickness and the width in mA is listed in coulumns 5 and 6. The measured absorption half width in velocity is presented in column 7. To illustrate our point we have added the 8th column which represents the gas temperature which is equivalent to the mean velocity in hydrogen gas given in column 7. This latter parameter says that for a rough order of magnitude calculation the gas cell is effective in removing emissions which are radiated by the cold core of the geocoronal

background. It can be seen that even filament temperatures in excess of 2000 degrees or more do not heat the gas appreciably to create absorption at various wavelength relevant to auroral velocities of the protons.

A continously glowing hot filament in the Hydrogen atmosphere poses a significant risk of failing during the 2 year mission life of the IMI satellite. It is thus concluded that in the present state of technology the benefit of the hydrogen cell is negated by the substantial risk.

3.3. Coded Aperture Spectrometer.

It was proposed that we should use the coded aperture technique to improve the spectral resolution and obtain the true Doppler profile of the Hydrogen emissions. The technique for FUV auroral imaging on a low earth orbit platform has been proposed by Mende and Fuselier [1993]. Although in principal this technique permits a large increase in the throughput or etendu our in depth study shows that we do not expect to obtain a signal to noise ratio enhancement.

In a conventional grating spectrograph consisting of a single entrance slit, a grating and a multichannel (imaging) detector, considerable light throughput advantage can be realized by replacing the single entrance slit with a mask. This can yield signal to noise ratio increase because of increased light collection over an extended area of the mask when compared to a single slit. The mask produces a spectrum on the detector which is the convolution of the mask pattern and the spectral distribution of the light source. In order to retrieve the spectrum, the spectrum has to be inverted. In special cases where emission spectra is superimposed on weak backgrounds, the signal to noise advantage is preserved through the inversion process. Thus, this technique is valuable in the observation of light sources which are produced by atomic or molecular emissions such as aurora, airglow,some interstellar emission, or laboratory spectra. However in the case of the Lyman alpha radiation where there is a strong emission at zero Doppler shift (121.6 nm) and relatively low emission at the high Doppler shift region (121.8 and up) the mixing of the two emissions are detrimental because the fluctuations from the high intensity part of the spectrum miox with the low level signals. In this chapter of the report we will demonstrate this in general terms and then summarize the results how it specifically applies to the Lyman alpha case.

Principles of coded aperture spectroscopy

Improvements to obtain better spectral resolution or increased light gathering power prompted the design of new types of spectrographs over the years. A conventional spectrograph consists of an entrance slit, a dispersive element such as a grating, and a means of detecting the resultant spectrum. In early instruments, photographic techniques were used to record all parts of a dispersed spectrum simultaneously. As a result of advances in electronic detection, especially the introduction of photomultipliers, photographic spectrographs were replaced by spectrometers using

single channel photomultiplier detectors. The higher sensitivity and linearity of the photomultiplier was regarded to be a more significant advantage than the multiplex advantage of the photographic technique. More recently, however, with the availability of multiplexed imaging photoelectric detectors, spectrographs are becoming instruments of preference. These new instruments permit simultaneous detection of many spectral features with the efficiency equivalent to many photomultipliers. In this report we discus further improvements in the efficiency of these multiplexed spectrometers by incorporating coded apertures to replace the input slits.

In spectrometers observing large area diffuse sources the spectral resolution is limited the spectral resolution and the size of the grating as discussed in chapter 3.1. In the dispersion direction the etendu limit be expressed as:

$$D\, d\alpha = D\, d\lambda x(\lambda^2 + l^2)^{-1/2} \sim D\ d\lambda/\lambda.$$

$d\lambda$ is directly related to the slit width dw as

$$d\lambda = \lambda \cot\theta\, dw / f$$

where θ = angle of dispersion produced by grating, f = focal length. Substituting into the equation above:

$$D\, d\alpha = D\ d\lambda/\lambda. = D\ \cot\theta\, dw / f$$

For a given grating this quantity is proportional to D dw/f = dw/F which is the slit width divided the F number of the optics which is focusing the light at the slit.

Thus we have shown that the light throughput (etendu or $A\Omega$) for a given spectral resolution is limited by the width of the slit. Since higher spectral resolutions demand narrower slits high spectral resolution spectrometers have small etendus for the same grating size. For faint sources or high throughput, wide slits are desirable but the width of the slit cannot be increased without reducing spectral resolution. In this report, we discuss the replacement of the single entrance slit by a mask consisting of a coded sequence of multiple slits that provide a signal-to-noise advantage related to the increase in the area. Such coded aperture spectrometers have been used in the past in conjuction with single channel detectors. We are applying the coded aperture mask technique to a spectrograph with multiplexed area detectors.

As we have stated, the efficiency of spectrometers employing a single channel photoelectric detector can be increased by using multiple entrance or exit slits in the form of a coded mask. Such multi-slits or masks were first constructed by Golay [1949, 1951]. Subsequently, grating instruments using various types of masks were developed by Girard [1960, 1963a, 1963b], and Mertz et al.[1963, 1965]. An extensive description of the technique of using a special type of masks so called Hadamard type masks for multiplexing is given by Harwit and Sloane [1979]. Presumably because their book was written prior to the widespread application of photoelectric

imaging detectors, the spectrometric and imaging techniques they discuss are related to systems using single channel detectors. In this application, a Hadamard mask is a pattern of opaque and transmissive slits of various widths represented by binary code (opaque = 0, transparent = 1) which, when cycled through all its cyclic permutations, produces a complete orthogonal set of patterns. The orthogonality property allows very simple mathematical inversion of the spectrum.

The binary codes needed to describe such a mask exhibiting the orthogonal property are contained in the rows of a matrix called an S matrix. S matrices are used when there is only one detector and the mask is a sequence of opaque and transparent elements. The so called Hadamard matrix would be used in a two detector case where the opaque areas were reflective and the reflected light were to be also collected into a second detector. For the practical simplicity of a single detector, the S-matrices are most suitable for these applications. More specifically, we are using the cyclic variety of the S matrices in which consecutive rows are generated by cycling any row through all its cyclic permutations. There are several known methods for constructing cyclic S matrices [1979].

Harwit and Sloane [1979] classify Hadamard spectrometers into two categories, the singly encoded and the doubly encoded type. In a singly encoded type, the exit slit is replaced by a cycled Hadamard mask. The doubly encoded type has two variations, non-imaging and imaging. In the non-imaging type, both the entrance and the exit slit is replaced by a Hadamard mask and one of the masks is stepped through by mechanically moving it. In the imaging type, both masks are stepped through and data is taken for a complete set of steps for each position of the other mask. The cycling of both masks provides enough information to solve for a spectrum for every slit position of the entrance mask. To obtain data for any of the spectrometer types each cycled mask must be moved mechanically or switched electro-optically through a complete orthogonal set of mask patterns.

In this chapter we will examine the performance of a singly encoded spectrograph using a fixed S matrix mask in place of the entrance slit and employing an imaging type detector on the output as illustrated on Fig. 3.3.1. In particular, we will discuss the advantages in signal-to-noise ratio improvement of this type of coded aperture spectrometer relative to a single slit design in the observation of diffuse sources. We will first show that the coded aperture design provides some improvement even when the imaging detector is an ideal detector limited only by statistical fluctuations in the input signal. This improvement applies only when the system is used in the measurement of bright emission lines embedded on a relatively weak continuum. More dramatic improvements can be obtained in systems which are limited by detector noise. However it is assumed that the detectors in the FUV spectral range near Lyman alpha would use photo electric devices which are limited by signal induced quantum noise.

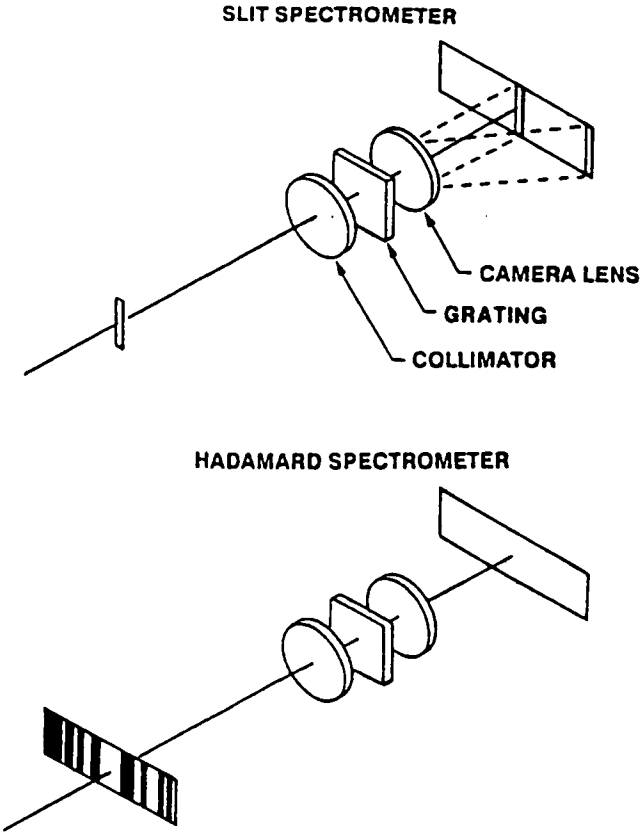


Fig. 3.3.1. Schematic drawing of a single slit spectrograph (top) and a Hadamard spectrograph (bottom). In the single slit spectrograph the first order spectral image of the slit is shown. The Hadamard spectrometer has a wide coded entrance aperture consisting of several slits. Each transparent slit produces a single slit spectrum, but offset in position.

Description of a Hadamard encoded coded aperture spectrograph.

The coded aperture spectrograph system, shown on the bottom of Fig. 3.3.1, produces an image on the detector. The image is a convolution of the coded mask pattern and the spectral profile of the light which is uniformly illuminating the coded mask. In general, the sequence of the clear (1) and opaque (0) slits in the mask could be chosen arbitrarily. In Fig. 3.3.2, we are illustrating the operation of a simple coded aperture spectrograph consisting of a seven element mask. We have chosen a coding pattern of 1110100, a seven element S matrix sequence [Harwit and Sloane, 1979]. The input light has a spectral distribution described by seven spectral intensities X_1, X_2 and X_7. The determination of the value of each of these intensities is the task of the spectrograph. Light from the mask elements is displaced by the spectrometer linearly with wavelength and the combined image is focused on the detector by the spectrograph. The superimposition scheme is illustrated schematically on Fig. 3.3.2. Spectral intensity X_1 is diffracted most to the left. This will deposit light proportional to X_1 to the leftmost seven detector elements from any open position in the mask. Spectral intensity X_2 will contribute to detector

elements 2 to 8 in a similar fashion. Thus, in general, light will appear at all 13 detector elements which will be a mixture of the different spectral intensities. Note that in the special case of the chosen seven element mask, the actual arrangement chosen for our illustration detector element 12 and 13 actually receive no light at all. However, in general, the spectrograph hardware transfers light into 13 detector elements. From hereon data can be handled by computer.

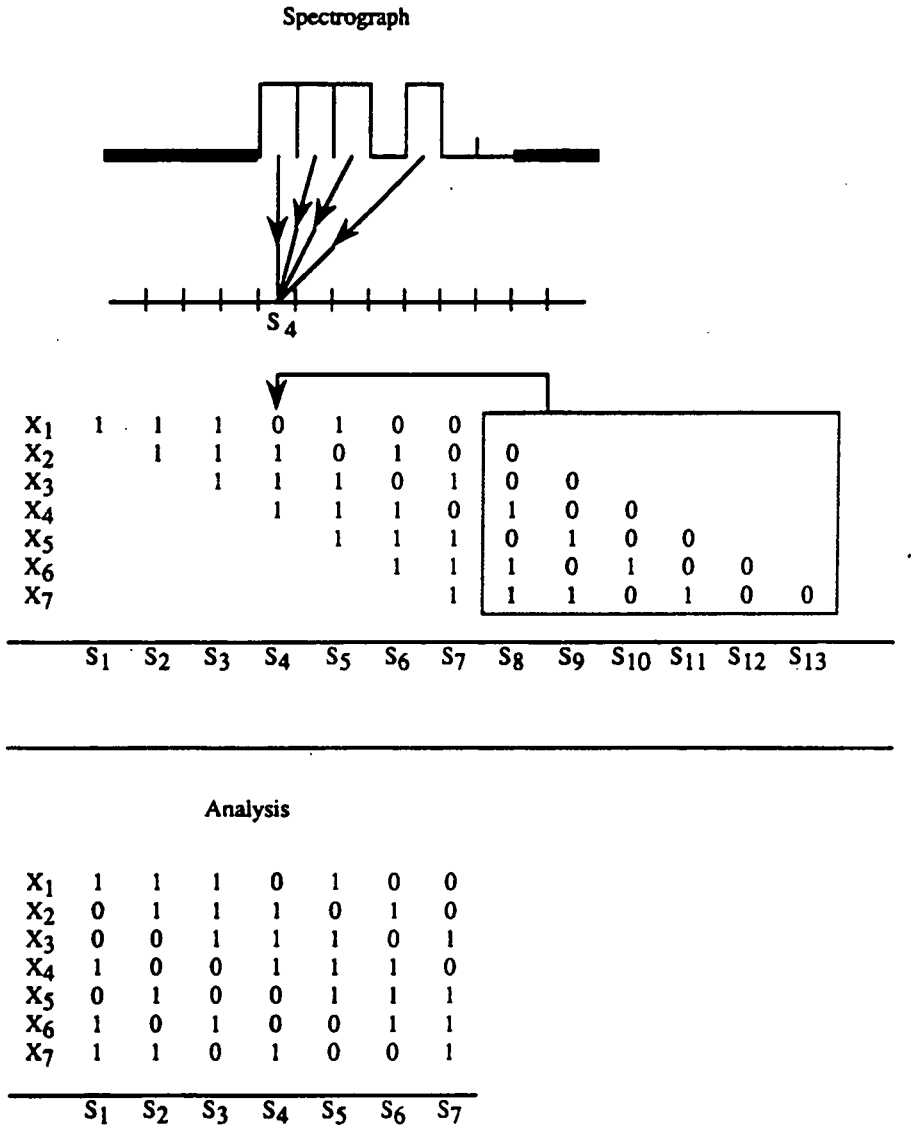


Fig. 3.3.2 Illustration of a seven element coded aperture spectrograph. Contributions to detector element 5, for example are indicated by arrows representing spectral intensities X1, X3, X4, and X5 each diffracted by a different amount. The array shows how each spectral intensity contributs to the detector elements according to the mask pattern.

Our choice, the S matrix coding of the aperture, allows us to simplify the analysis greatly. By adding the contents of column 8 to column 2, column 9 to column 3.... and column 13 to 7, we can reduce the number of (columns) to 7 and produce the following matrix:

The analysis simply consists of summing up all the data from the detector elements S_1 to S_7 while assigning the appropriate sign for each term according to the elements in the i-th row namely using a + for a 1 and a - for a 0. The resultant needs to be divided by (N+1)/2. Note there are (N+1)/2 clear apertures in an S matrix of order N.

As an illustration of the orthogonal property, we can follow how the third term X_3 propagates through the analysis. The X_3 term appears in the third row (0 0 1 1 1 0 1). In the analysis, when it is desired to obtain the value of X_3, sums are produced according to the signs associated with the third row (- - + + + - +). Applying it to the terms containing X_3, we simply get $4X_3$. In the seven element matrix N=7 and (N+1)/2=4. However, for any other row containing another unknown e.g. the row containing X_4 (1 0 0 1 1 1 0), when summed accordingly, the third row rule produces $-X_4 - 0 + 0 + X_4 + X_4 - X_4 + 0 = 0$. Similarly all terms not containing X_3 disappear when summed according to the third row rule.

If another coding scheme is used which does not have the orthogonal property, the solution could be more complex. Knowing the mask patterns and the spectral displacements, it is possible to construct equations for producing the sums at each detector element in terms of Xs, the unknown spectral intensity terms. A set of equations can be solved for the these terms and a spectral profile derived. Besides the convenience of the orthogonality, the S matrix coding scheme is also advantageous from the signal to noise stand point.

The analysis can be stated more formally by introducing the array a_{ij} for the S matrix components for the ith row and jth column and the spectral intensity I_i, the S_j sum in each detector element is:

$$S_j = \sum_{i=1}^{n} a_{ij} I_i \qquad (1)$$

The analysis simply consisting of obtaining X_i according to:

$$X_i = \frac{2}{n+1} \sum_{j=1}^{n} 2(a_{ij} - .5) S_j \qquad (2)$$

If there is no noise in the system $I_i = X_i$.

Thus, by recognizing that the signal in the detector elements is equivalent to the column sum of the matrix, the individual spectral intensities can be recovered from equation (2).

Signal to noise ratio propagation through a spectrograph using S matrix coded apertures.

It is most important to examine whether the increased etendu of the coded aperture results in any real increase in signal to noise ratio. Assume that the signal in each detector S_j is modified by some added uncertainty e_i at every wavelength. From equation (1) it can be seen that the sum (S_j) now is:

$$S_j = \sum_{i=1}^{n} a_{ij} I_i + e_j \tag{3}$$

If we proceed to evaluate X_i we obtain :

$$\begin{aligned} X_i &= I_i + \frac{2}{n+1}\left(\sum_{j=1}^{n} e_j\right) \\ \text{or} \\ X_i - I_i &+ \frac{2}{n+1}\left(\sum_{j=1}^{n} e_i\right) \end{aligned} \tag{4}$$

We do not know the values of e_j. However if we repeat the experiment many times we can obtain a statistical relationship between the root mean square error.

$$E = \sqrt{\overline{X_i - I_i}} = \frac{2}{n+1}\sqrt{\left(\sum_{i=1}^{n} \overline{e_i^2}\right)} \tag{5}$$

where $\overline{e_i^2}$ is the mean square signal fluctuation in the ith detector element.

If we assume that the noise is quantum noise of the input signal. The rms noise due to Poissonian fluctuations is

$$E_q = \frac{2}{n+1}\sqrt[7]{\sum_{i=1}^{n} I_i} \tag{6}$$

In case (1) of a continuum spectra, every value from I_1 to I_n is the same, equal to I_c, the mean squared error $\overline{e_i^2} = \frac{n+1}{2} I_c$ because in each detector element the signal is added from every open aperture. There are $\frac{n+1}{2}$ open apertures. In this case we can replace $\overline{e_i^2}$ by $I_c \times \frac{n+1}{2}$.

$$E_q = \frac{2}{n+1}\sqrt{\frac{n+1}{2}\sum_{i=1}^{n} I_c} = \sqrt{\left(\frac{2n}{n+1}\right)} I_c \tag{7}$$

Note that for a single slit spectrometer the noise is $\sqrt{I_c}$. Hence, when the input spectra is a continuum using the coded aperture technique, the signal to noise actually worsens by approximately a factor of $\sqrt{2}$.

If the spectrum consists of a single line I_s (case 2) then

$$E_q = \frac{2}{n+1}\sqrt{\sum_{i=1}^{n} I_s} = \frac{2}{n+1}\sqrt{nI_s} \tag{8}$$

This is approximately equal to $\frac{2}{\sqrt{n}}\sqrt{I_s}$. Since $\sqrt{I_s}$ is the noise for a single slit, the coded aperture technique produces a noise reduction by a factor of $\frac{2}{\sqrt{n}}$.

Although it is unlikely to be the case in the hydrogen detection at Lyman alpha we can consider detector background noise limited situation in which the detector noise is independent of the signal $\overline{e_i^2} = d_c^2$. In this case the Hadamard coded aperture spectrograph can make substantial improvements in signal to noise ratio.

$$E_{dc} = \frac{2}{n+1}\sqrt{\sum_{i=1}^{n} dc^2} = 2\frac{\sqrt{n}}{n+1} dc \qquad (9)$$

Thus the detector background noise in the coded aperture case is reduced by a factor of $2\sqrt{n}/(n+1)$.

The difficulty in using the coded aperture imager for Lyman alpha imaging can be demonstrated by looking at equation (6) above. This equation says that the fluctuations/noise, E_q, in any of the wavelength channels is a function of the sum of the illumination in all the channels In the case of the Lyman alpha detection there is a huge signal (several kRayleighs) in the zero Doppler shift channels while the other channels would only have a few hundred Rayleighs of signal. The signal to noise ratio in the single slit spectrometer case is determined from the fluctuations associated with the few hundred Rayleigh signal. However in the coded aperture case the signal to noise ratio would be determined by comparing the few hundred Rayleigh signal to the fluctuations generated by the entire band which includes the large component at zero Doppler shift.

4.0 Conclusions and Recommendations for Future Work

We have reviewed several technologies for making FUV images on the IMI mission. Our motivation is to design an instrument which would be the optimum combination for satisfying all the IMI FUV science goals. The main science goals of an FUV imager on the IMI satellite are as follows:

1. Determine the intensity of the geocoronal Lyman alpha
2. Measure the Size and extent of (electron) auroral precipitation regions.
3. Determine the morphology and intensity of global proton precipitation.
4. Determine the precipitating energy distribution over the polar cap.
5. Separate regions of soft and hard electron precipitaion.

This set of requirements define design parameters for the combined instrument.

The instrument has to have very high photon collection efficiency. IMI is in a high altitude orbit at 7-10 Re apogee while trying to resolve relatively small scale auroral behaviour on the ground. The exposue time is limited due to spinning of the spacecraft. The instrument has to reject out of band emissions to permit the quantitative interpretation of the auroral parameters. For example the instrument needs to reject the wavelength region higher than 200 nm to eliminate the huge solar

contribution from the daytime atmosphere below the aurora. In order to distuinguish between the quantitatively interpretable 135.6 nm line and the multiply scattered 1304 emission it is highly desirable that the instrument has a wavelength resolution of at least 3 nm. It is highly desirable that the instrument has a spectral resolution of aboiut 0.2 nm for making Hydrogen Doppler measurements to determine the energy of the proton precipitation. In addition it would be extremely useful to make measurements in the low and high wavelength band so that the electron energy could be derived frm the comparison of the two regions.

Baseline instrument parameters .

Parameter	Recommended requirement
Sensitivity	100-200 R
Rejection of out of band visible light	1 E-09
Spectral selectivity	3nm bandwidth
Spectral resolution for proton Doppler	0.2 nm
Simultaneous observations:	Simultaneous wide band imaging at 140-160 and 165- 185nmI

Combined Instrument recommendation

To satisfy these design requirement it appears that a spectrometer is required to provide the high resolution .2nm Doppler profile of the Hydrogen precipitation. There are no filtering techniques which could satisfy this requiremnent.

To have the maximum possible etendu from a rotating platform a wider field of view instrument is advantageous. An imager which has a wide field of view can keep the object in view during the time the object is rotating through the field of view. This property of the imager provides an advantage in light collection capability. It can be shown that for optimum performance the field of view of the imager should not be any larger than it is absolutely necessary from the point of view of tcovering the region. In summary we recommend that the FUV instrument should consist of a spectrometer with a wavelength resolutiuon of .2 nm combined with an imageer with a modest field of view and minimal spectral resolution.

In the above report we have modelled the spatial morphology of hydrogen auroras as they would be seen from above by the IMI satellite. Our model was derived from the mean of many measurmenmts taken by the DMSP satelllites. It can be seen that the proton precipitation patterns provide a ready means of observing the status of the magnetosphere.

We have also modelled the spectral profile of the Lyman alpha emissions from the aurora as it is seen superimpoosed on the geocorona. This modelling allowed the estimation of the spectral dispersions needed.

In this report we have also examined the idea of imaging through a hydrogen absorption cell. The technique of removing the geocoronal lyman alpha for auroral obserbvations by a hydrogen gas cell was used in several prior flight experiment. It was found from published technical data that the gas cell is effective in removing some of the emission which are radiated by the cold core of the geocoronal background. The geocoronal background has an intense cold core. It can be seen that even filament temperatures in excess of 2000 degrees or more do not heat the gas appreciably to create absorption at various wavelength relevant to auroral proton velocities.

A continously glowing hot filament in the Hydrogen atmosphere poses a significant risk of failing during the 2 year mission life of the IMI satellite. It is thus concluded that in the present state of technology the benefit of the hydrogen cell is negated by the substantial risk.

It was suggested that a coded aperture imager spectrometer should propduce an improvement in etendu and therefore such an instrument should be used on the IMI mission to image in the FUV. This idea was examined and it was demonstrated that from high altitudes e.g. from 6 Re or above the hydrogen geocorona produces a very strong background and the coded aperture scheme would not work well. An equation was derived to describe the performance of the coded aperture spectrometer signal to noise ratio and it was shown that in a coded aperture (Hadamard) spectrometer the fluctuations/noise, E_q, in any of the wavelength channels is a function of the sum of the illumination in all the channels In the case of the Lyman alpha detection there is a huge signal (several kRayleighs) in the zero Doppler shift channels while the other channels would only have a few hundred Rayleighs of signal from proton aurora. The signal to noise ratio in a single slit spectrometer is calculated from the fluctuations associated with the few hundred Rayleigh signal. However in the coded aperture spectrometer the signal to noise ratio is calcualated by dividing the few hundred Rayleigh signal by the large fluctuations of the average signal in the entire band which includes the large geocoronal component at zero Doppler shift. The nett result is therefore a signal to noise ratio loss rather than gain when comparede to the conventional single slit spectrometer.

In conclusion it is recommended that further work should be performed in the design of the combined geocoronal and FUV auroral spectrometer. Such work could include vacuum ultraviolet component testing and the construction of a model spectrometer.

5. References

Anger, C. D., J. S. Murphree, A. Vallance Jones, R. A. King, A. L. Broadfoot, L. L. Cogger, F. Creutzberg, R. L. Gattinger, G. Gustafsson, F. R. Harris, J. W. Haslett, E. J. Llewellyn, J. C. McConnell, D. J. McEwen, E. H. Richardson, G. Rostoker, B. R. Sandel, G. G. Shepherd, D. Venkatesan, D. D. Wallis and G. Witt, "Scientific results from the Viking ultraviolet imager: An introduction," *Geophys. Res. Lett.*, 14, 383-386, 1987.

Basu, B, J. R. Jasperse, R. M. Robinson, R. R. Vondrak, and D. S. Evans, Linear transport theory of auroral proton precipitation: A comparison with observations, *J. Geophys. Res., 92*, 5920, 1087.

Bertaux, J. L., F. Goutail, G. Kockarts, Observations of Lyman alpha emissions of Hydrogen and deuterium. Science, 225, 174-176, 1984

Bertaux, J. L., H. Le Texier, F. Goputail, R. Lallement, G. Kockartts, Lyman alpha observations of geocoronal ane interplanetary hydrogen from spacelab-1 : exospheric temperature and density

Blamont, J. E., and A. Vidal-Madjar, Monitoring of the Lyman alpha emission line of the sun during the year 1969, *J. Geophys. Res., 76,* 4311, 1971.

Chamberlain, J. W., Planetary coronae and atmospheric evaporation, *Planet. and Space Sci., 11,* 901-960, 1963.

Chakrabarti, S. C., Extreme and Far Ultraviolet Emissions From the Polar Cap, *J. Geophys. Res. 91*, 8065–8072, 1986.

Chakrabarti, S., F. Paresce, S. Bowyer, R. Kimble, and S. Kumar, The extreme ultraviolet day airglow, *J. Geophys. Res.*, 88, 4898-4904, 1983.

Creutzberg, F., Gattinger, R., Harris, F., Wozniak, S., and Valance Jones, A., Auroral studies with a chain of meridian scanning photometers 2. Mean distribution of proton and electron aurora as a function of magnetic activity, *J. Geophys. Res.*, 93, 14591-14601, 1988.

Eather, R. H., Auroral proton precipitation and hydrogen emissions, *Rev. Geophys. and Space Phys.*, 5, 207-285, 1967.

Edgar, B. C., W. T. Miles, and A. E. S. Green, "Energy deposition of protons in molecular nitrogen and application to proton auroral phenomena," *J. Geophys. Res.*, 78, 6595-6606, 1973.

Feldman, P. D., and E. P. Gentieu, The ultraviolet spectrum of an aurora 530-1520 Å, *J. Geophys. Res., 87,* 2453-2458, 1982.

Frank, L. A., and J. D. Craven, "Imaging results from dynamics explorer 1," *Reviews of Geophysics*, 26, 249-283, 1988.

Gattinger, R. L., private communication.

Girard, A., Nouveaux Dispositifs de Spectroscopie á Grande Luminositè, *Optica Acta 7,* 81–97 1960.

Girard, A., Spectromètre á Grilles, *Appl. Opt. 2*, 79–87, 1963.

Girard, A., Spectromètre á Grilles, *J. de Phys. 24,* 139–141 1963.

Golay, M. J. E., Multi-Slit Spectrometry, *J. Opt. Soc. Am. 39,* 437-444, 1949.

Golay, M. J. E., Static Multislit Spectrometry and Its Application to the Panoramic Display of Infrared Spectra, *J. Opt. Soc. Am. 41*, 468-472, 1951.

Hardy, D. A., M. S. Gussenhoven, and D. Brautigam, A statistical model of auroral ion precipitation, *J. Geophys. Res.*, 94, 370-392, 1989.

Harwit, M. and N. J. Sloane, *Hadamard Transform Optics*, Academic Press, New York, San Francisco, London, 1979.

Ishimoto, I., Meng, C. I., Romick, G. R., and Huffman, R. E., Doppler shift of auroral Lyα observed from a satellite, *Geophys. Res. Lett.*, 16, 143-146, 1989.

Lean, J. L., and A. Skumanich, Variability of the Lyman alpha flux with solar activity, *J. Geophys. Res., 88,* 5751, 1983.

Meier, R. R., R. R. Conway, P. D. Feldman, D. J. Strickland, and E. P. Gentieu, Analysis of nitrogen and oxygen far ultraviolet auroral emissions, *J. Geophys. Res., 87,* 2444-2452, 1982.

Meinel, A. B., Doppler shifted auroral hydrogen emission, *Astrophys. J., 113*, 50-54, 1951.

Mende, S. B.and R. H. Eather, "Photometric auroral particle measurements," in *Magnetospheric Processes*, B. M. McCormac, Ed. [D. Reidel Publ. Co., Dordrecht-Holland], 1972.

Mende, S. B., Claflin, E. S., Rairden, R. L., and G. R. Swenson, "Hadamard Spectroscopy with a Two-Dimensional Detecting Array", *Applied Optics*, in press 1993.

Mertz, L., N. O. Young, and J. Armitage, Mock Interferometry, pp. 51–57 of *Proc. Conference on Optical Instruments and Techniques*, K. J. Habell, ed. Wiley, New York, 1963.

Mertz, L., *Transformations in Optics* Wiley, New York, 1965.

Paresce, F., S. Chakrabarti, S. Bowyer, R. Kimble, The extreme ultraviolet spectrum of dayside and nightside aurorae: 800-1400A, *J. Geophys. Res.*, 88, 4905-4910, 1983.

Rairden, R. L., L. A. Frank, and J. D. Craven, Geocoronal imaging with dynamics explorer, *J. Geophys. Res., 91*, 13,613-13,630, 1986.

Rees, M. H. and Luckey, D., "Auroral electron energy derived from ratio of spectroscopic emissions. 1. Model computations," *J. Geophys. Res.*, 79, 5181-5186, 1974.

Strickland, D. J., Jasperse, J. R. and Whalen, J. A., Dependence of Auroral FUV emissions on the incident electron spectrum and neutral atmosphere, *J. Geophys. Res.*, 88, 8051-8062, 1983.

Strickland, D. J., R. E. Danielle, Jr., J. R. Jasperse, and B. Basu, Transport-theoretic model for the electron-proton-hydrogen atom aurora 2. Model results, *J. Geophys. Res., 98*, 21,533-21,548, 1993.

Thomas, G. E., and R. F. Krassa, OGO 5 measurements of the Lyman-alpha sky background in 1970 and 1971, *Astron. Astrophys., 30,* 223-232, 1974.

Vallance Jones, A., Overview of auroral spectroscopy, in Auroral Physics, edited by C.-I. Meng, M. J. Rycroft, and L. A. Frank, Cambridge UP, p. 15-28, 1991.

Van Zyl, B., M. W. Gealy, and H. Neumann, N_2^+ first negative emission cross sections for low energy H^+ and H impact on N_2, *Phys. Rev. A, 28*, 2141, 1983.

Van Zyl, B., M. W. Gealy, and H. Neumann, Prediction of photon yields for proton aurorae in an N_2 atmosphere, *J. Geophys. Res., 89*, 1701, 1984.

Van Zyl, B. and H. Neumann, Lyman a emission cross sections for low energy H and H^+ collisions with N_2 and O_2, *J. Geophys. Res., 93*, 1023, 1988.

Zukic, M., and D. G. Torr, Multiple Reflectors as Narrow-band and Broadband Vacuum Ultraviolet (VUV) Filters, *Applied Optics*, 31, 1588 1992.

6. List of Figures