Proceedings of 33rd Annual European Meeting on Atmospheric Studies by Optical Methods.
IRF Sci. Rep., 292, 105–109, 2008
http://www.irf.se/publications/proc33AM/aso-etal.pdf
ISSN 0284-1703, ISBN 978-91-977255-1-4, © Author(s) 2008

A proposed Bayesian model on the generalized tomographic inversion of aurora using multi-instrument data

Takehiko Aso[1], Björn Gustavsson[2], Kunio Tanabe[3], Urban Brändström[4], Tima Sergienko[4], and Ingrid Sandahl[1,4]

[1]National Institute of Polar Research
[2]Department of Physics, University of Tromsø
[3]Faculty of Science and Engineering, Waseda University
[4]Swedish Institute of Space Physics, Kiruna

Abstract.

The aim of aurora tomography is to reconstruct the 3D luminosity distribution of aurora from multiple monochromatic images taken with a multi view-point observing network. As a logical extension of this, we propose a generalized tomographic inversion of aurora in which different simultaneous signatures of particle precipitation are invoked in the calculation in order to retrieve more comprehensive information about the initial differential energy spectrum of precipitating particles at the top of the atmosphere. The method combines information from the luminosity distribution of aurora obtained by optical imagers, electron density data from the EISCAT radar, and cosmic noise absorption data from imaging riometers. All these data are integrated together in a statistical sense to infer the energy spectrum of the primary electrons causing the aurora. In this scheme, minimization or optimization in the inverse problem is converted to establishing a Bayesian model that gives the most probable function model, as well as the relative significance of each type of input data used in the calculation.

Keywords : aurora tomography, inverse problem, Bayes model, precipitating particle energy, EISCAT radar, imaging riometer

1 Introduction

We have been working for more than ten years on tomographic inversion of aurora in Antarctica, in Iceland (Aso et al., 1990, 1993), and since the start of the Swedish ALIS (Auroral Large Imaging System) in the Kiruna region (Aso et al., 1998; Gustavsson et al., 2001a; Brändström et al., 2003). In this analysis, both an analytical method and algebraic reconstruction techniques (ART) such as MART (Multiplicative ART) or SIRT (Simultaneous Iterative ART) are employed on up to six images viewing the aurora in the same volume from different viewing positions together with some constraints for this under-determined and ill-posed problem. To stabilize the solution, we use the a priori constraints that the aurora is excited by electrons traveling along the magnetic field lines and that the altitude profile of luminosity does not change abruptly. In August 2005, the Japanese Reimei satellite was launched and is taking monochromatic images of aurora from above, i.e. topside and lateral side of aurora. This makes it possible to use less constrained and time-dependent reconstruction of the auroral volume emission rate. This combination of ground based and in situ imaging is an advance in auroral tomography, but it is difficult to get observations in satisfactory conjunction. To go beyond the conventional tomography, it is here proposed to utilize in the solution of the inverse problem all multimodal data of parameters that result from precipitating aurora particles. These data are the simultaneously enhanced electron density observed by the EISCAT radar and the increased cosmic noise absorption detected by imaging riometer. It is even possible to include observations of the total electron content (TEC) which is an integration of electron density along the propagation path of a radio wave.

Inversion methods to retrieve the electron energy spectrum solely from ion production or electron density observed by the EISCAT radar have been developed by many workers, e.g. the CARD method by Brekke et al. (1989), SPECTRUM by Kirkwood (1988), and a time-dependent inversion by Semeter and Kamalabadi (2005). For TEC data, Raymund et al. (1990) compared the reconstruction of electron density distribution with incoherent scatter radar measurements. Imaging riometers give information about the harder part of the precipitating particle spectrum, and inversion analyses based on imaging riometer have been carried out by Kosch et al. (2001), Ashrafi et al. (2005).

For the ionization and excitation of atmospheric constituents by precipitating electrons, which is a basic “for-

Correspondence to: T. Aso (t-aso@nipr.ac.jp)

Published by the Swedish Institute of Space Physics.

ward" problem, many theoretical formulations have been developed. According to Rees (1963) (also in Rees, 1989), the ionization rate $\bar{q}(z)$ for the initial differential energy flux spectrum at the top of the atmosphere $f_0(E_0)$ can be approximately expressed by an integral over the relevant energy range as

$$\begin{aligned}\bar{q}(z) &= \frac{N(z)}{\Delta\epsilon_{ion}} \int \frac{\rho(R) f_0(E_0) E_o \lambda(E_0, \chi)}{R(E_0) N(R)} dE_0 \\ &= A(z) \int f_0(E_0) E_o Q(E_o) dE_0\end{aligned} \tag{1}$$

where $z = p/g$ is the scale height, ρ the atmospheric density, λ the normalized energy dissipation distribution function, χ the atmospheric depth or distance from the top, E_0 the initial electron energy, $\Delta\epsilon_{ion}$the ionization energy cost, N the number density of ionizable constituents, R the range or penetration depth for the particle with energy E_0, and $A(z)$ and $Q(E_0)$ collect the corresponding terms together. This simplified ionization rate is based on the empirical curves for energy dissipation distribution and effective range which give energy deposition as a function of fractional range. These together with scattering depth and neutral density yield an altitude-dependent ionization rate for electron beams, monoenergetic and others.

In a similar way, the volume emission rate or aurora luminosity profile $\Lambda(z)$ for $f_0(E_0)$ can likewise be expressed (Sergienko and Ivanov, 1993; Rees, 1963, 1989) as

$$\begin{aligned}\Lambda_\lambda(z) &= \frac{A_\lambda}{\sum A_\lambda} \frac{\rho(z) P_i(z)}{\epsilon^i_j(z)} \int \frac{f_0(E_0) E_o \lambda(E_0, \chi)}{\overline{R}(E_0)} dE_0 \\ &= B(z) \int f_0(E_0) E_o L(E_o) dE_0\end{aligned} \tag{2}$$

where $A_\lambda$is the Einstein emission coefficient, ϵ^i_j the excitation energy cost of the transition j for species i, P_i the excitation probability for species i, $\bar{R}$ the average range for E_0, and $B(z)$ and $L(E_0)$ collects the corresponding terms together. This is a simplified emission rate for $N_2{}^+$ 1NG 427.8 nm from a given altitude-dependent N_2 ionization rate for mono-energetic electrons. Incorporation of images with different wavelength in the present tomographic approach should be targeted for further study since ratios of emission intensity between particular wavelengths bear information on the characteristic energy of the precipitating particles and accordingly the altitude profile of the auroral emission.

As an inverse problem, Gustavsson et al. (2001b) expressed the above integral over energy as a simple "mixed determined" system of linear equations specified by the transfer matrix T as

$$\Lambda_\lambda(z_i) = \sum_i T_{i,j}\, f_0(E_j) \tag{1}$$

and estimated primary electron spectra from the altitude distribution that was derived from two station images with a thin sheet assumption (Gustavsson et al., 2001b). Subsequently, the electron distribution function is calculated by inverting T as

$$f_0(E) = T^{-1} \Lambda_{4278}(z) \tag{2}$$

We propose to combine these inversion approaches for excitation and ionization observations, i.e., observations of auroral luminosity and electron density enhancements, in a statistical sense to infer the energy spectrum of precipitating particles from the magnetosphere. In our scheme, minimization or optimization in this inverse problem is converted to finding a Bayes model that will give the most probable function model, as well as the relative significance of each type of input data used in the calculation.

In the following sections, we will present a formulation for comparing the excitation and ionization with observations of aurora, cosmic noise absorption, and electron densities; then propose an inversion algorithm based on Bayes principle to infer the particle spectrum of the primary auroral electrons.

2 Basic concepts and fundamental equations

An inverse problem in tomography is generally associated with the forward convolution process as

$$K\Lambda \equiv \int k(y, x) \Lambda(x) dx = \tilde{g}(y) \tag{3}$$

Here $\tilde{g}(y)$ is the observed column emission rate of image pixel at y, $\Lambda(x)$ the volume emission rate in a voxel at $x$and k is an integral kernel corresponding to the image projection from x to y. This integration inevitably implies smoothing and increase of entropy, i.e., loss of information and hence with a limited number of noisy observations the inversion becomes an ill-posed problem.

If the back projection or approximate inverse operator of K is $\bar{K}$, the i+1 th iteration for retrieving f can be expressed as

$$\Lambda^{(i+1)} = \Lambda^{(i)} + \bar{K}(\tilde{g} - K\Lambda^{(i)}) \tag{4}$$

where i is the index for iteration number. Algebraic reconstruction methods such as ART, MART and SIRT are used for this iteration. Our analysis has been mostly relying on the SIRT method for its better robustness to noise.

The other approach, the optimization model, defines an appropriate "functional" and optimizes it. A functional to be minimized is

$$|K\Lambda - \tilde{g}|^2 + \Omega(\Lambda) \tag{5}$$

where $\Omega(\Lambda)$ represents the functional term for unbiased constraints of the smoothness structure in Λ.

In the present approach of integrating available information on particle precipitation, we begin with an optimization model in which the following functional

$$pl(f_0; w_1, w_2, w_3) \equiv l(f_0;\ w_1, w_2, w_3) + \ \Omega(f_0) \quad (6)$$

is to be minimized with respect to f_0 for given hyper-parameter values w_1, w_2, w_3.

$$l(f_0;\ w_1, w_2, w_3) \equiv w_1 \sum_{i=1,\lambda}^{i=6} \sum_{u,v} \left| c_i(\theta, \lambda) \int_{l(\theta,\phi)} \frac{e^{-a_\lambda / \cos\theta}}{|\mathbf{r}(l) - \mathbf{r}_i|^2} \Lambda(\mathbf{r}(l),\ f_0) dl - \tilde{g}_i(u, v, t) \right|^2 + w_2 \left| \bar{q}(\mathbf{r},\ f_0) - \alpha \tilde{n}_e^2(\mathbf{r}, t) \right|^2 + w_3 \left| \int_{60-100km} a_\omega \frac{\nu}{\omega^2 + \nu^2} \sqrt{\frac{\bar{q}(\mathbf{r},\ f_0)}{\alpha}} ds - \tilde{b}_{64}(u, v, t) \right|^2 \quad (7)$$

Symbols are as follows. l: line of sight, $\mathbf{r}$: point in space, $\mathbf{r}_i$: observation point, a_λ: atmospheric absorptivity, $\tilde{b}_{64}$: observed radio wave absorption, c_i: sensitivity and vignetting factor, θ, φ: zenith and azimuthal angle, $\tilde{n}_e$: observed electron concentration, α: effective electron recombination rate, λ: wavelength, $w_1 - w_3$: weighting factors, a_ω: specific absorption at ω, ω: riometer angular frequency, ν: electron - neutral collision frequency. Also $\Lambda(r)$ and $\overline{q(r)}$ are determined by integrating $f_o(E_o)$ as in Eqs. (1) and (2). In the above $l(f_0;\ w_1, w_2, w_3)$ is composed of three parts: (1) An aurora tomography part with e.g. six images, (2) an electron density profile part for energy range of 0.2-50 keV, and (3) an imaging riometer part for energies greater than 10keV with e.g. 8x8 receiving beams. Of course, we should be deliberate on the differences in spatial resolution of the respective data sets and also on the horizontal inhomogeneity of precipitating particle energy. Specifically an EISCAT radar observation mode using beam scanning might be more preferable in discrete aurora which has large spatial gradients in electron precipitation, leading to sharp gradients in the ionosphere. To optimize the EISCAT observation for the inversion proposed here, antenna scanning will provide additional information about the spatial variation, widths and precipitation characteristics of such structures. For uniform diffuse aurora with less spatial gradients in the ionosphere scanning the radar beam will not give us much additional information, but might rather reduce the temporal resolution that might be more interesting for this case, e.g pulsating aurora. In the minimization, we usually add $\Omega(f_0)$ as a smoothness constraint. By putting Eqs. (1) and (2) into Eq. (7), we can solve for the electron energy spectrum $f_0(E_0)$ at the top of the ionosphere. In this case, atmospheric structures and composition can also be parameters to be inferred.

3 Conversion to a Bayesian model

Observed data are usually finite and limited and at the same time susceptible to stochastic errors. On the other hand, physical models expressed in the above integral include many assumptions and approximations. If we denote M as a model which delineates the forward process and D as observed data, the posterior probability $P(M|D)$ for model M when data D are at hand is expressed by the Bayes' theorem as

$$P(M \mid D) = \frac{P(D \mid M)P(M)}{P(D)} = \frac{P(D \mid M)P(M)}{\sum_M P(D \mid M)P(M)} \quad (8)$$

where $P(M)$and $P(D|M)$are the prior probability for model M taking place and the probability of data D for the assumed model M or the likelihood of observed data as a function of model parameters, respectively.

Relying on this, we will convert the optimization model into a Bayesian model formulation, adapted to "multimodal" stochastic data with the forward model based on our approximate understanding and knowledge of the ionosphere. Maximizing a posterior probability leads to model parameter determination with sophisticated adjustment of hyper parameters of the relative contribution of three terms, being based on data Tanabe (2004).

Now we define a Bayes model as

$$f \approx f^d(\phi) \quad (9)$$

that expresses model f by a finite number of parameters ϕ. For smoothness constraint, vanishing approximated second-order derivative might be appropriate.

$$\Omega^d(\phi) \equiv \Omega(f^d(\phi)) \quad (10)$$

Then the likelihood of data for assumed model $P(D|M)$ is defined as

$$L_\xi(\tilde{g}, \phi) \equiv \frac{\exp\left(-l(f^d(\phi))\right)}{N_l(w_1, w_2, w_3)} \quad (11)$$

in which $l$is a residual defined in Eq. (7) and N_l is a normalization factor which is an integration of the numerator by ϕ. Very little prior knowledge of parameters ϕ before getting data corresponding to $P(M)$ is likewise expressed as

$$\Pi(\phi) \equiv \frac{\exp\left(-\Omega(f^d(\phi))\right)}{N_\pi} \quad (12)$$

Here, N_π is a normalization factor. Hence the posterior distribution for parameters ϕ when data g are given or, equivalently, the probability of model parameter ϕ being realized

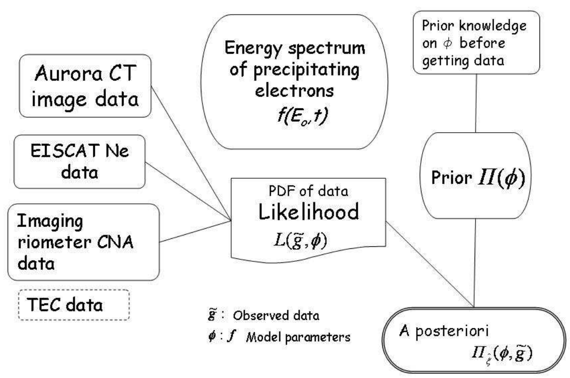

Fig. 1. Concept of generalized aurora CT

under the condition of given data g is expressed by Bayes formula Eq. (8) as

$$\Pi_{\xi}(\phi,\tilde{g}) = \frac{L_{\xi}(\tilde{g},\phi)\Pi(\phi)}{\int L_{\xi}(\tilde{g},\phi)\Pi(\phi)d\phi} \quad , \quad (\xi = (w_1, w_2, w_3)) \tag{13}$$

A denominator is a marginal likelihood which is a projection by integrating out model parameters and represents the probability of obtaining the present data $\tilde{g}$. Maximizing the above posterior distribution function with respect to ϕ, corresponds to the minimization of the optimization model.

The estimation procedure proceeds in the following way. We search for the hyper-parameter set $\xi = (w_1, w_2, w_3)$ which attains maximum of marginal likelihood $ML(\xi) = \int L_{\xi}(\tilde{g},\phi)\Pi(\phi)d\phi$ at $\hat{\xi} = (\hat{w}_1, \hat{w}_2, \hat{w}_3)$. This set determines the relative weight of the three input data sets, aurora brightness, electron density and cosmic noise absorption. In step two, we obtain an estimate $f \equiv f^d(\hat{\phi})$ by determining model parameters ϕ which maximize the posterior distribution function $\Pi_{\hat{\xi}}(\phi,\tilde{g})$ and minimizes the functional $pl(f^d(\phi); \hat{w}_1, \hat{w}_2, \hat{w}_3) \equiv l(f^d(\phi); \; \hat{w}_1, \hat{w}_2, \hat{w}_3) + \Omega(f^d(\phi))$. This completes the present algorithm. This two–step procedure is a so-called empirical Bayes method. In other words, the hyper-parameter set and $f_0(E_0)$ are searched that give the maximum marginal likelihood. The search is separated into an "inner" problem where for each set of hyper parameters we search for the optimal $f_0(E_0)$. This then makes up the optimal likelihood function for the corresponding set of hyper parameters which is the “outer” target function to maximize. The whole procedure is illustrated as a diagram in Fig. 1

The initial differential energy flux spectrum at the top of the atmosphere f_0 can be described either as

$$f_0(E_0)dE = CE_0^{\gamma}\exp(-E_0/\bar{E}_0)dE_0 \tag{14}$$

with model parameters γ and characteristic energy $\bar{E}_0$ in a possibly simple form or through node values of spline-type form. Also the terms $A(z)$ and $B(z)$ includes parameters for atmospheric structures and relevant ionization and excitation processes. The numbers of parameters depend on the available quantity of orthogonally independent data and also on a trade-off with the computing time for global search for maxima in the aforementioned procedures.

4 Discussions, suggestions and summary

A proposed model of generalized aurora tomography is given which aims at integrating information from ALIS, EISCAT and imaging riometers to estimate comprehensive differential electron energy spectra through a Bayesian statistical approach. The tentative formulation is targeting the initial energy spectrum at the top of the atmosphere through energy deposition and ionization. Janhunen (2001) did a generalization of spectroscopic ratio methods to off-zenith viewing directions to retrieve electron precipitation characteristics from a set of multiwavelength all-sky auroral images. In our approach, further comparison between the Bayesian method suggested and the conventional approach of solving the problem stepwise, i.e., first the tomographic inversion from the images to the three-dimensional volume emission followed by an estimate of the electron spectra remains. In both cases, establishing relevant forward models are really important. Also an intermediate way of integrating multi-instrument data can be suggested in which 3D aurora emission structures reconstructed from conventional aurora tomography is adjusted or modified by comparing electron density and absorption calculated by retrieved primary spectra with observations. Also, neutral composition can or should be a target for reconstruction together with energy spectrum at the top of the atmosphere.

Geophysics is in principle an inverse problem based on “observation” and then “induction” which derives theory or structures from data. Geophysics problems are frequently “under-determined” or mathematically “ill-posed”. Whereas “deduction” is a cause-&-effect argument as in e.g. computer simulation and is in principle a forward and logically well-defined approach, comprehensive induction exploiting heterogeneous data can do more convincing and versatile deconvolution. The present approach assumes some kind of steady-state condition and temporal variation is not seriously taken into account. But it will hopefully contribute to the reconstruction of auroral excitation in a statistical sense with less constraints and more flexibility and will help understand more comprehensibly aurora formation processes.

Acknowledgements. This research is partly supported by the National Institute of Polar Research as a Research Project P2 and also as a sub-theme of the research program entitled “Function and Induction research project“ at the Trans-disciplinary Research Integration Center of the Research Organization of Information and Systems.

The editors thank one referee for assistance in evaluating this paper.

References

Ashrafi, M., Kosch, M. J., and Honary, F.: Comparison of the characteristic energy of precipitating electrons derived from ground-based and DMSP satellite data, Annales Geophysicae, 23, 135–145, 2005.

Aso, T., Hashimoto, M. A., Ono, T., and Ejiri, M.: On the analysis of aurora stereo observations, J. Geomag. Geoelectr., 42, 579–595, 1990.

Aso, T., Ejiri, M., Miyaoka, H., Ono, T., Hashimoto, T., Yabu, T., and Abe, M.: Aurora stereo observations in Iceland, in: Proc. NIPR Symposium on Upper Atmos. Phys., vol. 6, pp. 1–14, National Institute of Polar Research, 1993.

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

Brändström, Urban, B., Lövgren, T., Most röm, A., Enell, C.-F., Gustavsson, B., Aso, T., Ejiri, M., Steen, Å., and Rydesäter, P.: Brief Report on ALIS (Auroral Large Imaging System), a new all -sky camera in Kiruna and Auroral Imaging using a mini-DV Camcorder, Proc. 28^{th} Annual European meeting on Atmospheric Studies by Optical Methods, pp. 92:89–92, 2003.

Brekke, A., Hall, C., and Hansen, T. L.: Auroral ionospheric conductances during dislurbed conditions, Ann. Geophys., 7, 269–280, 1989.

Gustavsson, B., Sergienko, T., Rietveld, M. T., Honary, F., Steen, Å., Brändström, B. U. E., Leyser, T. B., Aruliah, A. L., Aso, T., and Ejiri, M.: First Tomographic estimate of volume distribution of enhanced airglow emission caused by HF pumping, J. Geophys. Res., 106, 29 105–29 123, 2001a.

Gustavsson, B., Steen, Å., Sergienko, T., and Brändström, B. U. E.: Estimate of auroral electron spectra, the power of ground-based multi-station optical measurements, Phys. Chem. Earth, 26, 189–194, 2001b.

Janhunen, P.: Reconstruction of electron precipitation characteristics from a set of multiwavelength digital all-sky auroral images, J. Geophys. Res., 106, 18 505–18 516, doi:10.1029/2000JA000263, 2001.

Kirkwood, S.: SPECTRUM - a computer algorithm to derive the flux-energy spectrum of precipitating particles from EISCAT electron density Profiles,, Tech. Rep. IRF Technical Report 034, Swedish Institute of Space Physics, 1988.

Kosch, M. J., Honary, F., del Pozo, C. F., Marple, S. R., and Hagfors, T.: High-resolution maps of the characteristic energy of precipitating auroral particles, J. Geophys. Res., 106, 28 925–28 938, doi:10.1029/2001JA900107, 2001.

Raymund, T. D., Franke, S. J., Liu, C. H., Austen, J. R., and Klobuchar, J. A.: Application of computerized tomography to the investigation of ion ospheric structures, Radio Science, 25, 771–789, 1990.

Rees, M.: Auroral ionization and excitation by incident electrons, Planet. Space Sci., 11, 1209–1218, 1963.

Rees, M. H.: Physics and chemistry of the upper atmosphere, Cambridge and New York, Cambridge University Press, 1989, 297 p., 1989.

Semeter, J. and Kamalabadi, F.: Determination of primary electron spectra from incoherent scatter radar measurements of the auroral E region, Radio Sci., 40, doi:10.1029/2004RS003042, 2005.

Sergienko, T. I. and Ivanov, V. E.: A new approach to calculate the excitation of atmospheric gases by auroral electron impact, Ann. Geophys., 11, 717–727, 1993.

Tanabe, K.: Inductive inference and Bayesian method in Hierachical Bayes m odel and its environment,Supplementary chapter (in Japanese), chap. Inductive inference and Bayesian method, Iwanami Pub. Co., 2004.