Validation of Inner Magnetosphere Particle Transport and Acceleration Model (IMPTAM) with long-term GOES MAGED measurements of keV electron fluxes at geostationary orbit

N. Yu. Ganushkina[1,2], I. Sillanpää[2], D. Welling[1,3], J. Haiducek[1], M. Liemohn[1], S. Dubyagin[2], J. V. Rodriguez[4,5]

[1]University of Michigan, Ann Arbor, Michigan, USA.
[2]Finnish Meteorological Institute, Earth Observations, Helsinki, Finland.
[3]University of Texas at Arlington, Arlington, TX, USA
[4]Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, Colorado, USA.
[5]National Centers for Environmental Information, National Oceanic and Atmospheric Administration, Boulder, Colorado, USA.

Key Points:

- IMPTAM performs well, with the ratio between the GOES MAGED and modelled keV electron fluxes at 06 MLT close to one.
- Peaks of IMPTAM fluxes are shifted towards midnight due to the background field models and the sources and losses used inside IMPTAM
- Error is a factor of two based on median symmetric accuracy with largest difference of one order of magnitude, Heidke skill scores are low

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1029/2018SW002028

Corresponding author: N. Yu. Ganushkina, ganuna@umich.edu

Abstract

Surface charging by keV electrons can pose a serious risk for satellites. There is a need for physical models with the correct and validated dynamical behavior. 18.5 months (2013-2015) output from the continuous operation online in real time as a nowcast of the Inner Magnetosphere Particle Transport and Acceleration model (IMPTAM) is compared to the GOES 13 MAGED data for 40, 75, and 150 keV energies. The observed and modeled electron fluxes were organized by MLT and IMPTAM driving parameters, the observed IMF B_Z, B_Y, |B|, the solar wind speed V_{SW}, the dynamic pressure P_{SW}, and Kp and SYM-H indices. The peaks for modeled fluxes are shifted towards midnight but the ratio between the observed and modeled fluxes at around 06 MLT is close to one. All the statistical patterns exhibit very similar features with the largest differences of about one order of magnitude at 18-24 MLT. Based on binary event analysis, 20-78% of threshold crossings are reproduced but Heidke skill scores are low. The modeled fluxes are off by a factor of two in terms of the median symmetric accuracy. The direction of the error varies with energy: overprediction by 50% for 40 keV, overprediction by two for 75 keV, and underprediction by 18% for 150 keV. The revealed discrepancies are due to the boundary conditions developed for ions but used for electrons, absence of substorm effects, representations of electric and magnetic fields which can result in not enough adiabatic acceleration, and simple models for electron lifetimes.

1 Introduction

According to the Index of Objects Launched into Outer Space (http://www.unoosa.org/oosa/osoindex/) maintained by the United Nations Office for Outer Space Affairs (UNOOSA), there were about 1980 active satellites in orbit in April 2018. Many of them traverse the variable radiation environment in the magnetosphere. One of the primary constituents of the radiation environment is the electrons with energies ranging from 1 to tens of keVs. One obvious example of their importance is their role as the seed population, being further accelerated to MeV energies by various processes in the Earth's radiation belts (e.g., Horne et al., 2005; Y. Chen et al., 2007; Li et al., 2014; Jaynes et al., 2015; Boyd et al., 2016). At the same time, plasma sheet electron and ion distributions get altered into unstable forms, exciting various plasma waves (notably VLF chorus and EMIC waves) that can either energize or scatter relativistic particles (Kennel & Petschek, 1966; Kennel & Thorne, 1967; Green & Kivelson, 2001, 2004; Y. Chen et al., 2006; Shprits et al., 2006; Usanova et al., 2014; Foster et al., 2017). MeV electrons are one of the major sources of damaging space weather effects on space assets inside the radiation belts (see, for example, Baker et al. (2018) and references therein).

The electrons with energies of 10's of keVs do not penetrate deep into the satellite materials but stay near the surface, posing a serious risk for satellites in the form of surface charging (Garrett, 1981; Lanzerotti et al., 1998; Davis et al., 2008; Thomsen et al., 2013). The electron fluxes at these keV energies vary significantly with geomagnetic activity on the scale of minutes or even shorter. Their dynamics is determined by convective and substorm-associated electric fields in the magnetosphere (Mauk & Meng, 1983; Kerns et al., 1994; Liemohn et al., 1998; Ganushkina et al., 2013, 2014). When a satellite anomaly due to surface charging occurs, the radiation environment may be more extreme than that given by the specification models used for design (Iucci et al., 2005; Mato-Vlez et al., 2018). However, data may not be available at the location of the satellite to determine the cause of the anomaly. Thus, there is a need for physical models with the correct dynamical behavior that can be used to reconstruct the radiation environment at any location at any satellite orbit. Prediction models of MeV electron fluxes do daily averaging (Balikhin et al., 2016), even though less than one hour variability is important for them. This was taken into account in VERB (Subbotin & Shprits, 2009) (http://rbm.epss.ucla.edu/realtime-forecast/) and BAS (Glauert et al., 2014) (http://fp7-spacecast.eu/index.php?page=he_forecasts) radiation belt codes, for example. For keV electron fluxes, smaller scale variations do not allow averaging over an orbit/day/hour and they must be considered while modeling the fluxes.

Several modeling attempts for keV electron dynamics have been made (e.g., Jordanova & Miyoshi, 2005; Miyoshi et al., 2006; Y. Chen et al., 2006; Jordanova et al., 2014) focusing mainly on the application to specific events. A couple of models, namely, the Fok Ring Current Model (FRC) (Fok & Moore, 1997; Fok et al., 1999, 2001) and the Comprehensive Inner-Magnetosphere Ionosphere (CIMI) model (Fok et al., 2001, 2011, 2014) run online at the Community Coordinated Modeling Center (CCMC) (http://ccmc.gsfc.nasa.gov/index.php) in near real time but without real time comparison with the observations. The purely empirical model for electron flux for 1 eV to 40 keV at GEO (Denton et al., 2015, 2016, 2017) based on LANL data (http://gemelli.spacescience.org/mdenton/) and dependent on the Kp index, daily F10.7 index, and $-V_{SW}B_Z$ is not well suited for modeling of the specific events and of the fast variations of keV electrons due to its limited number of driving parameters. Another empirical model, the MSSL Electron Population Model, based on Cluster PEACE and EFW instrument data from 2001-2014 provides the omni-directional 10 eV to 40 keV electron population parameterised by solar wind velocity and Kp-index at MEO (L=4-6) and GEO (L=6-7). It is not accessible without registration to the ESA Space Radiation Expert Service Centre and the resolutions of the grids in MLT, energy, and driving parameters are quite low. A very different approach is used in the SNB3GEO models (e.g., Balikhin et al., 2011) (http://www.ssg.group.shef.ac.uk/USSW/UOSSW.html) based on Multi-Input Single-Output (MISO) Nonlinear AutoRegressive Moving Average with eXogenous inputs (NARMAX) methodologies (Leontaritis & Billings, 1985a, 1985b). Boynton et al. (2016) extended the forecast to lower energies of 30-600 keV electrons using MAGED GOES satellite data which is now shown under the H2020 PROGRESS project (https://ssg.group.shef.ac.uk/progress2/html/index.phtml). In general, it is challenging to forecast keV electrons one day ahead because the same day variations in the solar wind affect the current electron flux. In Boynton et al. (2016), the past 24 hour averages for each hour were computed and they represented one hour forecasts but without smaller-time-scale variations.

The Inner Magnetosphere Particle Transport and Acceleration model (IMPTAM) (Ganushkina et al., 2013, 2014, 2015) was developed for low energy (<200 keV) electrons and has been operating online in real time since February 2013 under the EU-funded projects (http://fp7-spacecast.eu, imptam.fmi.fi) and at http://csem.engin.umich.edu/tools/imptam with the most recent version running at https://ssg.group.shef.ac.uk/progress2/html/index.phtml and at http://citrine.engin.umich.edu/imptam/. The model covers the whole inner magnetosphere from 3 R_E up to 10 R_E distances. It is driven by the real time solar wind and IMF parameters and geomagnetic indices and provides the outputs of the keV electron fluxes at a given time step at all L-shells and at all satellite orbits within the computational domain. So far, the output of IMPTAM is compared with the only data set available in real time for keV electrons in the inner magnetosphere which is the geostationary GOES 13 or GOES 15 (whenever available) MAGED data on electron fluxes at three energies (40, 75, and 150 keV). A preliminary validation study (Ganushkina et al., 2015) demonstrated that IMPTAM provides a now-cast of keV electrons comparable to the observations so that the same order of magnitude variations of the observed fluxes were reproduced. At the same time, the validation study was done only for four months of IMPTAM performance.

The quality of any model is determined by how well this model predicts the quantities being modeled as compared to real data and how much it deviates from the observations. The direct data-model, or observed-modeled electron flux, comparison alone cannot fully quantitatively reveal the model performance. There are several metrics to assess the model's quality. In the validation study by Ganushkina et al. (2015), we computed (1) the Normalized Root-Mean-Squared Deviation (NRMSD) (Walther & Moore, 2005; Wilks, 2006) and the associated standard deviations of the observations and (2) the binary event tables (Jolliffe & Stephenson, 2012) and Heidke Skill Score (HSS) (Heidke, 1926; Doswell III et al., 1990; Balch, 2008) based on them. For four months of IMPTAM performance, the NRMSD ranged from 0.015 to 0.0324 and the hit rates were reasonable (0.159-0.739) with the best hit rate reached for 75 keV electrons (0.367-0.739) but the Heidke Skill Scores were rather small (0.17 and below).

There is a need to evaluate the model performance on larger data sets and with more appropriate metrics. In the case of keV electron fluxes, there are several orders of magnitude differences at different locations along the geostationary orbit and during quiet and disturbed conditions with different levels of variability. Therefore, using of scale-dependent accuracy measures as simple model error or mean error can be problematic, since it can result in very large values due to the outliers in the data and in the model (Morley et al., 2018).

In the present paper we extend the study of Ganushkina et al. (2015) on the performance of IMPTAM by analyzing 18.5 months of IMPTAM output during its continuous operation online in near real time. In Section 2, the GOES 13 MAGED data used in the study are briefly described along with the method of determining the flight direction integrated differential electron fluxes following (Sillanp et al., 2017). Section 3 presents IMPTAM settings and driving parameters (solar wind and IMF parameters and geomagnetic indices) which were kept unchanged during the whole period analyzed in the paper. The comparative analysis of long-term variations of keV electron fluxes modeled by IMPTAM and measured by the GOES 13 MAGED instrument as dependent on IMPTAM driving parameters is given in Section 4. To evaluate quantitatively IMPTAM's overall performance, independent of its driving parameters, the appropriate metrics are introduced and computed in Section 5. The obtained results are discussed and conclusions are given in Section 6.

2 Data for IMPTAM validation: GOES MAGED electron fluxes at geostationary orbit

The only data on keV electrons in the inner Earth's magnetosphere which can be used for comparison with modeled electron fluxes by IMPTAM in real time are the measurements by the geostationary GOES 13 (or GOES 14 and 15, whenever available) MAGED instrument. The MAGED (MAGnetospheric Electron Detector) instrument is a set of nine collimated solid state telescopes (Hanser, 2011; Rowland & Weigel, 2012). The nine detectors, or telescopes, each with a 30° full-angle conical field-of-view, form a cruciform field-of-regard with the central telescope 1 pointing anti-Earthward. Each telescope measures electron fluxes in five energy channels of 30-50 keV, 50-100 keV, 100-200 keV, 200-350 keV, and 350-600 keV. The MAGED archival data are provided as directional differential electron fluxes in units of $cm^{-2}\ sec^{-1}\ sr^{-1}\ keV^{-1}$ determined for the midpoint of the five energy ranges (i.e., at 40, 75, 150, 275, and 475 keV) and given separately for all nine telescopes, as well as the pitch angles calculated from the GOES Magnetometer 1 data (Rodriguez, 2014). We consider the first three energy channels. Using electron fluxes measured by separate telescopes provides sparse information on the full distribution function at the GOES location, although they can be used to estimate the complete pitch-angle distribution (Hartley et al., 2013). Coverage of pitch angles of electrons entering a certain telescope varies with time, magnetic field changes being one of the reasons for that. Instead of determining the pitch angles measured by separate telescopes and using the corresponding fluxes of nine separate values from the nine telescopes, we compute one omni-directionally averaged flux value for each of the energies of 40 keV, 75 keV, and 150 keV, flight direction integrated differential electron fluxes, following the method presented in (Sillanp et al., 2017). Here, we briefly summarize the procedure.

Several assumptions are made when computing the flight direction integrated differential electron flux, namely, that the directional electron fluxes are (1) cylindrically symmetric with respect to the direction of the magnetic field (i.e., fluxes are uniform in all directions with the same pitch angle) and (2) symmetrically reflected with respect to the plane perpendicular to the magnetic field (i.e., fluxes for pitch angles α from 0° to 90° are the same as from 90° to 180°, $J(180° - \alpha) = J(\alpha)$).

The flight direction integrated differential electron flux for each energy channel can be computed using the directional differential electron fluxes of individual telescopes in order to get the differential fluxes in all directions, then, integrating these fluxes over the full solid angle of 4π. To avoid the confusion which may arise due to differences in units for the computed flight direction integrated differential electron fluxes (Roberts, 1965), the directional differential electron fluxes provided by separate telescopes and fluxes modeled by IMPTAM, we obtain the flight direction integrated differential electron flux J in units of $cm^{-2}\ sec^{-1}\ sr^{-1}\ keV^{-1}$ by normalizing the computed values by 4π:

$$J = \frac{1}{4\pi}\int_{4\pi} J(\Omega)d\Omega = \frac{1}{4\pi}\cdot 2\cdot 2\pi\int_0^{\pi/2} J(\alpha)\sin(\alpha)d\alpha =$$

$$= \sum_{i=1}^{n} J_i \int_{\alpha_{i0}}^{\alpha_{i1}} \sin(\alpha)d\alpha = \sum_{i=1}^{n} J_i[-\cos(\alpha_{i0}) - (-\cos(\alpha_{i1}))], \quad (1)$$

where

$$J_i = \frac{\sin(\alpha_{i0})\cdot J_{i0} + \sin(\alpha_{i1})\cdot J_{i1}}{\sin(\alpha_{i0}) + \sin(\alpha_{i1})}, \quad (2)$$

and $J(\Omega)$ is the directional flux as a function of the solid angle Ω, J_i is the differential flux for each pitch angle interval i which is the actual pitch angles of the telescopes, J_{i0} is the differential flux by a detector at the beginning of a pitch angle interval i and J_{i1} is the differential flux at the end of the interval with the corresponding pitch angles α_{i0} and α_{i1}, respectively.

In the present study we use the GOES 13 MAGED data of electron fluxes and the data for the pitch angles of each telescope with 5 minute averaging from http://satdat.ngdc.noaa.gov/sem/goes/data/

3 IMPTAM setup for modeling of keV electron fluxes at GOES 13 locations

The Inner Magnetosphere Particle Transport and Acceleration Model (IMPTAM), version for electrons (Ganushkina et al., 2013, 2014, 2015), traces distributions of electrons in the drift approximation (1st and 2nd adiabatic invariants conserved) with arbitrary pitch angles from the plasma sheet (starting at 10 R_E) to the inner L-shell regions (3 R_E) with energies reaching up to hundreds of keVs in time-dependent magnetic and electric fields. We obtain the changes in the electron distribution function $f(R, \phi, t, E_{kin}, \alpha)$, where R and ϕ are the radial and azimuthal coordinates in the equatorial plane, respectively, t is the time, E_{kin} is the particle energy, and α is the particle pitch angle, considering the drift velocity as a combination of the $\mathbf{E}\times\mathbf{B}$ drift velocity and the velocities of gradient and curvature drifts. Even the grid for distance is in R, the L-values are computed inside IMPTAM. Liouville's theorem is used to gain information of the entire distribution function with losses taken into account. For electron losses, we consider the convection outflow and pitch angle diffusion. In IMPTAM we do not use the pitch angle diffusion coefficients directly, but electron lifetimes computed from them. When running IMPTAM online in real time, we used two model representations for the electron lifetimes τ, one of M. W. Chen et al. (2005) at distances from 10 R_E, where our IMPTAM outer boundary was located, to 6 R_E and the other of Shprits et al. (2007) at distances from 6 R_E to 3 R_E, which was the IMPTAM inner boundary. The M. W. Chen et al. (2005) representation does not include any dependence on the geomagnetic activity but it includes an MLT-dependence and it can be applied when we model electron motion from the plasma sheet to geostationary orbit. The Shprits et al. (2007) representation does not include an MLT-dependence but it includes the Kp-dependence which is important when we apply these electron lifetimes at distances inside geostationary orbit. Shprits et al. (2007), and addressed only interactions due to chorus waves, hiss waves are not taken into account but this is acceptable for the comparison between the modeled and observed electron fluxes at geostationary orbit. For the obtained distribution function, we apply radial diffusion by solving the radial diffusion equation (Schulz & Lanzerotti, 1974). Kp-dependent radial diffusion coefficients D_{LL} for the magnetic field fluctuations are computed following Brautigam and Albert (2000). After that, we repeat the order of calculation: first, we solve transport with losses and then apply the diffusion. More detailed description of IMPTAM is given in Ganushkina et al. (2014) and Ganushkina et al. (2015).

The IMPTAM nowcast (imptam.fmi.fi) for low energy (1-200 keV) electrons in the inner magnetosphere has been operating online since February 2013 in near-real time under the FP7 SPACECAST (http://fp7-spacecast.eu), SPACESTORM (http://www.spacestorm.eu/) and H2020 PROGRESS (https://ssg.group.shef.ac.uk/progress2/html/) projects funded by the European Commission. Real time geostationary GOES 13 MAGED data on electron fluxes for three energies of 40, 75 and 150 keV have been used for comparison and validation of IMPTAM running online (Ganushkina et al., 2015). IMPTAM is driven by the solar wind and IMF parameters and geomagnetic indices obtained in real time.

Inside IMPTAM, the set of models which was found to provide best agreement with the measured electron fluxes at geostationary orbit is used, namely, (1) a dipole model for the internal magnetic field, (2) T96 model (Tsyganenko, 1995) for the external magnetic field, and (3) (Boyle et al., 1997) polar cap potential mapped to the magnetosphere. The T96 model uses the Dst index, solar wind pressure P_{SW}, and IMF B_Y and B_Z components as input parameters. We re-compute the magnetic field configuration in the entire modeling domain every 5 minutes using the observed, 5 minute-averaged P_{SW} and IMF B_Y and B_Z and, instead of hourly Dst index, we use 5 minute SYM-H index for consistency with other parameters. Wanliss and Showalter (2006) showed that the Dst and SYM-H indices correlate with a coefficient higher than 0.9, indicating that they can be used interchangeably. Furthermore, Katus and Liemohn (2013) demonstrated that, during storm times, these indices are close to each other but can vary from each by up to 20%. This is an acceptable difference that allows for a higher-time resolution of this input parameter to the T96 model. The electric field (Boyle et al., 1997) is determined using the solar wind speed V_{SW}, the IMF strength |B| and its components B_Y and B_Z (via IMF clock angle θ_{IMF}) dependent on radial distance and MLT. We set the model boundary at 10 R_E and use the kappa electron distribution function. Parameters of the kappa distribution function are the number density n and temperature T in the plasma sheet given by the empirical model derived from Geotail data by (Tsyganenko & Mukai, 2003). In IMPTAM simulation, the electron n is assumed to be the same as that for ions in the model, but $T_e/T_i = 0.2$ is taken into account. The (Tsyganenko & Mukai, 2003) model uses as input parameters the solar wind speed V_{SW} and density N_{SW} as well as the B_Z component of IMF. Kp-index is a parameter for the radial diffusion coefficients D_{LL} and (Shprits et al., 2007) electron lifetimes. Thus, the IMPTAM driving parameters are (1) the IMF B_Z and (2) B_Y components, (3) the IMF strength |B|, (4) the solar wind speed V_{SW} and (5) dynamic pressure P_{SW}, (6) Kp and (7) SYM-H indices. These parameters are of primary interest in data-model comparison. The comparison between the keV electron fluxes modeled by IMPTAM and measured by GOES 13 MAGED instrument presented here is for the period from 20 September 2013 (by then, the initial checks of IMPTAM running online which started in February 2013 were done) to 31 March 2015. During this period, the model's settings were not changed. For the IMPTAM input parameters, we used the openly available ACE data (http://services.swpc.noaa.gov/text/) together with data from OMNIWeb (http://omniweb.gsfc.nasa.gov/) and the World Data Center for Geomagnetism, Kyoto (http://wdc.kugi.kyoto-u.ac.jp/wdc/Sec3.html).

4 Comparative analysis of long-term variations of keV electron fluxes modeled by IMPTAM and measured by GOES 13 MAGED instrument at geostationary orbit

We use the 5 minute averaging for GOES 13 MAGED data and the 5 minute IMPTAM output as flight-direction integrated differential fluxes for energies of 40, 75, and 150 keV that are directly comparable during the period between September 20, 2013 and March 31, 2015. The direct data-model comparison during two periods, two months of July-August 2013 and four months of January-April 2014 was analyzed in Ganushkina et al. (2015). Time series of the observed and modeled fluxes over a 18.5 months period are presented in Figure 1 together with IMPTAM driving parameters. Since keV electron fluxes vary at rather short time scales, the conclusions which can be made from this Figure 1 are limited to the following:

(1) the modeled 40 keV electron fluxes vary within the range observed by GOES 13 MAGED but, at the same time, sharp dropouts are not reproduced;

(2) the modeled 75 keV electron fluxes have a narrower range than observed, but fail to fit the dropouts and smaller fluxes;

(3) in general, statement (2) is true also for 150 keV electrons.

Looking at this Figure 1, it is very difficult to make any conclusions about the influence of driving parameters upon the modeled fluxes. Therefore, we analyze in details the observed and modeled electron fluxes organized by MLT along the GOES 13 orbit and the IMPTAM driving parameters (IMF B_Z, B_Y components and strength |B|, V_{SW} and P_{SW}, Kp and SYM-H), instead of direct data-model comparison for the modeled period. This approach can provide more insights into the influence of the different parameters on the IMPTAM performance quality. Figures 2-8 present the comparison results. The MAGED electron fluxes (panels (a), (d), and (g)) and the IMPTAM modeled electron fluxes (panels (b), (e), and (h)) for the three energies of 40, 75, and 150 keV are plotted in the same logarithmic scale. Panels (c), (f), and (i) present the ratio between the modeled and observed fluxes in the logarithmic scale. Bottom panel (j) shows the data counts for the occurrence of a corresponding driving parameter.

Figure 2 shows the modeled (panels on the left) and the observed (panels in the middle) electron fluxes binned by MLT with 1 hour step and IMF B_Z with 1 nT step. The fluxes were computed as the average fluxes from all datapoints which fall into certain bins but plotted in the logarithmic scale. In addition, the ratio between the modeled and observed fluxes, after averaging those fluxes in each bin, is shown in panels on the right, also plotted in logarithmic scale. The way how this ratio was computed, when one average (of modeled fluxes in a bin) was divided by another average (of observed fluxes in a bin), results in higher fluxes being given more weight in it. The ratio of the averaged values ($\frac{\sum IMPTAMflux}{\sum GOESflux}$) will not be equal to the averaged ratio of the same values ($\sum \frac{IMPTAMflux}{GOESflux}$) in which lower fluxes will have more influence. In our present study, we compute the ratio between the averaged values, since we wanted to focus on the ability of IMPTAM to reproduce the higher fluxes which can be reached by keV electrons at the geostationary orbit. This focus is due to the fact that the surface potential of a spacecraft can become significant ranging from several to ten kV as long as electron fluxes exceed a spacecraft-dependent threshold level. For specific spacecraft and their surfaces, certain electron energies are of most importance and the threshold depends on them. For example, at the LANL satellites, the most important energies for surface charging were found to be ranging from 5 to 50 keV (Thomsen et al., 2013; Mato-Vlez et al., 2018). For GOES, we do not possess readily such information, therefore, the range of higher fluxes, observed and modeled, was given special attention here and the ratio was computed between the averaged values.

Since this Figure 2 contains all the points with corresponding IMF B_Z values, Figure 2j gives the distribution of data counts within the observed range of MLT and IMF

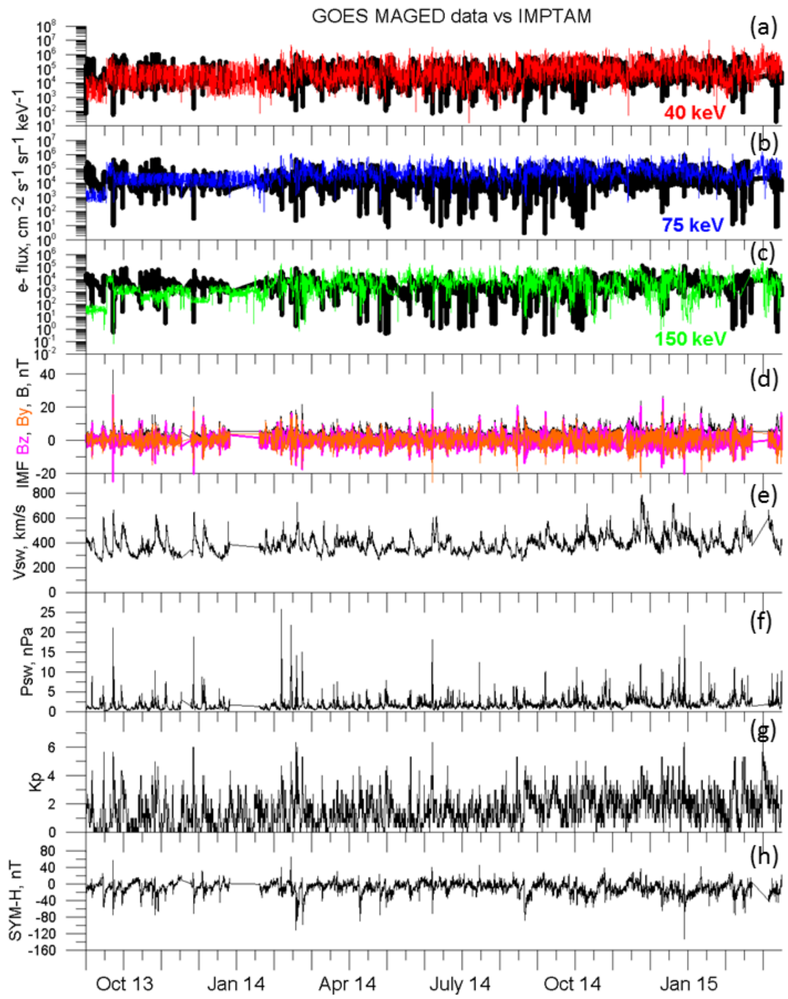


Figure 1. IMPTAM performance run in real time: the observed fluxes (black lines) at GOES 13 together with the modeled fluxes for (a) 40 keV (red line), (b) 75 keV (blue line), and (c) 150 keV electrons (green line) with model driving parameters as observed (d) IMF Bz (pink line), By (orange line) and B (black line), (e) solar wind velocity, and (f) solar wind dynamic pressure and geomagnetic indices (g) Kp and (h) SYM-H.

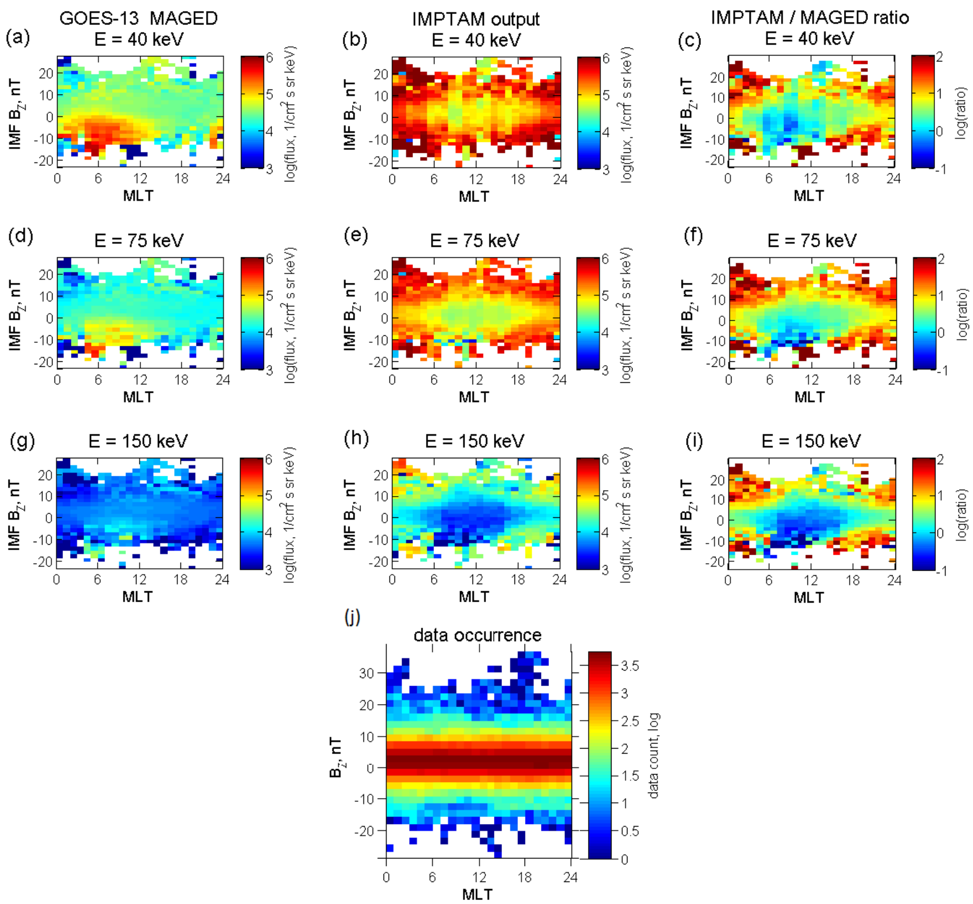


Figure 2. Flight-direction integrated differential electron fluxes in logarithmic scale for the energies of 40, 75, and 150 keV computed from the GOES 13 MAGED data (panels (a), (d), and (g)) and modeled by the IMPTAM (panels (b), (e), and (h)) binned by MLT and IMF B_Z, and then averaged, together with the ratio between them in logarithmic scale (panels (c), (f), and (i)). Bottom panel (j) shows the data counts for the IMF B_Z occurrence.

B_Z. From Figure 2j, we can see that the maximum occurrence of data points is for IMF B_Z from 0 to +5 nT with about 10^4 points per bin and all points above +10 nT and below -8 nT constitute less than 10% of the maximum number of points in that MLT range. Points with IMF B_Z above +20 nT and below -15 nT are already less than 1% of the maximum number of points. Therefore, for our analysis, the main attention will be paid to the modeled and observed fluxes which fall into the IMF B_Z range of -10 to 10 nT (same absolute values for negative and positive B_Z are chosen to make the analysis of Figure 2 easier).

The observed 40 keV electron fluxes (Figure 2a) exhibit the clear peak reaching to 10^6 cm^{-2} sec^{-1} sr^{-1} keV^{-1} for negative IMF B_Z (0 to -10 nT) located at a rather wide midnight-dawn-noon sector of 00-12 MLT. The 40 keV electron flux for positive IMF B_Z in this MLT sector and for all values of IMF B_Z in the noon-dusk-midnight sector is about the same, being of $5{-}8{\cdot}10^4$ cm^{-2} sec^{-1} sr^{-1} keV^{-1}. For higher energies the pattern of electron flux dependence on MLT and IMF B_Z is very similar with fluxes being lower. The peak values for 75 keV electrons (Figure 2d) are around $5{\cdot}10^5$ cm^{-2} sec^{-1} sr^{-1} keV^{-1} and for 150 keV electrons (Figure 2g) they are about $5 \cdot 10^4$ cm^{-2} sec^{-1} sr^{-1} keV^{-1} and located on the dawn sector. In general, the observed geostationary keV electron fluxes are very clearly organized by IMF B_Z with maximum fluxes located at around 06 MLT.

One particular location of higher observed fluxes can be seen very close to 04-06 MLT for IMF B_Z of about -18 to -12 nT. Comparing this location to the number of data points presented in Figure 2j tells us that such high fluxes can be the result of averaging over a small number of points where higher values of the observed fluxes get larger weights. This can be unrealistic and very different if there would have been more, statistically valuable data points. The same is true for smaller peaks seen at 12-16 MLT for IMF B_Z above 20 nT.

Keeping in mind the number of actual data points corresponding to different IMF B_Z is especially important when analyzing the modeled fluxes. If we concentrate at the range of -10 to 10 nT of IMF B_Z, it can be seen that the modeled electron fluxes have similar peaks for negative IMF B_Z (Figures 2b, e and h) but the maxima of the peaks are located not at around 06 MLT as observed but shifted towards midnight being between 00 and 06 MLT. The modeled fluxes have peaks at large ($>$ 10 nT) positive IMF B_Z at around 18-06 MLT for all three energies which are not seen in the observed fluxes. At the locations of these peaks, the difference of one to two orders of magnitude can be seen (Figures 2c, f and i). As was stated above, this is the IMF B_Z range where the number of data points was less than 10% of the maximum number of points in that MLT range. For negative IMF B_Z, the ratio can also reach one to two orders of magnitude but it is mainly for IMF B_Z below -10 nT. At the same time, the ratio between the modeled and the observed fluxes at 00-12 MLT where the observed peak is located is close to one and up to 10 for several values of IMF B_Z for the presented statistics.

In a similar way as presented in Figure 2, Figure 3 shows the modeled and the observed electron fluxes binned by MLT and IMF B_Y, and then averaged, together with the ratio between them and the distribution of data counts within the observed range of MLT and IMF B_Y. Following the same estimates as for Figure 2j, we can say that all points above +12 nT and below -10 nT constitute less than 10% of the maximum number of points in any given MLT range, so our analysis is concentrated at the range between -10 and +10 nT for IMF B_Y. The observed 40 keV electron fluxes (Figure 3a) show the x-shaped peak, again located at around 06 MLT, with values of about $5{\cdot}10^5$ cm^{-2} sec^{-1} sr^{-1} keV^{-1}. The peak widens in MLT (from midnight to noon) with the increase of negative and positive values of IMF B_Y in magnitude being narrower ($\pm$ 2 hours from 06 MLT) for IMF B_Y close to zero. Similar peaks, but an order of magnitude lower and shifted a little more towards noon than the previous ones, are visible for 75 keV (Figure 3d) and 150 keV electrons (Figure 3g). The modeled fluxes exhibit very similar x-shaped structure but shifted towards midnight (Figures 3b, e, and h). Due to this shift, the modeled fluxes are one

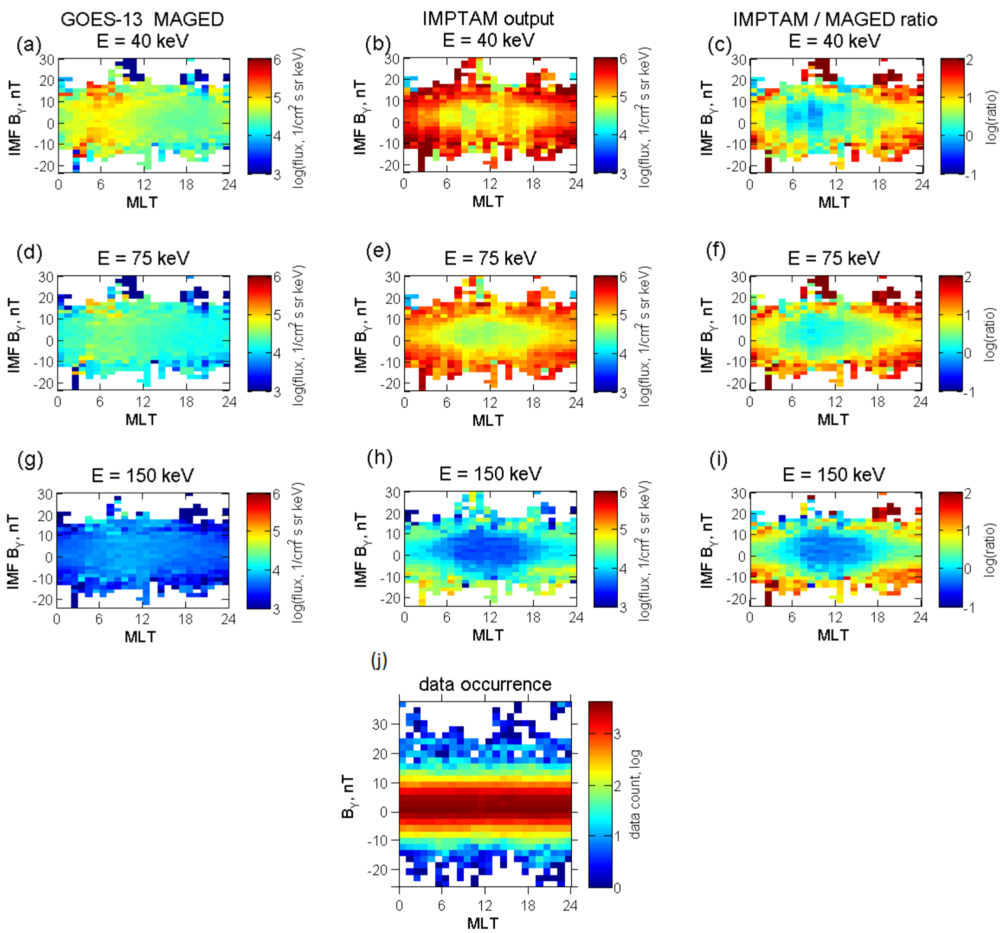


Figure 3. Similar to Figure 2 but the observed and modeled electron fluxes are binned by MLT and IMF B_Y and then averaged with IMF B_Y data occurrence.

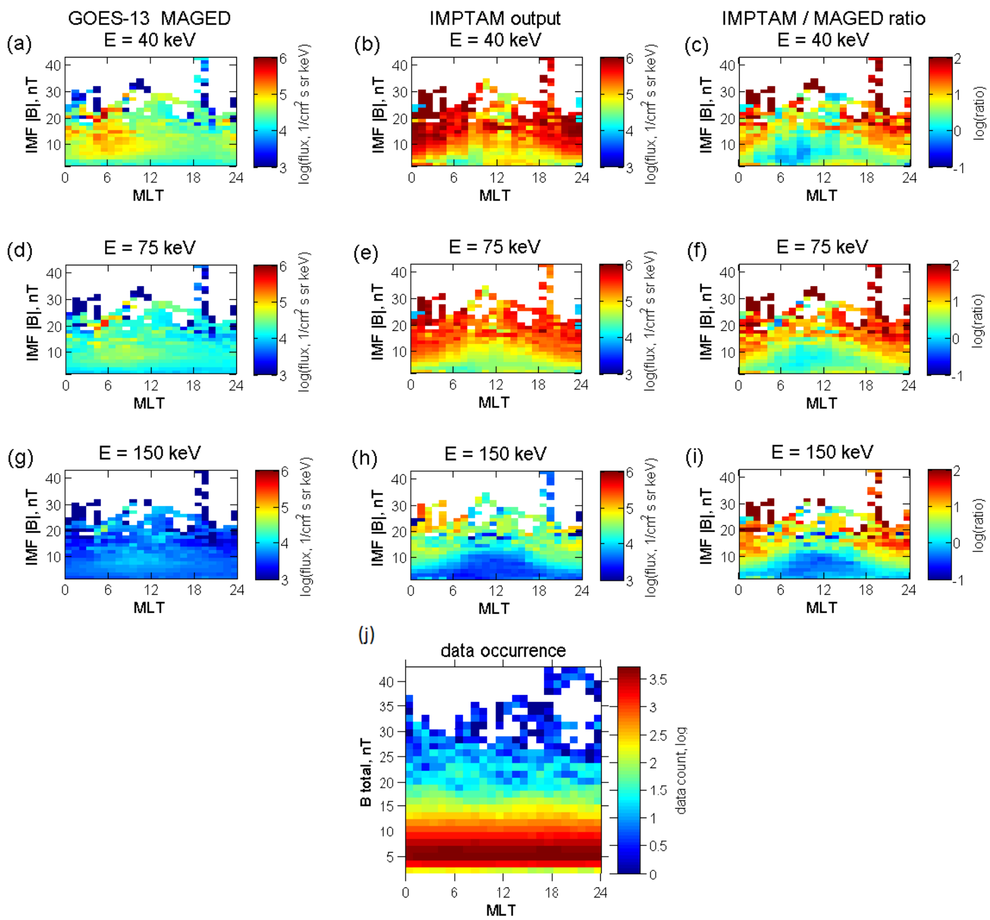


Figure 4. Similar to Figure 2 but the observed and modeled electron fluxes are binned by MLT and IMF |B| and then averaged with IMF |B| data occurrence.

to two orders of magnitude higher than the observed ones at around 18-02 MLT for both positive and negative IMF B_Y values. At 06-12 MLT the ratio is close to one or smaller indicating the difference in fluxes with the modeled smaller than the observed up to one order in magnitude (Figures 3c, f, and i).

Figure 4 presents the modeled and observed electron fluxes binned by MLT and IMF total strength |B|, and then averaged, together with the ratio between them and the distribution of data counts within the observed range of MLT and IMF |B|. Figure 4j indicates that the data-model comparison needs to be done for IMF |B| below about 20 nT. The observed fluxes show quite similar features as in Figure 2 with peaks at 00-12 MLT but with inverted-V shapes and with an order of magnitude lower and shifted more towards noon than the previous ones with energy (Figures 4a, d, g). The modeled fluxes can reach of one to two orders of magnitude difference at 18-06 MLT for larger (10 to 20 nT) values of IMF |B| (Figures 4b, e, h) but at 06-12 MLT for IMF |B| $<$ 15 nT, the ratio between them and the observed ones is close to one (Figures 4c, f, i).

Figure 5 presents the modeled and observed electron fluxes binned by MLT with 1 hour step and solar wind speed V_{SW} with 20 km/s step, and then averaged, together with the ratio between them and the distribution of data counts within the observed range of MLT and V_{SW}. Based on Figure 5j, datapoints with corresponding V_{SW} above 700

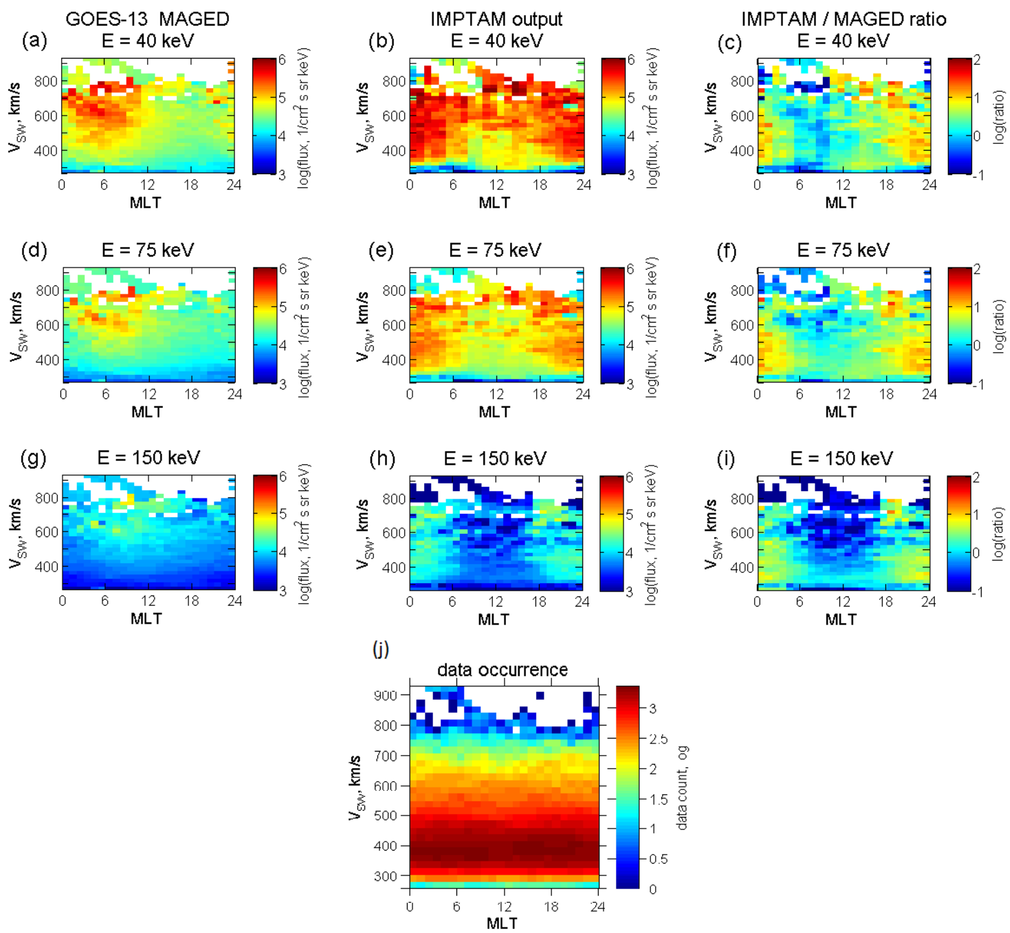


Figure 5. Similar to Figure 2 but the observed and modeled electron fluxes are binned by MLT and solar wind speed V_{SW} and then averaged with V_{SW} data occurrence.

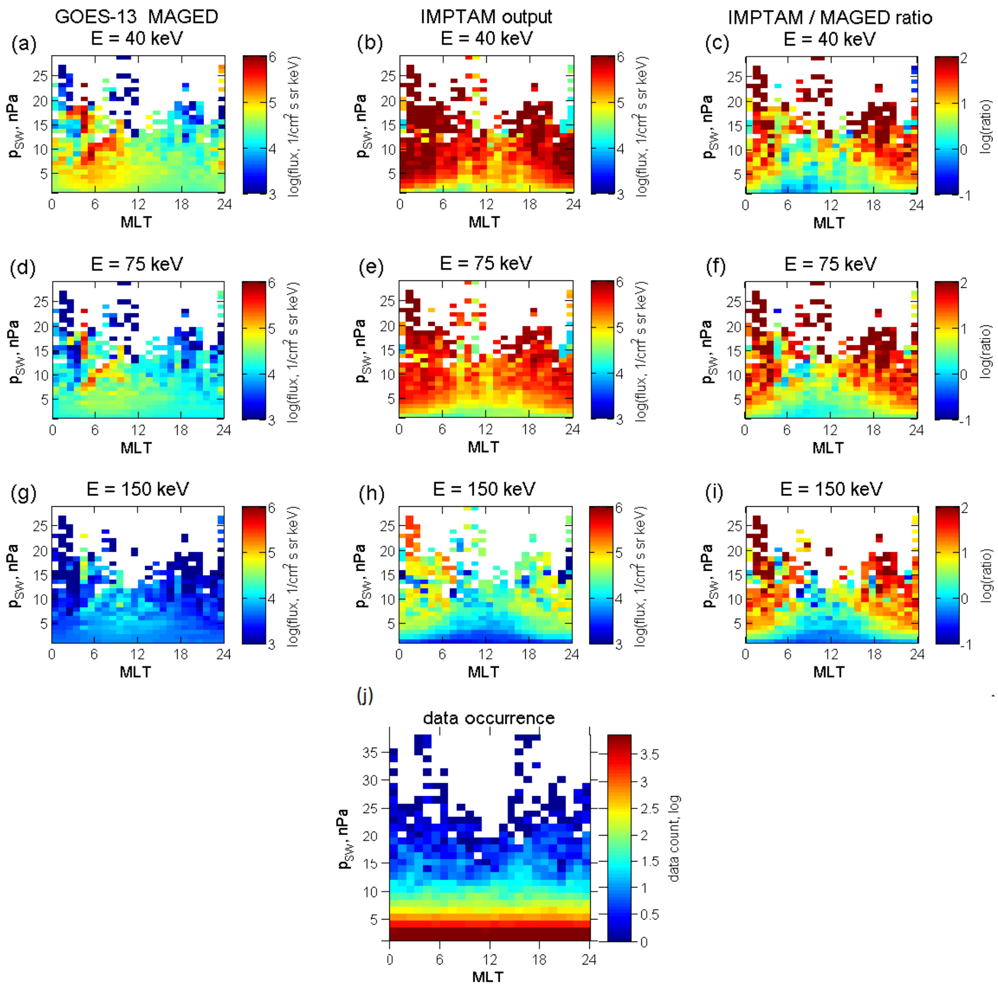


Figure 6. Similar to Figure 2 but the observed and modeled electron fluxes are binned by MLT and solar wind dynamic pressure P_{SW} and then averaged with P_{SW} data occurrence.

km/s constitute less than 10% from the maximum number of points per bin and the corresponding structures in the observed and modeled fluxes can be disregarded. The U-shaped peaks in the observed electron fluxes are located at 00-12 MLT as in previous figures and the fluxes increase with the increase of V_{SW} covering larger range of MLT. The modeled fluxes of about $5 \cdot 10^5 - 10^6$ cm^{-2} sec^{-1} sr^{-1} keV^{-1} for 40 keV electrons are present at a wider than observed range of MLTs (20 -04) for V_{SW} above 200 km/s. The same is true for 75 keVs but with order of magnitude lower fluxes (or 2 orders of magnitude for 150 keV). Shifts of the peaks to midnight instead of dawn are also present. Looking at the ratio we can see that the modeled fluxes are rather close to the observed ones at 06-12 MLT. The main over-estimation is seen at around midnight with about one order of magnitude.

Figure 6 demonstrates the modeled and observed electron fluxes binned by MLT with 1 hour step and solar wind dynamic pressure P_{SW} with 1 nPa step, and then averaged, together with the ratio between them and the distribution of data counts within the observed range of MLT and P_{SW}. As can be seen in Figure 6j, analyzing the observed and modeled fluxes with P_{SW} above 10 nPa can lead to unreasonable conclusions, since the number of points there is less than 10% from the maximum number of points per bin. The largest observed 40 keV electron fluxes (Figure 6a) are located at 00-12 MLT peaking at around 06 MLT and increasing with the increase of P_{SW}. Similar features are seen for 75 and 150 keV electron fluxes (Figures 6d and g) but with peaks shifted towards noon and with order of magnitude smaller values as in all figures described above. The modeled fluxes are higher than the observed ones at 18-06 MLT with the difference reaching about 1.5 orders of magnitude at around midnight and 18 MLT. Again, in the MLT sector of 06-12 for $P_{SW} < 10$ nPa, the ratio between the modeled and observed fluxes can be close to one.

In addition to the IMF and solar wind parameters, we present the statistical dependencies on the geomagnetic indices Kp and SYM-H which are also the driving parameters for IMPTAM. Figure 7 presents the modeled and observed electron fluxes binned by MLT with 1 hour step and Kp-index with 4 steps when moving from one Kp-value to the next, and then averaged, together with the ratio between them and the distribution of data counts within the observed range of MLT and Kp. Contrary to the IMF and solar wind parameters, many more datapoints need to be considered in our analysis, except of those with Kp>5 as can be seen in Figure 7j. The similar pattern how the observed electron fluxes depend on the Kp-index along the geostationary orbit was previously reported using LANL MPA data (Korth et al., 1999) and Polar HYDRA data (Friedel et al., 2001). It is rather similar to the one for V_{SW} (Figure 5) with the U-shaped peaks on the dawnside with fluxes increasing as Kp increases. The modeled fluxes exhibit two orders of magnitude difference at around midnight for Kp greater than 5 but these correspond to statistically less meaningful bins. They are close to the observed fluxes at 06-12 MLT with the ratio of one or less.

Figure 8 shows the modeled and observed electron fluxes binned by MLT with 1 hour step and SYM-H index with 5 nT step, and then averaged, together with the ratio between them and the distribution of data counts within the observed range of MLT and SYM-H. According to Figure 8j, we take into account the datapoints with SYM-H below 50 nT and above -60 nT. The observed 40 keV fluxes (Figure 8a) exhibit a clear peak for negative SYM-H values located at 00-06 MLT. This peak is present for 75 keV (Figure 8b) and 150 keV (Figure 8c) electron fluxes with an order of magnitude smaller fluxes but similar MLT location. The modeled fluxes again show the shift towards midnight and order of magntiude over-estimates at 18-24 MLT. The ratio is close to one at around 06-12 MLT.

5 Metrics for model performance

The quality of any model is determined by how well this model predicts the quantities being modeled as compared to the real data and how much it deviates from the observations. There are several metrics to assess the model's quality and many of them have been successfully applied to terrestrial weather forecast models (Murphy, 1993; Thornes & Stephenson, 2001; Jolliffe & Stephenson, 2012). With the intense development of space weather forecast models, similar metrics can be applied for them, too (e.g., Lopez et al., 2007; Welling & Ridley, 2010; Pulkkinen et al., 2013; Ganushkina et al., 2015; Morley, 2016; Morley et al., 2018).

Before computing the necessary metrics, in Figure 9, we present the scatter plots of GOES MAGED electron fluxes vs. fluxes by IMPTAM for (a) 50, (b) 75, and (c) 150 keV. We overplot the fluxes with the scatter density which converts the population density of the data into a logarithmic gradient. This logarithmic gradient of the points is denoted by the colorbar in these plots. As expected, there is no obvious one-to-one correlation. The observed dropouts (lowest fluxes for all three energies) are not reproduced (modeled fluxes stay high). It is also seen that there are times of low modeled fluxes that are not observed. These are dropouts from magnetopause shadowing in the model that

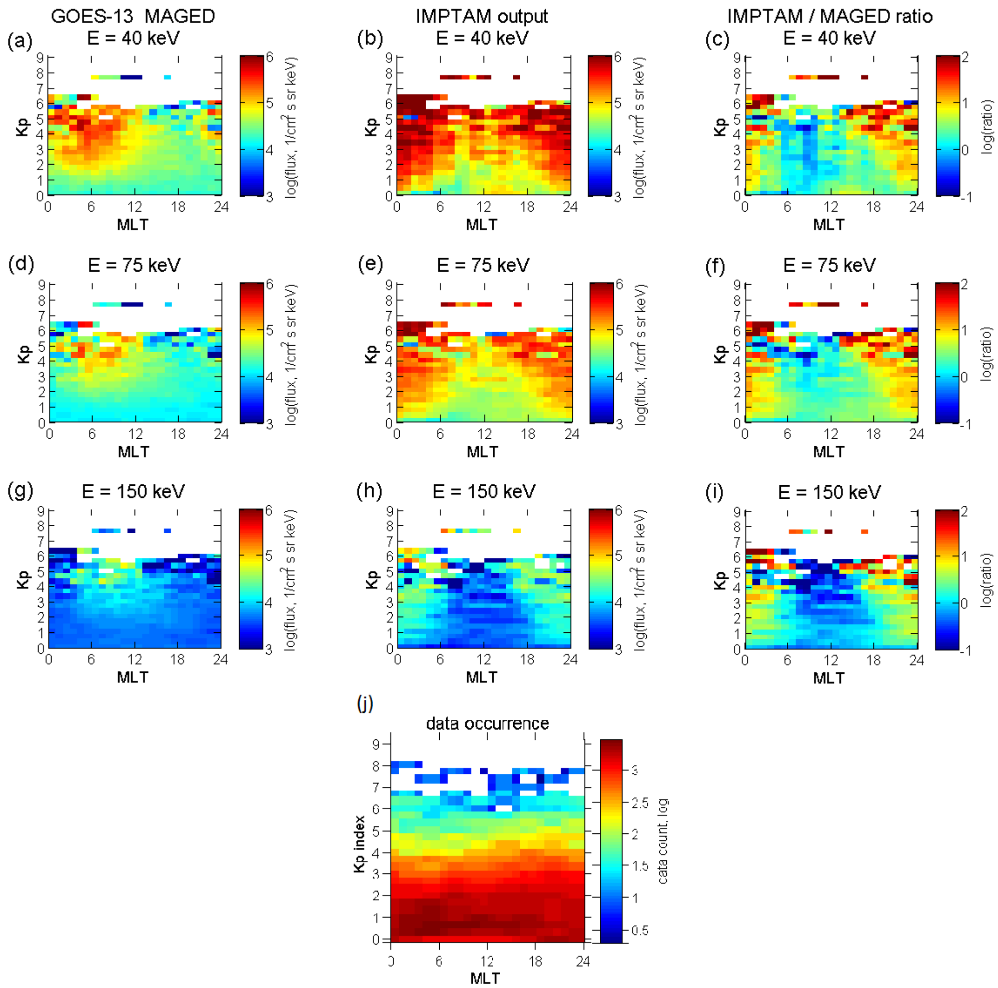


Figure 7. Similar to Figure 2 but the observed and modeled electron fluxes are binned by MLT and Kp-index and then averaged with Kp data occurrence.

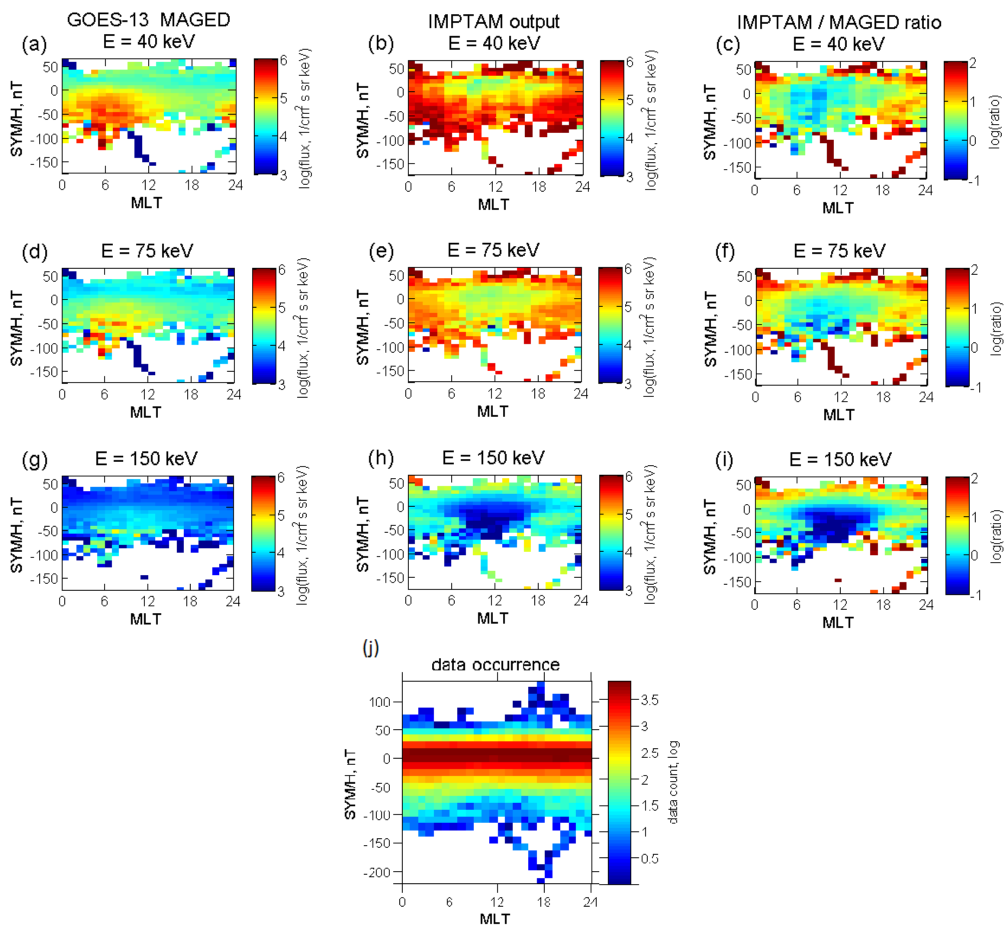


Figure 8. Similar to Figure 2 but the observed and modeled electron fluxes are binned by MLT and SYM-H index and then averaged with SYM-H data occurrence.

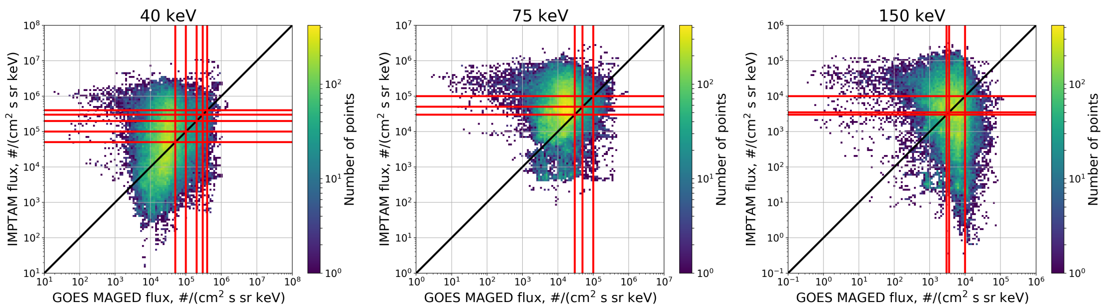


Figure 9. Scatter plots of GOES MAGED electron fluxes vs. modeled fluxes by IMPTAM for (a) 50, (b) 75, and (c) 150 keV overplotted with population density of the data, together with thresholds used for binary event analysis marked by red lines.

were not seen at GOES. Despite these "wings" of the distribution, there is a large cloud of points within an order of magnitude along the one-to-one black diagonal line.

To evaluate the quality of the electron flux forecasts made by the IMPTAM, we employ the binary event analysis (Jolliffe & Stephenson, 2012). This methodology first divides the time series data into non-overlapping time windows. Each interval is then categorized by the behavior of the model and observation with respect to a given threshold: it is considered a "Hit" if the model and data both cross the threshold, a "Miss" if the observation does but the model does not, a "False Positive" of the model does but the data does not, and a "True Negative" if neither cross. The thresholds for each energy level are given in the first column of Table 1 and Figure 9 shows them as red lines over the scatter plots. Ideally, these thresholds need to be meaningful for applications based on the fact that the surface charging can begin when electron fluxes exceed the threshold level which is spacecraft and energy dependent. Since we do not know them for GOES MAGED data, the selection of threshold levels is somewhat arbitrary. Any particular percentile of the observed flux is no more meaningful, either, since the surface potential on a satellite is not determined by a specific percentile. Therefore, in the present study we select several thresholds so that binary event metrics have enough events (i.e., threshold crossings) to be useful and the thresholds correspond to our previous analysis (Ganushkina et al., 2015) to be able to compare the results.

Following the our previous work (Ganushkina et al., 2015), the window width is set to one hour. One hour is rather long as compared to the model output every 5 minutes, the flux can vary significantly within an hour, but in the present study, the selected window is chosen to test "bulk activity". Columns 2-5 in Table 1 contain the actual numbers of Hits, False Positives, Misses and True Negatives. Descriptive metrics and skill scores can be calculated from them. One metric is the "Hit Rate", or the ratio of correctly predicted threshold crossings to all observed crossings. It ranges from 0 to 1, with 1 being perfect. Next is "False Alarm Rate", or the fraction of false alarms to all non-event intervals. Here, 0 is perfect and 1 indicates the model predicts a crossing at all times. Finally, we list the "Heidke Skill Score", which is the fraction of correct predictions when adjusted for those expected from pure random chance, is calculated. This value has the range [-1,1] where 1 is perfect and zero corresponds to a performance that is indistinguishable from random chance. The results for this analysis are shown in Table 1. In general, the model has an appreciable hit rate (20-78% of threshold crossings are reproduced depending on energy and threshold, see Table 1) for all energy levels and all thresholds. However, this is offset by considerable false alarm rates (crossings were incorrectly predicted during 14% to 50% of non-event times, see Table 1), which keep the Heidke skill scores modest at best. For 40 keV electrons, the model correctly forecasts 6 to 16% more

Table 1. Binary event analysis results for each energy channel as a function of flux threshold

Threshold, $cm^{-2}\ sec^{-1}\ sr^{-1}\ keV^{-1}$	Hit	False Positive	Miss	True Negative	Hit Rate	False Alarm Rate	Heidke Skill Score
40 keV electron fluxes							
$5 \cdot 10^4$	2051	3458	868	3419	0.703	0.503	0.159
$1 \cdot 10^5$	801	3217	553	5225	0.592	0.381	0.115
$2 \cdot 10^5$	346	2154	344	6952	0.501	0.237	0.120
$3 \cdot 10^5$	180	1702	197	7717	0.477	0.181	0.102
$4 \cdot 10^5$	84	1403	128	8181	0.396	0.146	0.063
75 keV electron fluxes							
$3 \cdot 10^4$	1707	5048	473	2598	0.783	0.660	0.070
$5 \cdot 10^4$	634	4303	394	4495	0.617	0.489	0.048
$1 \cdot 10^5$	154	2753	226	6693	0.405	0.291	0.027
150 keV electron fluxes							
$3 \cdot 10^3$	3717	1790	2931	730	0.559	0.710	-0.133
$3.5 \cdot 10^3$	3112	2062	2996	998	0.509	0.674	-0.153
$1 \cdot 10^4$	299	2561	1125	5183	0.210	0.331	-0.086

events than what is expected from a random forecast (and a bit lower values, 3-7%, for 75 keV). For 150 keV electrons, the performance is worse than a random forecast as the numbers are negative. In summary, the model is performing best at the 40 keV channel and for lower thresholds. It struggles at the 150 keV channel. The scores are in line with those reported in Ganushkina et al. (2015), with improvements found in the 40 keV channel predictions.

The IMPTAM performance level presented above is rather expected, since in case of electron fluxes observed by GOES MAGED, there are several orders of magnitude differences between the fluxes at different locations along geostationary orbit and during quiet and disturbed conditions with different levels of variability. For this reason, for example, using the scale-dependent accuracy measures such as simple model error or mean error can be problematic, since it can result in very large values due to the outliers in the data and in the model. Outliers influence the model performance significantly more than small deviations from the observations (Morley et al., 2018). Morley et al. (2018) presented a very thorough analysis of other descriptive metrics which can help better to illustrate the performance of the model. In our study, we follow the Morley et al. (2018) findings.

The first metric used here is the well-known Pearson correlation coefficient, a measure of linear correlation between the observations and model results. Next is “Median Symmetric Accuracy” (designated as ζ) expressed as

$$\zeta = 100(exp(M(\mid log_e(Q_i) \mid)) - 1), \tag{3}$$

where $Q_i = \frac{y_i}{x_i}$ is the accuracy ratio, which is the ratio between the modeled y_i and the observed x_i fluxes. As was shown in Tofallis (2015), $log_e(Q)$ is best for the data with the variance depending on the magnitude of the variable which is the case for radiation belt electron fluxes and where this metric has been previously used (e.g., Morley, 2016; Reeves et al., 2011). Absolute values of $log_e(Q)$ makes sure that the metric is symmetric (when the values of the modeled and observed fluxes are switched, the error is the same). The median function M and then exponent is used to return to the original units and scale.

One is subtracted so that the metric is in the $[0, \infty)$ range, multiplying by 100 gives the equivalent percentage error. Median symmetric accuracy can be interpreted as the median percentage error. For example, if $\zeta = 50\%$, the model is most frequently reporting values that are 50% larger or smaller than the observation at any given point.

The final metric used here is Symmetric Signed Percentage Bias ($SSPB$). The bias describes the difference between the average model output and the average observation. A negative bias indicates a systematic under-prediction, whereas a positive bias indicates a systematic over-prediction. Morley et al. (2018) presented a new measure of bias based on the log accuracy ratio

$$SSPB = 100sgn(M(log_e(Q_i)))(exp(\mid M(log_e(Q_i)) \mid) - 1). \tag{4}$$

The magnitude of the bias is estimated by taking the absolute value of $M(log_e(Q_i))$, one is subtracted so that the lower limit is zero, the direction of the bias is found using the signum function, and the metric is multiplied by 100 to express it as a percentage. This value reports the median bias of the model as a percentage of the observed value. For example, if $SSPB = -50\%$, the model is biased towards underprediction, most frequently reporting values that are 50% less than the corresponding observations. Both ζ and $SSPB$ are defined in detail by Morley et al. (2018). While other measures of bias and accuracy exist, these are robust to data that spans orders of magnitude, as is the case with inner magnetosphere electron fluxes.

The correlation, accuracy, and bias metrics for the IMPTAM dataset compared to GOES 13 are shown in Table 2. Overall correlation is weak and appears inversely proportional to electron energy. ζ values demonstrate that the predictions are typically off by 200% (of almost 300% for 150 keV electrons). $SSPB$ shows that the direction of the error varies with energy. Considered together with the binary event analysis, performance is best at the 40 keV level.

Table 2. Descriptive metrics for each energy channel.

	Energy Channel		
	40keV	75keV	150keV
Corr. Coeff	0.1300	0.0390	-0.1227
Accuracy (ζ)	232.75%	244.36%	292.40%
Bias ($SSPB$)	049.04%	189.34%	-17.86%

6 Discussion and Conclusions

We presented the validation study of the performance of the model for electrons with energies of 1 to few hundreds of keVs (IMPTAM) at geostationary orbit. keV electrons are important constituents of the near-Earth's radiation environment being the seed population for further acceleration to MeV energies in the radiation belts and posing a serious risk of surface charging for satellites. The 18.5 months of IMPTAM output taken from its continuous operation online in real time was compared to the corresponding data from the GOES 13 MAGED instrument for the flight direction integrated differential fluxes for energies of 40, 75, and 150 keV. In addition to the direct data-model comparison during the entire modeled period (as was done in Ganushkina et al. (2015)), the observed and modeled electron fluxes were organized by MLT along the GOES 13 orbit and the solar wind and IMF parameters and geomagnetic indices (IMF B_Z, B_Y components and strength |B|, V_{SW} and P_{SW}, Kp and SYM-H) which are the driving parameters for IMPTAM and then compared. This approach provided more insights into the influence of the different parameters on the IMPTAM performance quality.

All the statistical patterns for all three energies binned by MLT and IMPTAM driving parameters have their peaks in electron fluxes at the dawnside as would be expected from the motion of electrons in the inner magnetosphere but the peaks for modeled fluxes are located not at around 06 MLT as for the observed fluxes but shifted towards midnight being between 00 and 06 MLT. This does not mean that the electrons in IMPTAM do not drift dawnward (the ratio between the observed and modeled fluxes at around 06 MLT is very close to one). This indicates that the modeled flux at around midnight is too high. There are several possible reasons for this. One of them is the representation of electron losses by introducing electron lifetimes as a combination of M. W. Chen et al. (2005) and Shprits et al. (2007) electron lifetimes for strong and weak diffusion, respectively. The Shprits et al. (2007) representation does not include the MLT-dependence but it has the Kp-dependence which is important when we apply these electron lifetimes at distances inside geostationary orbit. There is no dependence on geomagnetic activity in the M. W. Chen et al. (2005) representation but the MLT-dependence is present (although rather homogeneous and weak as can be seen in Figure 5 of M. W. Chen et al. (2005)) and it can be applied when we model electron motion from the plasma sheet to geostationary orbit. The model for electron lifetimes used in the present paper lacks the realistic distribution of waves as compared to, for example, the model of electron lifetimes due to interactions with chorus waves by Orlova and Shprits (2014) and with hiss waves by Orlova et al. (2016). These models are now incorporated into the new version of IMPTAM. For the 18.5 months of IMPTAM run, we used what was available at that time and the run was done without any changes.

Another reason for the excessive amount of electrons around midnight is the symmetry in the models used inside IMPTAM. For the electric field model, we used the Boyle et al. (1997) polar cap potential dependent on IMF and solar wind parameters but applied this to a Volland-Stern type two-cell convection pattern. Our choice was based on the need for dependence on IMF and solar wind parameters yet keeping it a rather simple model. There exist numerous models which can be used for the global convection electric field in the magnetosphere. In reality, particle transport from the plasma sheet does not occur in the Boyle-type potential. There are studies on penetration electric field (e.g., Ridley & Liemohn, 2002; Liemohn et al., 2004), concentrations of potential in narrow channels resulting in a fast transport of plasma sheet particles to the inner magnetosphere (M. W. Chen et al., 2003), existence of an extra potential well near local midnight (Fok et al., 2001, 2003). Usage of a simple representation for the electric field contributes to the presence of higher than observed fluxes at midnight. We are now in the process of testing the Weimer (2005) electric field model incorporated into IMPTAM which depends on the IMF clock angle, IMF total field and components, V_{SW}, P_{SW}, and AL index. For magnetic field, several of the latest models, such as the TA15 (Tsyganenko & Andreeva, 2015) model and the RBF (Radial Basis Function) model (Andreeva & Tsyganenko, 2016; Tsyganenko & Andreeva, 2016) are now being considered.

The third reason is related to using the Tsyganenko and Mukai (2003) model for boundary conditions at 10 R_E in the plasma sheet. Limitations of Tsyganenko and Mukai (2003) applied for electrons are discussed in Dubyagin et al. (2016). The modeled fluxes are affected by the model's parameterization for plasma sheet density and temperature and its simple $sin^2(MLT)$ dependence. Dawn-dusk asymmetric terms are not included which sets the maximum location of density and temperature at around midnight. As for model parameters, for example, there will be an influx of electrons during both negative and positive IMF B_Z, and for positive IMF B_Z, the dependence is still proportional to the absolute value of B_Z. The distribution at the boundary fitted by the kappa shape with parameters as the electron number density and temperature in the plasma sheet which were obtained at distances between 6 and 11 R_E based on THEMIS data was given in Dubyagin et al. (2016) empirical model, usage of it will improve critically the IMPTAM outputs.

It needs to be mentioned that the version of IMPTAM used in the present paper did not include the effects from the substorm-associated electromagnetic fields. Substorms are a crucial factor in the transport and acceleration of keV electrons. Many satellite anomalies due to surface charging at geostationary orbit occur at night and early dawn (e.g., Fennell et al., 2001; O'Brien, 2009) where a hot plasma is injected from the magnetotail during substorms. Ganushkina et al. (2013, 2014), when modeling specific storm events, launched electromagnetic pulses given by Sarris et al. (2002) at each substorm onset determined from the AE index and scaled the amplitude according to the maximum values of the AE index. Addition of effects from substorms can influence the long-term IMPTAM performance.

All the statistical patterns for all three energies binned by MLT and IMPTAM driving parameters exhibit very similar shapes for the observed and modeled fluxes. The differences of one to two orders of magnitude are present, though. At the same time, the largest differences are mainly seen for such ranges of driving parameters when the number of datapoints (observed and modeled fluxes) is much less than 10% from the maximum number of points in a bin in that MLT range. For example, unrealistically high modeled fluxes were obtained at large (> 10 nT) positive IMF B_Z at around 18-06 MLT. If during our analysis we concentrate only at the ranges of IMPTAM driving parameters where the number of datapoints is statistically significant and disregard those which constitute less than 10%, the average difference will be about one order of magnitude. At the same time, as was mentioned above, the ratio between the observed and modeled fluxes at around 06 MLT is very close to one.

(Sillanp et al., 2017) conducted the analysis of GOES 13 MAGED data for five years (2011-2015) and developed an empirical model for the electron fluxes at geostationary orbit. They found that IMF B_z and solar wind speed V_{SW} with time delay of 1.5 hours were the parameters that produced the best correlation between the modeled and observed electron fluxes, so the model used those two driving parameters. Both parameters are the driving parameters in IMPTAM. The ratio between the modeled and the observed fluxes at 00-12 MLT is close to one (with upper value of up to 10) for IMF B_Z range of -10 to 10 nT which has most of the datapoints (Figure 2). The same is true for modeled fluxes corresponding to $V_{SW} < 700$ km/s: main over-estimation of about one order of magnitude is seen at around midnight (Figure 5). This reasonable agreement between the MAGED and IMPTAM fluxes is a valuable achievement for IMPTAM validation.

To evaluate the quality of the electron flux forecasts made by the IMPTAM, we employed the binary event analysis. The window width was set to one hour which is rather long, since the flux can vary significantly within an hour, but in the present study, the selected window is chosen to test "bulk activity". It was found that, in general, IMPTAM performs with the hit rate of 20-78% of threshold crossings reproduced depending on energy and threshold, see Table 1). The Heidke skill scores are rather low (0.159 at best for 40 keV electrons and negative values (-0.133) for 150 keV electrons) due to considerable false alarm rates (incorrect predictions during 14% to 50% of non-event times). The model is best at the 40 keV channel and for lower thresholds. This is very similar to that what was found in the previous study (Ganushkina et al., 2015), although some improvements are present for the 40 keV electrons. Three more metrics, namely, correlation, accuracy, and bias metrics, were used for the IMPTAM output compared to GOES 13 data. Overall correlation is rather weak and appears inversely proportional to electron energy. Median Symmetric Accuracy values demonstrate that the modeled fluxes are off by a factor of two (up to 3 for 150 keV electrons). Symmetric Signed Percentage Bias shows that the direction of the error varies with energy: the model overpredicts by 50% for 40 keV, underpredicts by 18% for 150 keV and overpredicts by almost 200% for 75 keV electrons. As was mentioned in Section 5, it is hard to expect the perfect performance of IMPTAM due to variations of several orders of magnitudes seen in keV electron fluxes which are strongly dependent on location and geomagnetic conditions. The main factors influencing the IMPTAM performance, especially at 150 keV, are the (1) boundary conditions were developed for ions but used here for electrons, (2) absence of substorm effects, (3) representations of electric and magnetic fields which can result in not enough adiabatic acceleration, and (4) effects from wave-particle interactions introduced as simple electron lifetimes. Ongoing work for IMPTAM improvement takes into account these factors. The Heidke skill scores are also influenced by the somewhat arbitrary selection of the thresholds for its calculation and the window width. The analysis conducted here provides insights into the representation of physical processes inside the IMPTAM. Special attention should be paid to these issues when improving IMPTAM in the future.

It needs to be stressed here that the analysis presented is for "nowcast" IMPTAM output, which is in contrast to "pastcast" when finalized, not real time driving parameters can be used and the IMPTAM setup can be varied to achieve the best fit to the data. The present study analyzes the IMPTAM output when it was run online in real time continuously, without introducing any changes into its structure and with the driving parameters always taken as real time parameters. It was a specific intention to present the IMPTAM performance on a sufficiently long time period without any interventions into its operation and without any "pastcast"-type approach.

Keeping in mind the points discussed above, the conclusions are the following:

1. The peaks for IMPTAM modeled fluxes are located not at around 06 MLT as for the observed GOES 13 MAGED fluxes but shifted towards midnight at all statistical patterns binned by MLT and IMPTAM driving parameters for all three energies.

2. All the statistical patterns for all three energies binned by MLT and IMPTAM driving parameters exhibit very similar features for the observed and modeled fluxes with the largest differences of about one order of magnitude. Differences of two orders of magnitude are seen for all IMPTAM parameters when the number of datapoints is less than 10% from the maximum number per bin. At the same time, the ratio between the observed and modeled fluxes at around 06 MLT is very close to one.

3. The IMF B_z and solar wind speed V_{SW} are the parameters which organize best the observed and modeled electron fluxes.

4. The applied metrics demonstrate that (a) in binary event analysis, 20-78% of threshold crossings are reproduced depending on energy and threshold but Heidke skill scores are not higher than 0.159 for 40 keV electrons and negative for 150 keV electrons due to incorrect predictions during 14% to 50% of non-event times; (b) the correlations are weak; (c) modeled fluxes are off by 200% (and up to 300% for 150 keV electrons) in terms of the median symmetric accuracy; and (d) symmetric signed percentagebias shows that the direction of the error varies with energy: overprediction by 50% (40 keV), overprediction by 200% (75 keV), underprediction by 18% (150 keV). Performance is best at the 40 keV level.

5. The revealed discrepancies are due to the models inside IMPTAM, such as (1) boundary conditions developed for ions but used for electrons, (2) absence of substorm effects, (3) representations of electric and magnetic fields which can result in not enough adiabatic acceleration, and (4) effects from wave-particle interactions introduced as simple electron lifetimes.

There is a further need to evaluate the model performance on larger data sets and with more appropriate metrics. The models like IMPTAM provide the information about the radiation environment which is vital and necessary to have in order to estimate the surface charging effects on satellites. When an anomaly occurs, the radiation environment may be more extreme than that given by the specification models used for design.

The existence of an operational model, fully validated and run in real time, is extremely important for determining the possible reason for that anomaly.

Acknowledgments

The projects leading to these results have received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement No 606716 SPACESTORM and from the European Union's Horizon 2020 research and innovation program under grant agreement No 637302 PROGRESS. N. Ganushkina thanks the International Space Science Institute in Bern, Switzerland, for their support of the international teams on "Analysis of Cluster Inner Magnetosphere Campaign data, in application of the dynamics of waves and wave-particle interaction within the outer radiation belt" and "Ring current modeling: Uncommon Assumptions and Common Misconceptions". The contribution by S. Dubyagin has been supported by the framework of the Finnish Centre of Excellence in Research of Sustainable Space (Academy of Finland decision numbers 312351 and 312390), which we gratefully acknowledge. The work at the University of Michigan was partly supported by the National Aeronautics and Space Administration under grant agreement NNX17AI48G and by the National Science Foundation under grant agreement NSF 1663770. The work at the University of Colorado was supported by the National Centers for Environmental Information under cooperative agreement NA17OAR4320101.

The IMPTAM output can be found at https://umich.box.com/s/axce0pxezdrudkwldfn5r6vwg92zylzz.

References

Andreeva, V. A., & Tsyganenko, N. A. (2016). Reconstructing the magnetosphere from data using radial basis functions. *Journal of Geophysical Research: Space Physics*, *121*(3), 2249-2263. doi: 10.1002/2015JA022242

Baker, D., Erickson, P. J., Fennell, J. F., Foster, J. C., Jaynes, A. N., & Verronen, P. T. (2018). Space Weather Effects in the Earths Radiation Belts. *Space Science Reviews*, *214*, 17. doi: 10.1007/s11214-017-0452-7

Balch, C. C. (2008). Updated verification of the Space Weather Prediction Center's solar energetic particle prediction model. *Space Weather*, *6*(1). doi: 10.1029/2007SW000337

Balikhin, M. A., Boynton, R. J., Walker, S. N., Borovsky, J. E., Billings, S. A., & Wei, H. L. (2011). Using the NARMAX approach to model the evolution of energetic electrons fluxes at geostationary orbit. *Geophysical Research Letters*, *38*(18). doi: 10.1029/2011GL048980

Balikhin, M. A., Rodriguez, J. V., Boynton, R. J., Walker, S. N., Aryan, H., Sibeck, D. G., & Billings, S. A. (2016). Comparative analysis of NOAA REFM and SNB3GEO tools for the forecast of the fluxes of highenergy electrons at GEO. *Space Weather*, *14*, 22-31. doi: 10.1002/2015SW001303

Boyd, A. J., Spence, H. E., Huang, C.-L., Reeves, G. D., Baker, D. N., Turner, D. L., ... Shprits, Y. Y. (2016). Statistical properties of the radiation belt seed population. *Journal of Geophysical Research: Space Physics*, *121*(8), 7636-7646. doi: 10.1002/2016JA022652

Boyle, C. B., Reiff, P. H., & Hairston, M. R. (1997). Empirical polar cap potentials. *Journal of Geophysical Research: Space Physics*, *102*(A1), 111-125. doi: 10.1029/96JA01742

Boynton, R. J., Balikhin, M. A., Sibeck, D. G., Walker, S. N., Billings, S. A., & Ganushkina, N. (2016). Electron flux models for different energies at geostationary orbit. *Space Weather*, *14*(10), 846-860. doi: 10.1002/2016SW001506

Brautigam, D. H., & Albert, J. M. (2000). Radial diffusion analysis of outer radiation belt electrons during the October 9, 1990, magnetic storm. *Journal of Geophysical Research: Space Physics*, *105*(A1), 291-309. doi:

10.1029/1999JA900344

Chen, M. W., Schulz, M., Anderson, P. C., Lu, G., Germany, G., & West, M. (2005). Storm time distributions of diffuse auroral electron energy and X-ray flux: Comparison of drift-loss simulations with observations. *Journal of Geophysical Research: Space Physics*, *110*(A3). doi: 10.1029/2004JA010725

Chen, M. W., Schulz, M., Lu, G., & Lyons, L. R. (2003). Quasi-steady drift paths in a model magnetosphere with AMIE electric field: Implications for ring current formation. *Journal of Geophysical Research: Space Physics*, *108*(A5). doi: 10.1029/2002JA009584

Chen, Y., Friedel, R. H. W., & Reeves, G. D. (2006). Phase space density distributions of energetic electrons in the outer radiation belt during two Geospace Environment Modeling Inner Magnetosphere/Storms selected storms. *Journal of Geophysical Research: Space Physics*, *111*(A11). doi: 10.1029/2006JA011703

Chen, Y., Reeves, G. D., & Friedel, R. H. W. (2007). The energization of relativistic electrons in the outer Van Allen radiation belt. *Nature Physics*, *3*, 614-617. doi: 10.1038/nphys655

Davis, V. A., Mandell, M. J., & Thomsen, M. F. (2008). Representation of the measured geosynchronous plasma environment in spacecraft charging calculations. *Journal of Geophysical Research: Space Physics*, *113*(A10). doi: 10.1029/2008JA013116

Denton, M. H., Henderson, M. G., Jordanova, V. K., Thomsen, M. F., Borovsky, J. E., Woodroffe, J., . . . Pitchford, D. (2016). An improved empirical model of electron and ion fluxes at geosynchronous orbit based on upstream solar wind conditions. *Space Weather*, *14*(7), 511-523. doi: 10.1002/2016SW001409

Denton, M. H., Reeves, G. D., Larsen, B. A., Friedel, R. H. W., Thomsen, M. F., Fernandes, P. A., . . . Sarno-Smith, L. K. (2017). On the origin of low-energy electrons in the inner magnetosphere: Fluxes and pitch-angle distributions. *Journal of Geophysical Research: Space Physics*, *122*(2), 1789-1802. doi: 10.1002/2016JA023648

Denton, M. H., Thomsen, M. F., Jordanova, V. K., Henderson, M. G., Borovsky, J. E., Denton, J. S., . . . Hartley, D. P. (2015). An empirical model of electron and ion fluxes derived from observations at geosynchronous orbit. *Space Weather*, *13*(4), 233-249. doi: 10.1002/2015SW001168

Doswell III, C. A., Davies-Jones, R., & Keller, D. L. (1990). On summary measures of skill in rare event forecasting based on contingency tables. *Weather Forecasting*, *5*, 576-585. doi: 10.1175/1520-0434(1990)005⟨0576:OSMOSI⟩2.0.CO;2

Dubyagin, S., Ganushkina, N. Y., Sillanp, I., & Runov, A. (2016). Solar wind-driven variations of electron plasma sheet densities and temperatures beyond geostationary orbit during storm times. *Journal of Geophysical Research: Space Physics*, *121*(9), 8343-8360. doi: 10.1002/2016JA022947

Fennell, J. F., Koons, H. C., Roeder, J. L., & Blake, J. B. (2001). *Spacecraft Charging: Observations and Relationship to Satellite Anomalies* (Vol. 476). European Space Agency, Noordwijk, Netherlands.

Fok, M.-C., Buzulukova, N. Y., Chen, S.-H., Glocer, A., Nagai, T., Valek, P., & Perez, J. D. (2014). The Comprehensive Inner Magnetosphere-Ionosphere Model. *Journal of Geophysical Research: Space Physics*, *119*(9), 7522-7540. doi: 10.1002/2014JA020239

Fok, M.-C., Glocer, A., Zheng, Q., Horne, R. B., Meredith, N. P., Albert, J. M., & Nagai, T. (2011). Recent developments in the radiation belt environment model. *Journal of Atmospheric and Solar-Terrestrial Physics*, *73*(11), 1435-1443. doi: https://doi.org/10.1016/j.jastp.2010.09.033

Fok, M.-C., & Moore, T. E. (1997). Ring current modeling in a realistic magnetic field configuration. *Geophysical Research Letters*, *24*(14), 1775-1778. doi: 10.1029/97GL01255

Fok, M.-C., Moore, T. E., & Delcourt, D. C. (1999). Modeling of inner plasma sheet and ring current during substorms. *Journal of Geophysical Research: Space Physics*, *104*(A7), 14557-14569. doi: 10.1029/1999JA900014

Fok, M.-C., Moore, T. E., Wilson, G. R., Perez, J. D., Zhang, X. X., Brandt, P. C., ... Wolf, R. A. (2003). Global ENA Image Simulations. *Space Science Reviews*, *109*(1), 77-103. doi: 10.1023/B:SPAC.0000007514.56380.fd

Fok, M.-C., Wolf, R. A., Spiro, R. W., & Moore, T. E. (2001). Comprehensive computational model of Earth's ring current. *Journal of Geophysical Research: Space Physics*, *106*(A5), 8417-8424. doi: 10.1029/2000JA000235

Foster, J. C., Erickson, P. J., Omura, Y., Baker, D. N., Kletzing, C. A., & Claudepierre, S. G. (2017). Van Allen Probes observations of prompt mev radiation belt electron acceleration in nonlinear interactions with VLF chorus. *Journal of Geophysical Research: Space Physics*, *122*(1), 324-339. doi: 10.1002/2016JA023429

Friedel, R. H. W., Korth, H., Henderson, M. G., Thomsen, M. F., & Scudder, J. D. (2001). Plasma sheet access to the inner magnetosphere. *Journal of Geophysical Research: Space Physics*, *106*(A4), 5845-5858. doi: 10.1029/2000JA003011

Ganushkina, N. Y., Amariutei, O. A., Shprits, Y. Y., & Liemohn, M. W. (2013). Transport of the plasma sheet electrons to the geostationary distances. *Journal of Geophysical Research: Space Physics*, *118*(1), 82-98. doi: 10.1029/2012JA017923

Ganushkina, N. Y., Amariutei, O. A., Welling, D., & Heynderickx, D. (2015). Nowcast model for low-energy electrons in the inner magnetosphere. *Space Weather*, *13*(1), 16-34. doi: 10.1002/2014SW001098

Ganushkina, N. Y., Liemohn, M. W., Amariutei, O. A., & Pitchford, D. (2014). Low-energy electrons (550 keV) in the inner magnetosphere. *Journal of Geophysical Research: Space Physics*, *119*(1), 246-259. doi: 10.1002/2013JA019304

Garrett, H. B. (1981). The charging of spacecraft surfaces. *Reviews of Geophysics*, *19*(4), 577-616. doi: 10.1029/RG019i004p00577

Glauert, S. A., Horne, R. B., & Meredith, N. P. (2014). Three-dimensional electron radiation belt simulations using the BAS Radiation Belt Model with new diffusion models for chorus, plasmaspheric hiss, and lightning-generated whistlers. *Journal of Geophysical Research: Space Physics*, *119*(1), 268-289. doi: 10.1002/2013JA019281

Green, J. C., & Kivelson, M. G. (2001). A tale of two theories: How the adiabatic response and ulf waves affect relativistic electrons. *Journal of Geophysical Research: Space Physics*, *106*(A11), 25777-25791. doi: 10.1029/2001JA000054

Green, J. C., & Kivelson, M. G. (2004). Relativistic electrons in the outer radiation belt: Differentiating between acceleration mechanisms. *Journal of Geophysical Research: Space Physics*, *109*(A3). doi: 10.1029/2003JA010153

Hanser, F. A. (2011). *EPS/HEPAD calibration and data handbook* (Tech. Rep. GOESN-ENG-048D). Assurance Technology Corporation, Carlisle, Mass. Retrieved from `http://www.ngdc.noaa.gov/stp/satellite/goes/documentation.html`

Hartley, D. P., Denton, M. H., Green, J. C., Onsager, T. G., Rodriguez, J. V., & Singer, H. J. (2013). Case studies of the impact of high-speed solar wind streams on the electron radiation belt at geosynchronous orbit: Flux, magnetic field, and phase space density. *Journal of Geophysical Research: Space Physics*, *118*(11), 6964-6979. doi: 10.1002/2013JA018923

Heidke, P. (1926). Berechnung Des Erfolges Und Der Gte Der Windstrkevorhersagen Im Sturmwarnungsdienst. *Geografiska AnnalerNature*, *8*, 301-349. doi: 10.1080/20014422.1926.11881138

Horne, R. B., Thorne, R. M., Shprits, Y. Y., Meredith, N. P., Glauert, S. A., Smith,

A. J., ... Decreau, P. M. E. (2005). Wave acceleration of electrons in the Van Allen radiation belts. *Nature*, *437*, 227-230. doi: doi:10.1038/nature03939

Iucci, N., Levitin, A. E., Belov, A. V., Eroshenko, E. A., Ptitsyna, N. G., Villoresi, G., ... G., Y. V. (2005). Space weather conditions and spacecraft anomalies in different orbits. *Space Weather*, *3*, S01001. doi: doi:10.1029/2003SW000056

Jaynes, A. N., Baker, D. N., Singer, H. J., Rodriguez, J. V., Loto'aniu, T. M., Ali, A. F., ... Reeves, G. D. (2015). Source and seed populations for relativistic electrons: Their roles in radiation belt changes. *Journal of Geophysical Research: Space Physics*, *120*(9), 7240-7254. doi: 10.1002/2015JA021234

Jolliffe, I. T., & Stephenson, D. B. (Eds.). (2012). *Forecast verification, A practinioner's guide in atmospheric science* (Second ed.). Wiley.

Jordanova, V. K., & Miyoshi, Y. (2005). Relativistic model of ring current and radiation belt ions and electrons: Initial results. *Geophysical Research Letters*, *32*(14). doi: 10.1029/2005GL023020

Jordanova, V. K., Yu, Y., Niehof, J. T., Skoug, R. M., Reeves, G. D., Kletzing, C. A., ... Spence, H. E. (2014). Simulations of inner magnetosphere dynamics with an expanded RAM-SCB model and comparisons with Van Allen Probes observations. *Geophysical Research Letters*, *41*(8), 2687-2694. doi: 10.1002/2014GL059533

Katus, R. M., & Liemohn, M. W. (2013). Similarities and differences in low- to middle-latitude geomagnetic indices. *Journal of Geophysical Research: Space Physics*, *118*(8), 5149-5156. doi: 10.1002/jgra.50501

Kennel, C. F., & Petschek, H. E. (1966). Limit on stably trapped particle fluxes. *Journal of Geophysical Research*, *71*(1), 1-28. doi: 10.1029/JZ071i001p00001

Kennel, C. F., & Thorne, R. M. (1967). Unstable growth of unducted whistlers propagating at an angle to the geomagnetic field. *Journal of Geophysical Research*, *72*(3), 871-878. doi: 10.1029/JZ072i003p00871

Kerns, K. J., Hardy, D. A., & Gussenhoven, M. S. (1994). Modeling of convection boundaries seen by CRRES in 120-ev to 28-keV particles. *Journal of Geophysical Research: Space Physics*, *99*(A2), 2403-2414. doi: 10.1029/93JA02767

Korth, H., Thomsen, M. F., Borovsky, J. E., & McComas, D. J. (1999). Plasma sheet access to geosynchronous orbit. *Journal of Geophysical Research: Space Physics*, *104*(A11), 25047-25061. doi: 10.1029/1999JA900292

Lanzerotti, L. J., LaFleur, K., Maclennan, C. G., & Maurer, D. W. (1998). Geosynchronous spacecraft charging in January 1997. *Geophysical Research Letters*, *25*(15), 2967-2970. doi: 10.1029/98GL00987

Leontaritis, I. J., & Billings, S. A. (1985a). Input-output parametric models for non-linear systems Part I: deterministic non-linear systems. *International Journal of Control*, *41*(2), 303-328. doi: 10.1080/0020718508961129

Leontaritis, I. J., & Billings, S. A. (1985b). Input-output parametric models for non-linear systems Part II: stochastic non-linear systems. *International Journal of Control*, *41*(2), 329-344. doi: 10.1080/0020718508961130

Li, W., Thorne, R. M., Ma, Q., Ni, B., Bortnik, J., Baker, D. N., ... Claudepierre, S. G. (2014). Radiation belt electron acceleration by chorus waves during the 17 March 2013 storm. *Journal of Geophysical Research: Space Physics*, *119*(6), 4681-4693. doi: 10.1002/2014JA019945

Liemohn, M. W., Khazanov, G. V., & Kozyra, J. U. (1998). Banded electron structure formation in the inner magnetosphere. *Geophysical Research Letters*, *25*(6), 877-880. doi: 10.1029/98GL00411

Liemohn, M. W., Ridley, A. J., Gallagher, D. L., Ober, D. M., & Kozyra, J. U. (2004). Dependence of plasmaspheric morphology on the electric field description during the recovery phase of the 17 April 2002 magnetic storm. *Journal of Geophysical Research: Space Physics*, *109*(A3). doi: 10.1029/2003JA010304

Lopez, R. E., Hernandez, S., Wiltberger, M., Huang, C.-L., Kepko, E. L., Spence, H., ... Lyon, J. G. (2007). Predicting magnetopause crossings at geosynchronous orbit during the Halloween storms. *Space Weather*, *5*(1). doi: 10.1029/2006SW000222

Mato-Vlez, J.-C., Sicard, A., Payan, D., Ganushkina, N., Meredith, N. P., & Sillanpa, I. (2018). Spacecraft surface charging induced by severe environments at geosynchronous orbit. *Space Weather*, *16*(1), 89-106. doi: 10.1002/2017SW001689

Mauk, B. H., & Meng, C.-I. (1983). Characterization of geostationary particle signatures based on the "Injection Boundary" Model. *Journal of Geophysical Research: Space Physics*, *88*(A4), 3055-3071. doi: 10.1029/JA088iA04p03055

Miyoshi, Y. S., Jordanova, V. K., Morioka, A., Thomsen, M. F., Reeves, G. D., Evans, D. S., & Green, J. C. (2006). Observations and modeling of energetic electron dynamics during the October 2001 storm. *Journal of Geophysical Research: Space Physics*, *111*(A11). doi: 10.1029/2005JA011351

Morley, S. K. (2016). *Alternatives to accuracy and bias metrics based on percentage errors for radiation belt modeling applications* (Tech. Rep. LA-UR-16-24592). Los Alamos, NM 87545, USA: Los Alamos National Laboratory. doi: doi:10.2172/1260362

Morley, S. K., Brito, T. V., & Welling, D. T. (2018). Measures of Model Performance Based On the Log Accuracy Ratio. *Space Weather*, *16*(1), 69-88. doi: 10.1002/2017SW001669

Murphy, A. H. (1993). What Is a Good Forecast? An Essay on the Nature of Goodness in Weather Forecasting. *Weather and Forecasting*, *8*(2), 281-293. doi: 10.1175/1520-0434(1993)008⟨0281:WIAGFA⟩2.0.CO;2

O'Brien, T. P. (2009). SEAES-GEO: A spacecraft environmental anomalies expert system for geosynchronous orbit. *Space Weather*, *7*(9). doi: 10.1029/2009SW000473

Orlova, K., & Shprits, Y. (2014). Model of lifetimes of the outer radiation belt electrons in a realistic magnetic field using realistic chorus wave parameters. *Journal of Geophysical Research: Space Physics*, *119*(2), 770-780. doi: 10.1002/2013JA019596

Orlova, K., Shprits, Y., & Spasojevic, M. (2016). New global loss model of energetic and relativistic electrons based on Van Allen Probes measurements. *Journal of Geophysical Research: Space Physics*, *121*(2), 1308-1314. doi: 10.1002/2015JA021878

Pulkkinen, A., Rasttter, L., Kuznetsova, M., Singer, H., Balch, C., Weimer, D., ... Weigel, R. (2013). Community-wide validation of geospace model ground magnetic field perturbation predictions to support model transition to operations. *Space Weather*, *11*(6), 369-385. doi: 10.1002/swe.20056

Reeves, G. D., Morley, S. K., Friedel, R. H. W., Henderson, M. G., Cayton, T. E., Cunningham, G., ... Thomsen, D. (2011). On the relationship between relativistic electron flux and solar wind velocity: Paulikas and Blake revisited. *Journal of Geophysical Research: Space Physics*, *116*(A2), A02213. doi: 10.1029/2010JA015735

Ridley, A. J., & Liemohn, M. W. (2002). A model-derived storm time asymmetric ring current driven electric field description. *Journal of Geophysical Research: Space Physics*, *107*(A8), SMP 2-1-SMP 2-12. doi: 10.1029/2001JA000051

Roberts, C. S. (1965). On the relationship between the unidirectional and omnidirectional flux of trapped particles on a magnetic line of force. *Journal of Geophysical Research*, *70*(11), 2517-2527. doi: 10.1029/JZ070i011p02517

Rodriguez, J. V. (2014). *GOES 13-15 MAGE/PD pitch angles* (Algorithm Theoretical Basis Document, version 1.0). NOAA NESDIS NGDC. Retrieved from `http://www.ngdc.noaa.gov/stp/satellite/goes/documentation.html`

Rowland, W., & Weigel, R. S. (2012). Intracalibration of particle detectors on a three-axis stabilized geostationary platform. *Space Weather*, *10*(11). doi: 10.1029/2012SW000816

Sarris, T. E., Li, X., Tsaggas, N., & Paschalidis, N. (2002). Modeling energetic particle injections in dynamic pulse fields with varying propagation speeds. *Journal of Geophysical Research: Space Physics*, *107*(A3), SMP 1-1-SMP 1-10. doi: 10.1029/2001JA900166

Schulz, M., & Lanzerotti, L. J. (1974). *Particle Diffusion in the Radiation Belts* (Vol. 7). Springer-Verlag, New York.

Shprits, Y. Y., Meredith, N. P., & Thorne, R. M. (2007). Parameterization of radiation belt electron loss timescales due to interactions with chorus waves. *Geophysical Research Letters*, *34*(11). doi: 10.1029/2006GL029050

Shprits, Y. Y., Thorne, R. M., Horne, R. B., & Summers, D. (2006). Bounce-averaged diffusion coefficients for field-aligned chorus waves. *Journal of Geophysical Research: Space Physics*, *111*(A10). doi: 10.1029/2006JA011725

Sillanp, I., Ganushkina, N. Y., Dubyagin, S., & Rodriguez, J. V. (2017). Electron Fluxes at Geostationary Orbit From GOES MAGED Data. *Space Weather*, *15*(12), 1602-1614. doi: 10.1002/2017SW001698

Subbotin, D. A., & Shprits, Y. Y. (2009). Three-dimensional modeling of the radiation belts using the Versatile Electron Radiation Belt (VERB) code. *Space Weather*, *7*(10). doi: 10.1029/2008SW000452

Thomsen, M. F., Henderson, M. G., & Jordanova, V. K. (2013). Statistical properties of the surface-charging environment at geosynchronous orbit. *Space Weather*, *11*(5), 237-244. doi: 10.1002/swe.20049

Thornes, J. E., & Stephenson, D. B. (2001). How to judge the quality and value of weather forecast products. *Meteorological Applications*, *8*(3), 307-314. doi: 10.1017/S1350482701003061

Tofallis, C. (2015). A better measure of relative prediction accuracy for model selection and model estimation. *Journal of the Operational Research Society*, *66*(8), 1352-1362. doi: 10.1057/jors.2014.103

Tsyganenko, N. A. (1995). Modeling the Earth's magnetospheric magnetic field confined within a realistic magnetopause. *Journal of Geophysical Research: Space Physics*, *100*(A4), 5599-5612. doi: 10.1029/94JA03193

Tsyganenko, N. A., & Andreeva, V. A. (2015). A forecasting model of the magnetosphere driven by an optimal solar wind coupling function. *Journal of Geophysical Research: Space Physics*, *120*(10), 8401-8425. doi: 10.1002/2015JA021641

Tsyganenko, N. A., & Andreeva, V. A. (2016). An empirical RBF model of the magnetosphere parameterized by interplanetary and ground-based drivers. *Journal of Geophysical Research: Space Physics*, *121*(11), 10,786-10,802. doi: 10.1002/2016JA023217

Tsyganenko, N. A., & Mukai, T. (2003). Tail plasma sheet models derived from Geotail particle data. *Journal of Geophysical Research: Space Physics*, *108*(A3). doi: 10.1029/2002JA009707

Usanova, M. E., Drozdov, A., Orlova, K., Mann, I. R., Shprits, Y., Robertson, M. T., ... Wygant, J. (2014). Effect of EMIC waves on relativistic and ultrarelativistic electron populations: Ground-based and Van Allen Probes observations. *Geophysical Research Letters*, *41*(5), 1375-1381. doi: 10.1002/2013GL059024

Walther, B. A., & Moore, J. L. (2005). The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. *Ecography*, *28*(6), 815-829. doi: 10.1111/j.2005.0906-7590.04112.x

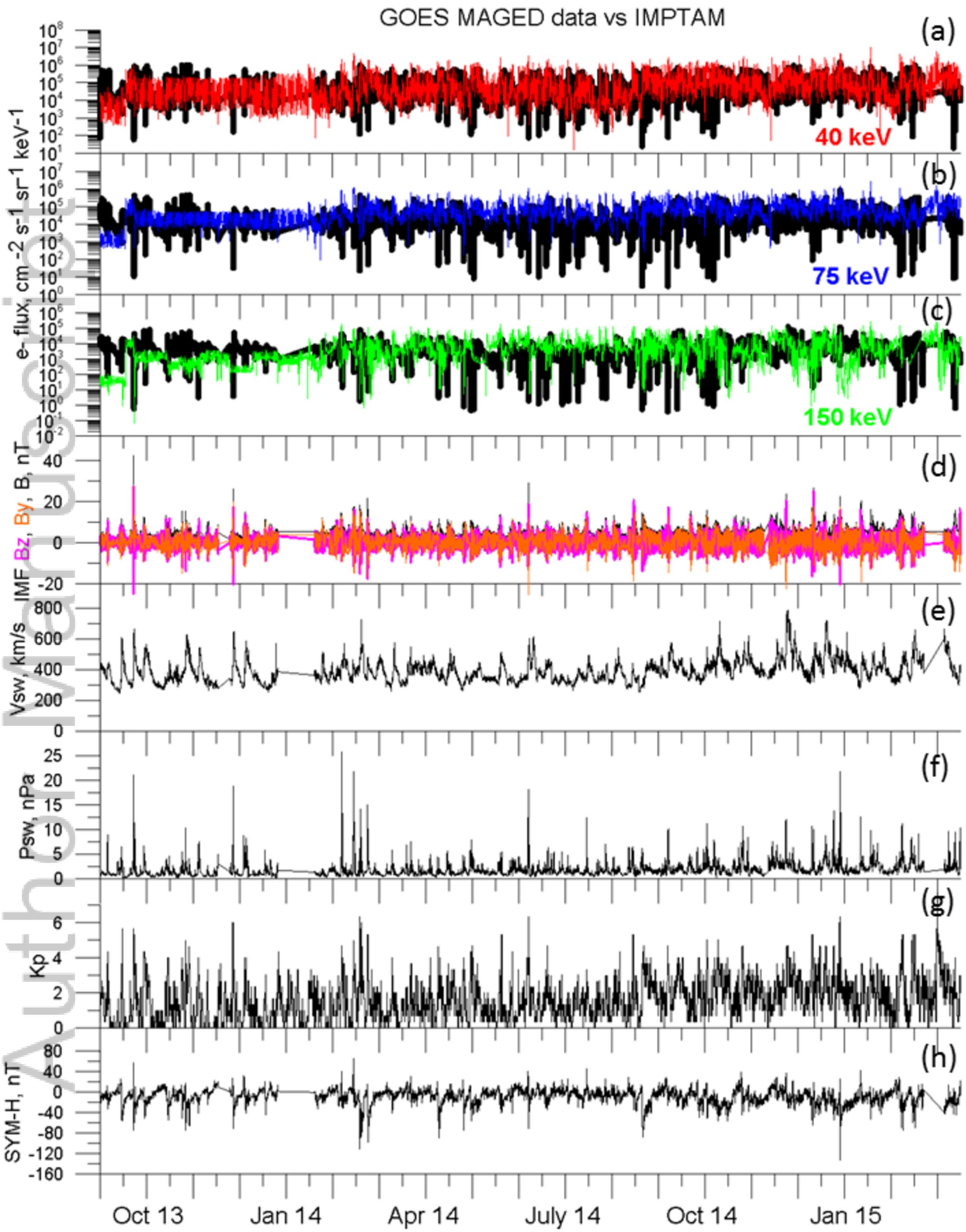
Wanliss, J. A., & Showalter, K. M. (2006). High-resolution global storm index: Dst versus SYM-H. *Journal of Geophysical Research: Space Physics*, *111*(A2). doi: 10.1029/2005JA011034

Weimer, D. R. (2005). Improved ionospheric electrodynamic models and application to calculating Joule heating rates. *Journal of Geophysical Research: Space Physics*, *110*(A5). doi: 10.1029/2004JA010884

Welling, D. T., & Ridley, A. J. (2010). Validation of SWMF magnetic field and plasma. *Space Weather*, *8*(3). doi: 10.1029/2009SW000494

Wilks, D. S. (2006). *Statistical methods in the atmospheric sciences* (Second ed.). Academic Press.

Figure 1.



GOES MAGED data vs IMPTAM
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
40 keV
75 keV
150 keV
e- flux, cm-2 s-1 sr-1 keV-1
IMF Bz, By, B, nT
Vsw, km/s
Psw, nPa
Kp
SYM-H, nT
Oct 13
Jan 14
Apr 14
July 14
Oct 14
Jan 15

Figure 2.

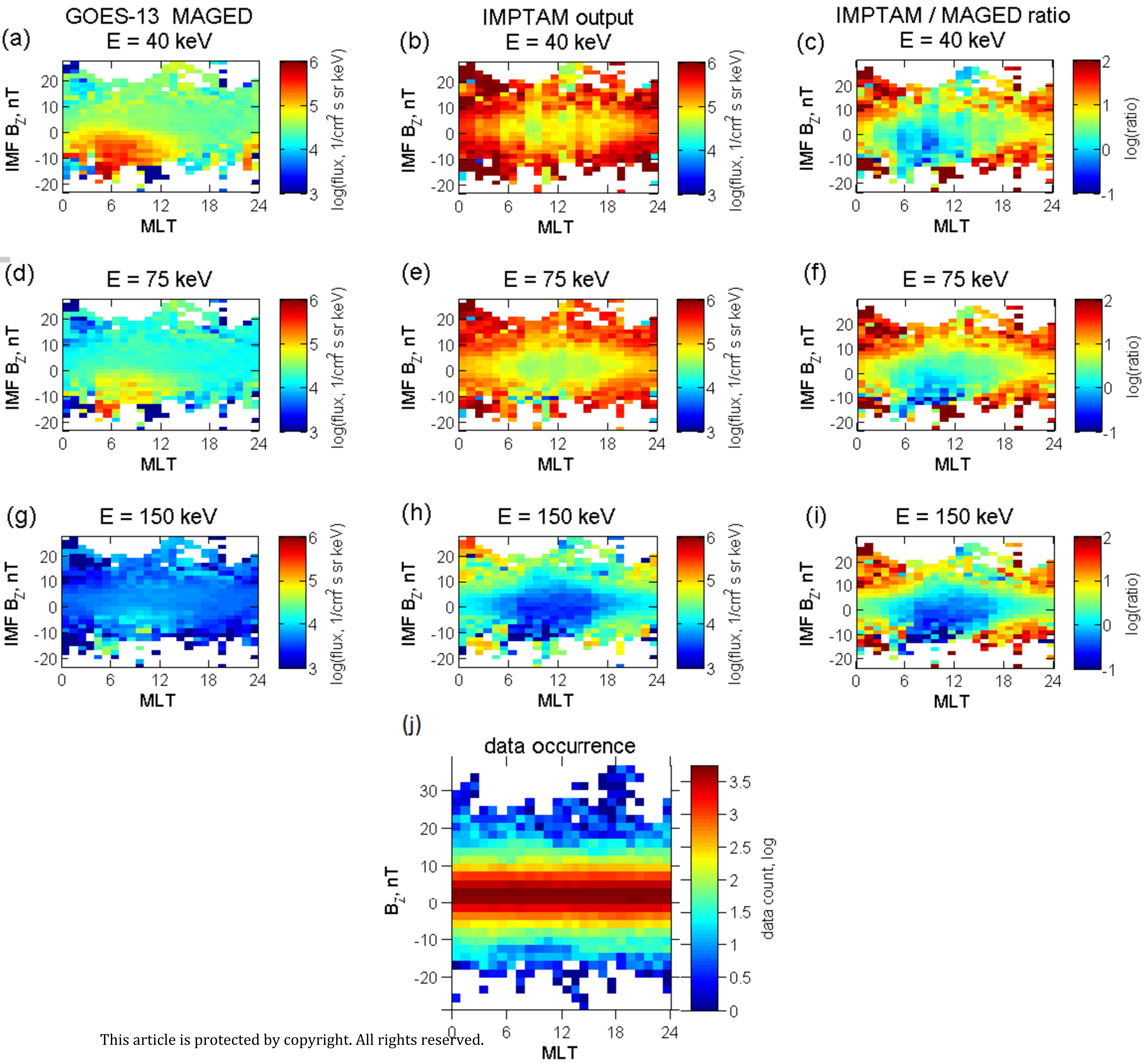

GOES-13 MAGED
E = 40 keV
IMPTAM output
E = 40 keV
IMPTAM / MAGED ratio
E = 40 keV
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
E = 75 keV
E = 150 keV
IMF B_Z, nT
MLT
log(flux, 1/cm^2 s sr keV)
log(ratio)
data occurrence
B_Z, nT
data count, log

Figure 3.

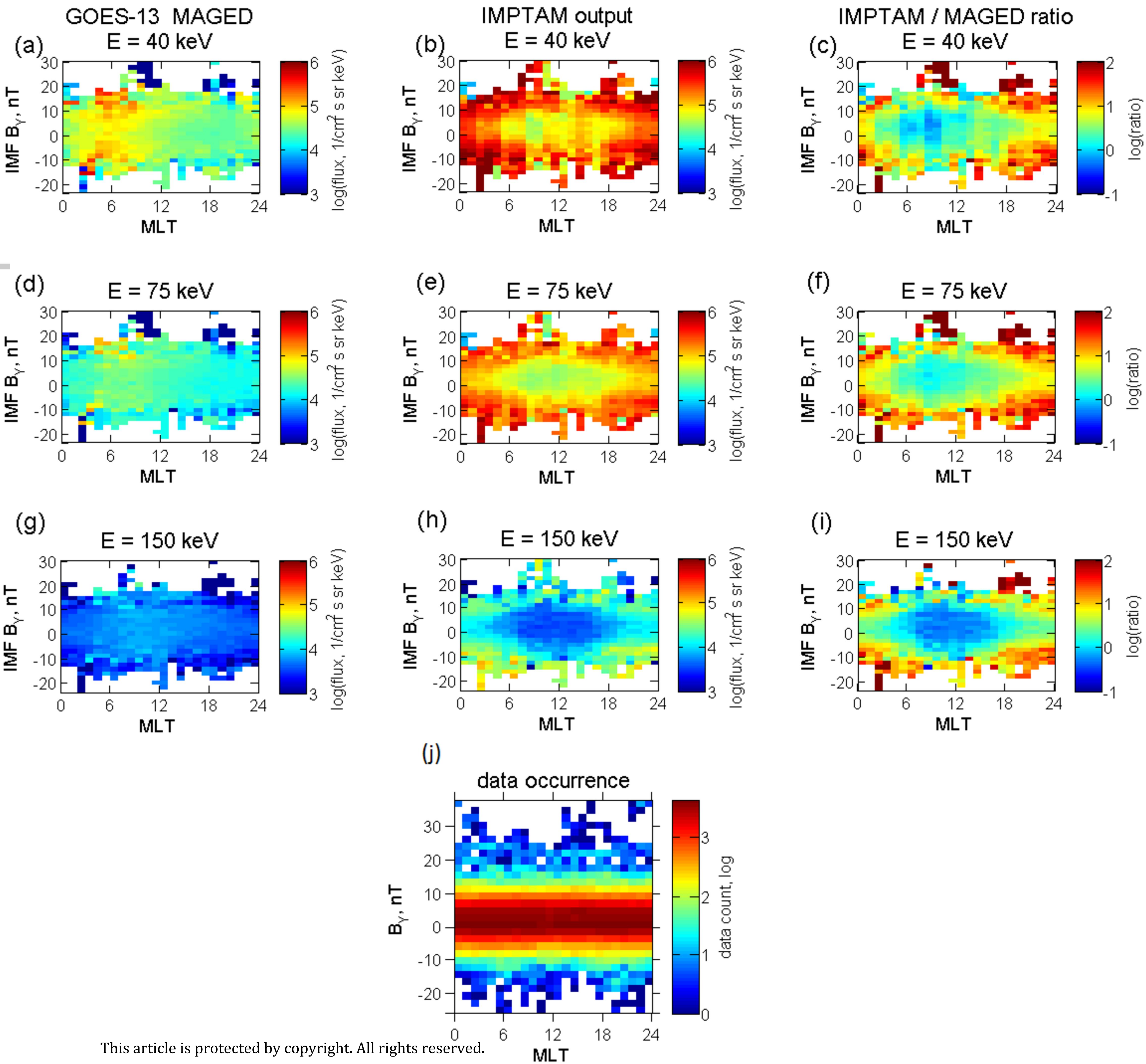

GOES-13 MAGED
IMPTAM output
IMPTAM / MAGED ratio
(a)
E = 40 keV
(b)
E = 40 keV
(c)
E = 40 keV
(d)
E = 75 keV
(e)
E = 75 keV
(f)
E = 75 keV
(g)
E = 150 keV
(h)
E = 150 keV
(i)
E = 150 keV
IMF B$_Y$, nT
MLT
log(flux, 1/cm^2 s sr keV)
log(ratio)
(j)
data occurrence
B$_Y$, nT
data count, log

Figure 4.

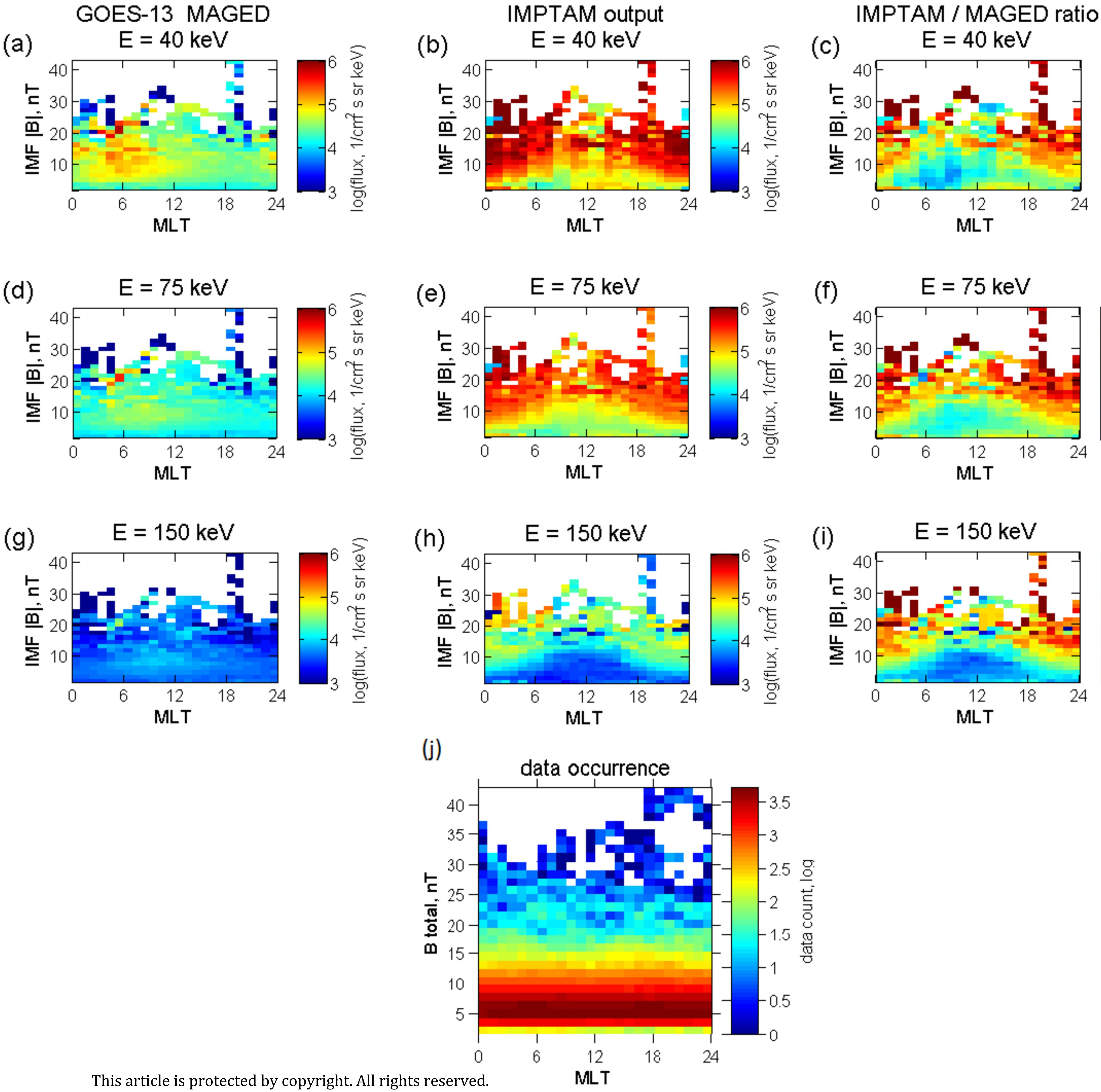

GOES-13 MAGED
IMPTAM output
IMPTAM / MAGED ratio
(a) E = 40 keV
(b) E = 40 keV
(c) E = 40 keV
(d) E = 75 keV
(e) E = 75 keV
(f) E = 75 keV
(g) E = 150 keV
(h) E = 150 keV
(i) E = 150 keV
(j) data occurrence
IMF |B|, nT
MLT
log(flux, 1/cm² s sr keV)
log(ratio)
B total, nT
data count, log

Figure 5.

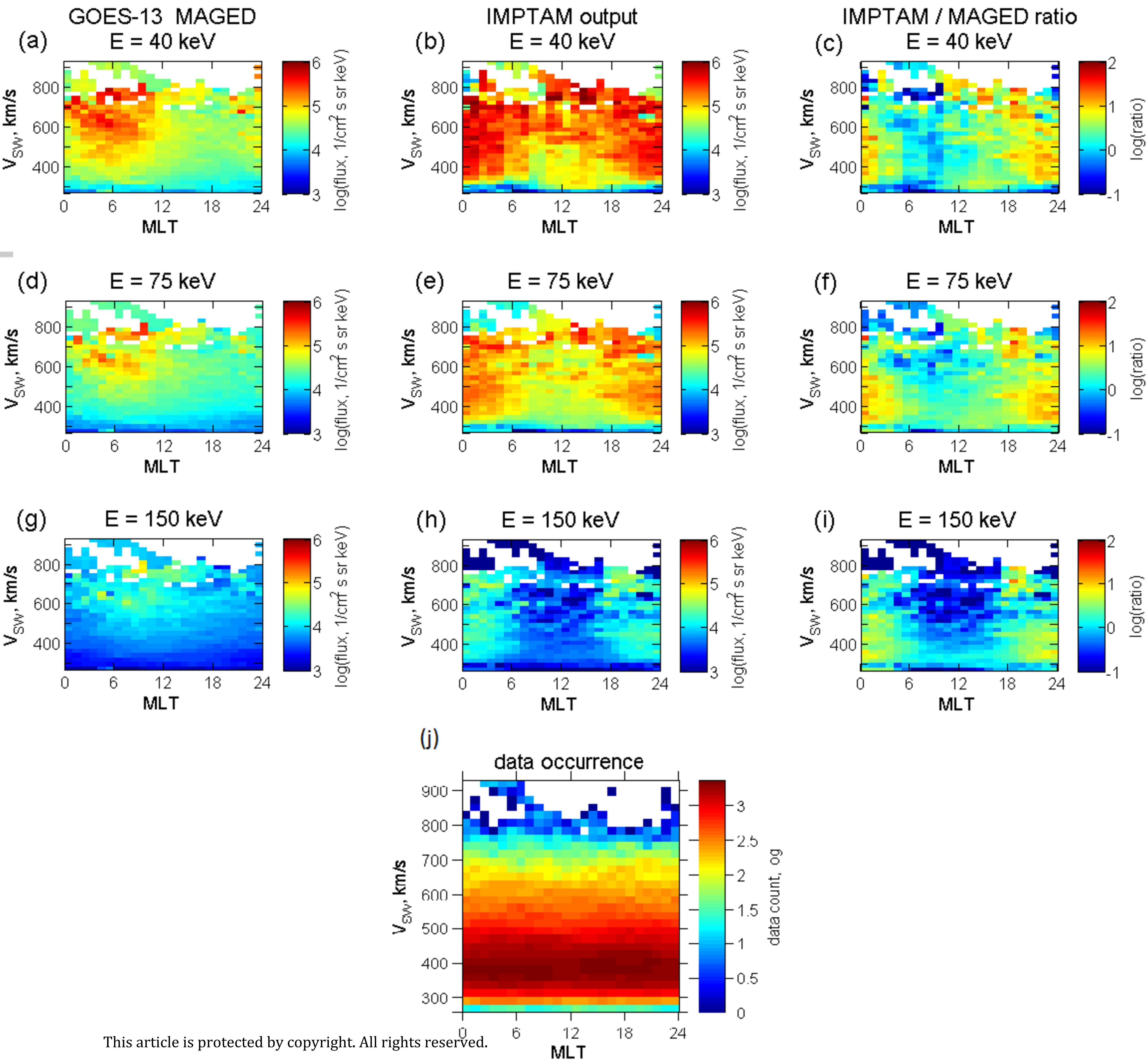

GOES-13 MAGED
IMPTAM output
IMPTAM / MAGED ratio
(a)
E = 40 keV
(b)
E = 40 keV
(c)
E = 40 keV
(d)
E = 75 keV
(e)
E = 75 keV
(f)
E = 75 keV
(g)
E = 150 keV
(h)
E = 150 keV
(i)
E = 150 keV
(j)
data occurrence
V_SW, km/s
MLT
log(flux, 1/cm² s sr keV)
log(ratio)
data count, og
0
6
12
18
24
400
600
800
300
500
700
900

Figure 6.

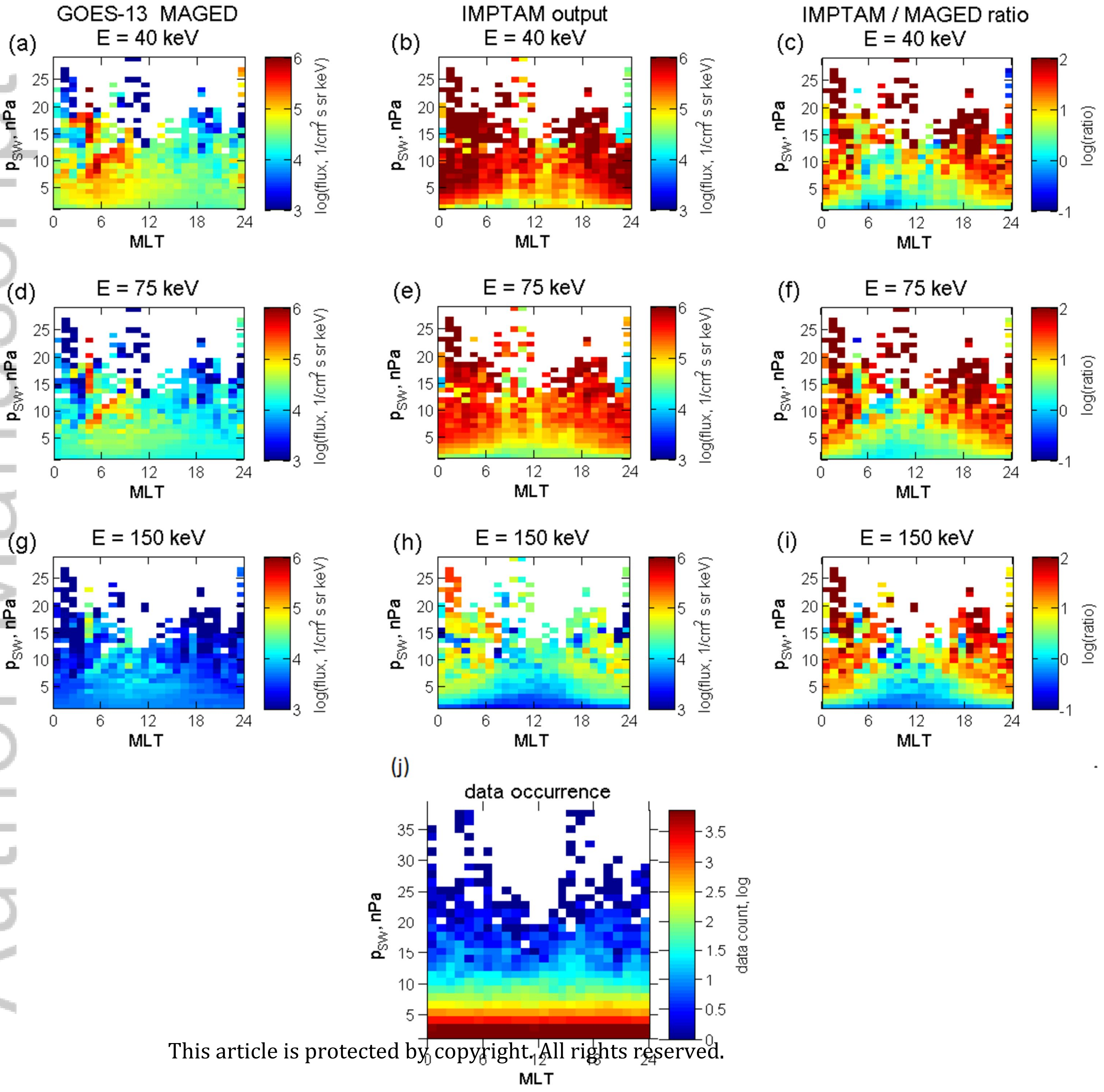

GOES-13 MAGED
IMPTAM output
IMPTAM / MAGED ratio
(a) E = 40 keV
(b) E = 40 keV
(c) E = 40 keV
(d) E = 75 keV
(e) E = 75 keV
(f) E = 75 keV
(g) E = 150 keV
(h) E = 150 keV
(i) E = 150 keV
(j) data occurrence
p_SW, nPa
MLT
log(flux, 1/cm² s sr keV)
log(ratio)
data count, log

Figure 7.

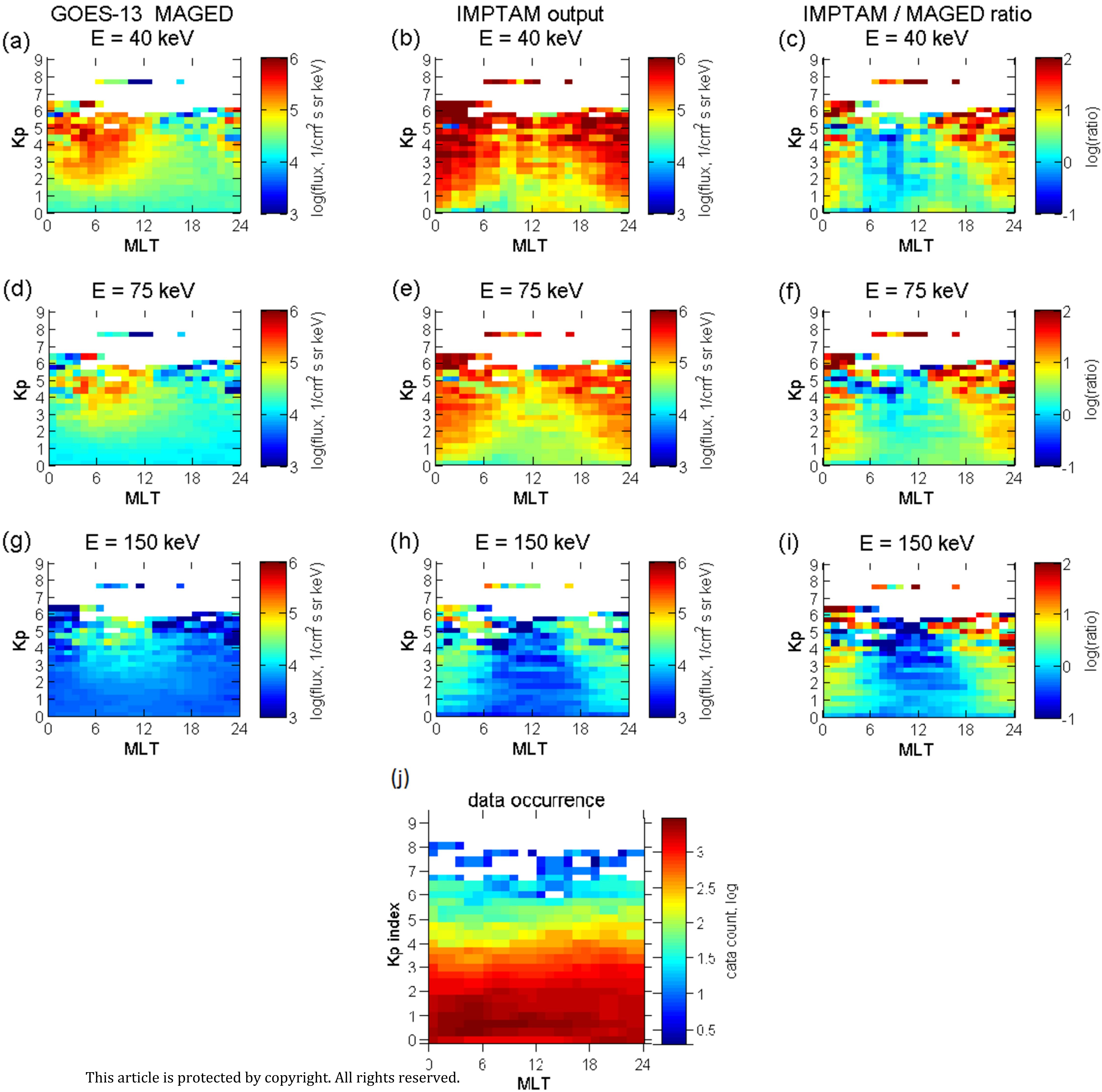
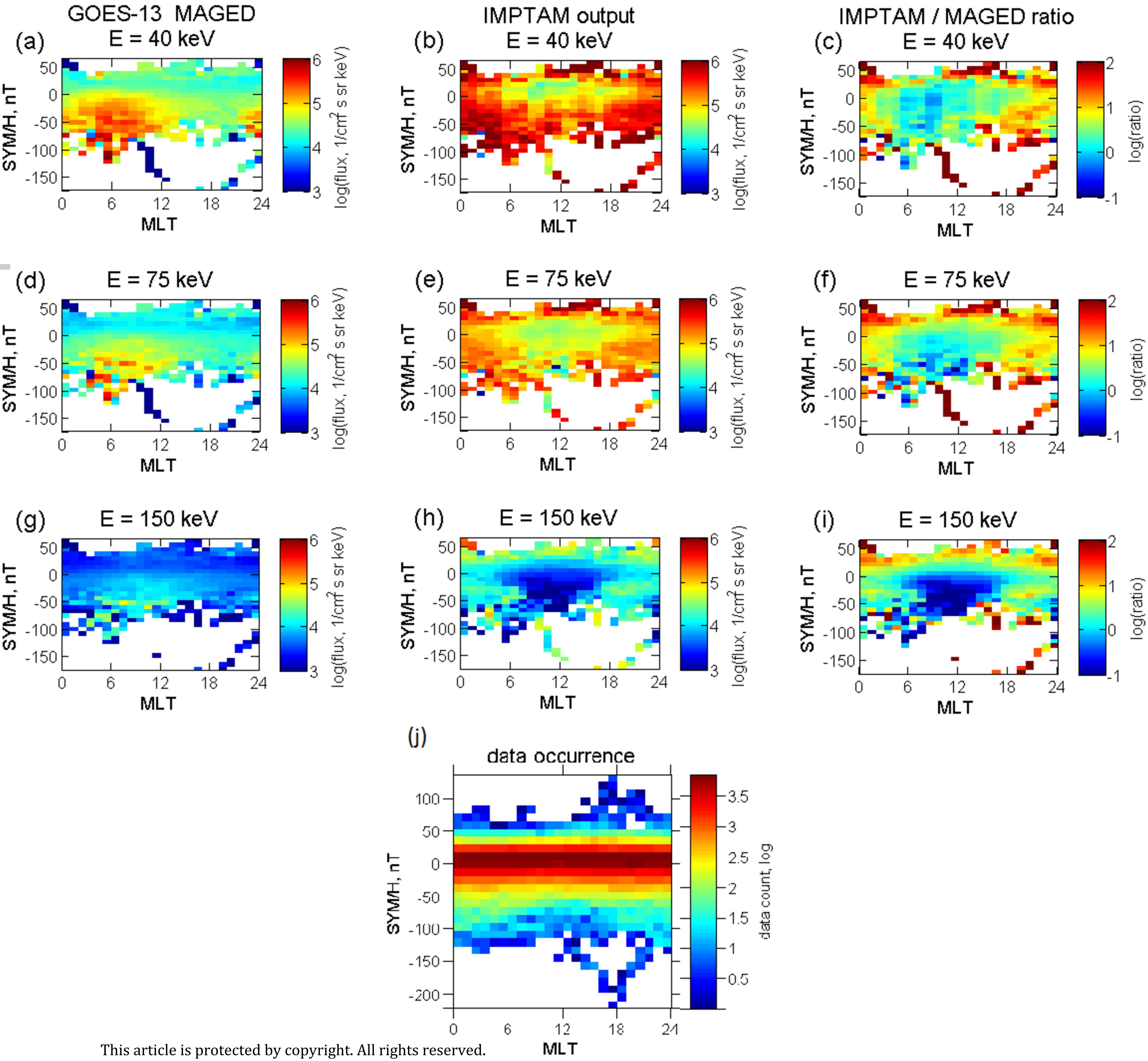

GOES-13 MAGED
IMPTAM output
IMPTAM / MAGED ratio
(a) E = 40 keV
(b) E = 40 keV
(c) E = 40 keV
(d) E = 75 keV
(e) E = 75 keV
(f) E = 75 keV
(g) E = 150 keV
(h) E = 150 keV
(i) E = 150 keV
(j) data occurrence
Kp
Kp index
MLT
log(flux, 1/cm2 s sr keV)
log(ratio)
data count, log

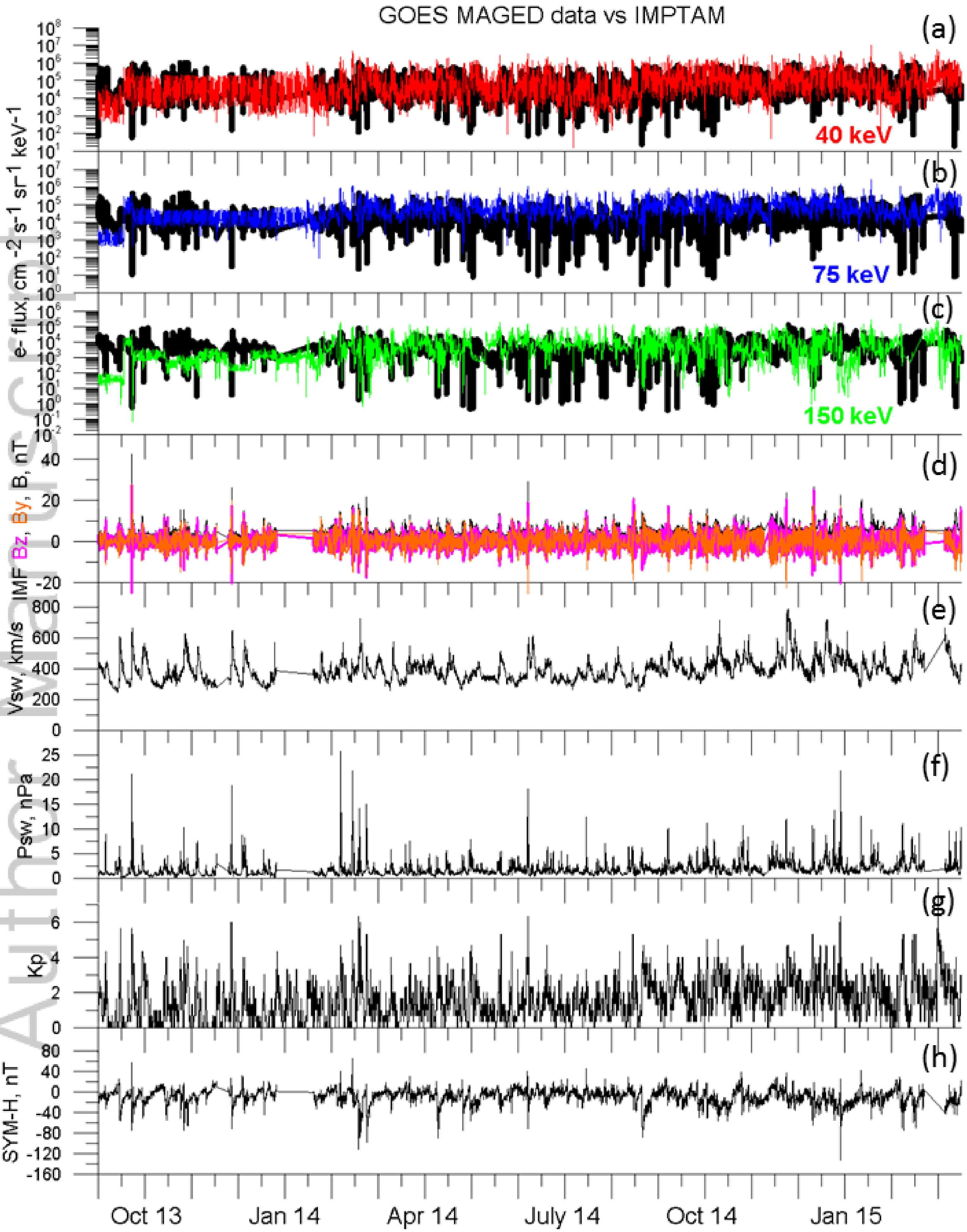
Figure 8.



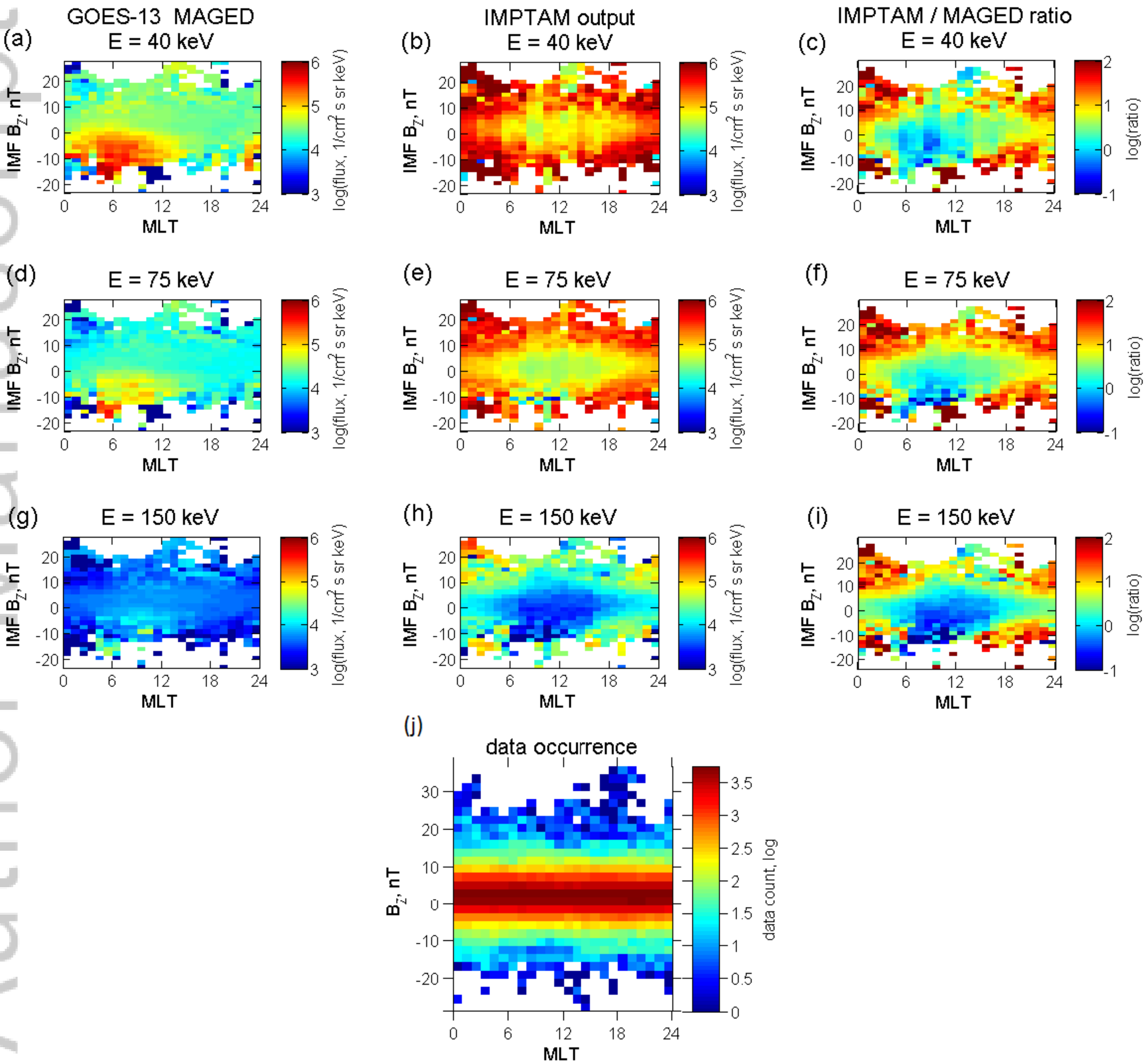
GOES-13 MAGED
IMPTAM output
IMPTAM / MAGED ratio
(a) E = 40 keV
(b) E = 40 keV
(c) E = 40 keV
(d) E = 75 keV
(e) E = 75 keV
(f) E = 75 keV
(g) E = 150 keV
(h) E = 150 keV
(i) E = 150 keV
(j) data occurrence
SYM/H, nT
MLT
log(flux, 1/cm² s sr keV)
log(ratio)
data count, log

Figure 9.

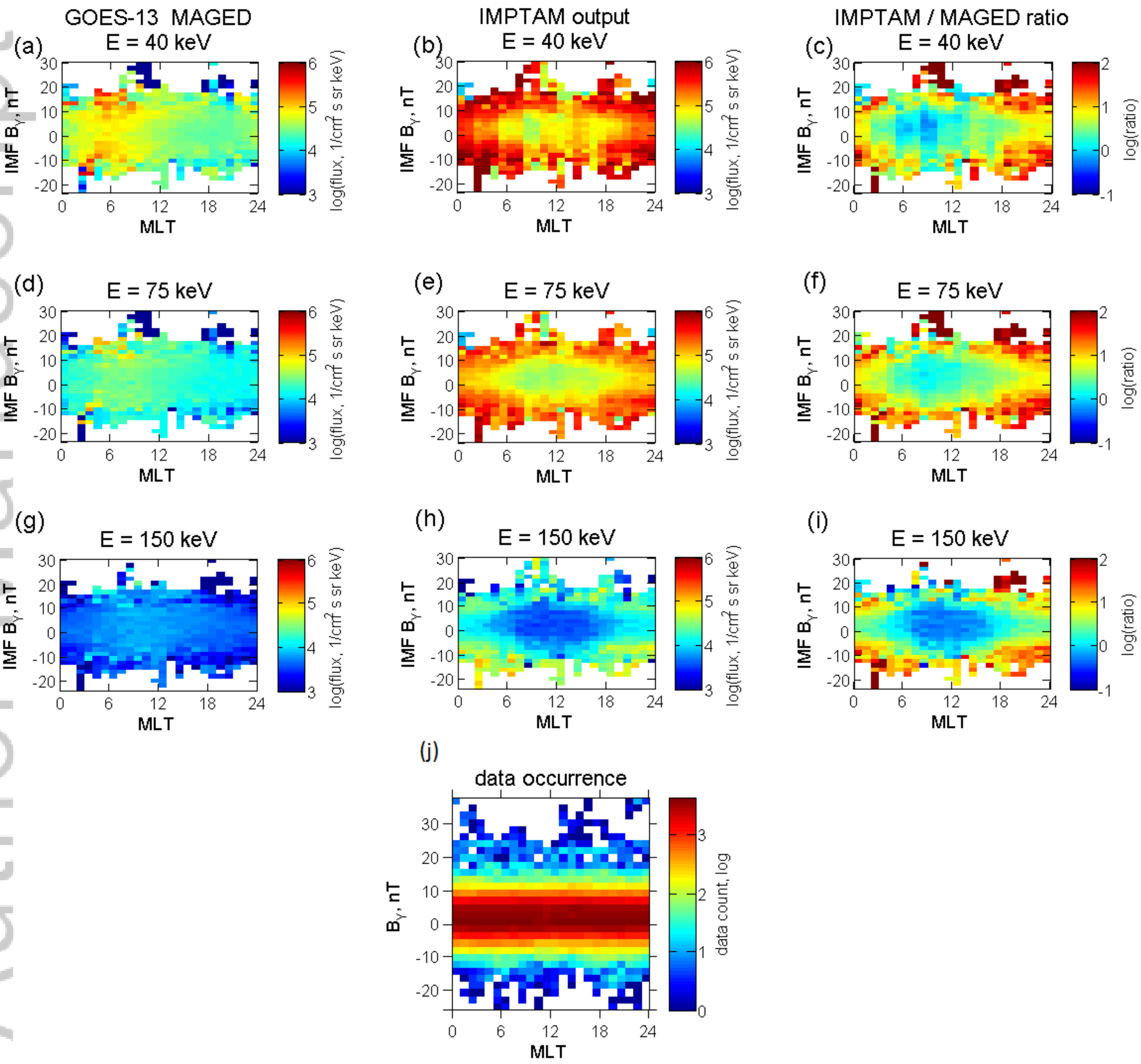
40 keV
75 keV
150 keV
IMPTAM flux, #/(cm² s sr keV)
GOES MAGED flux, #/(cm² s sr keV)
Number of points



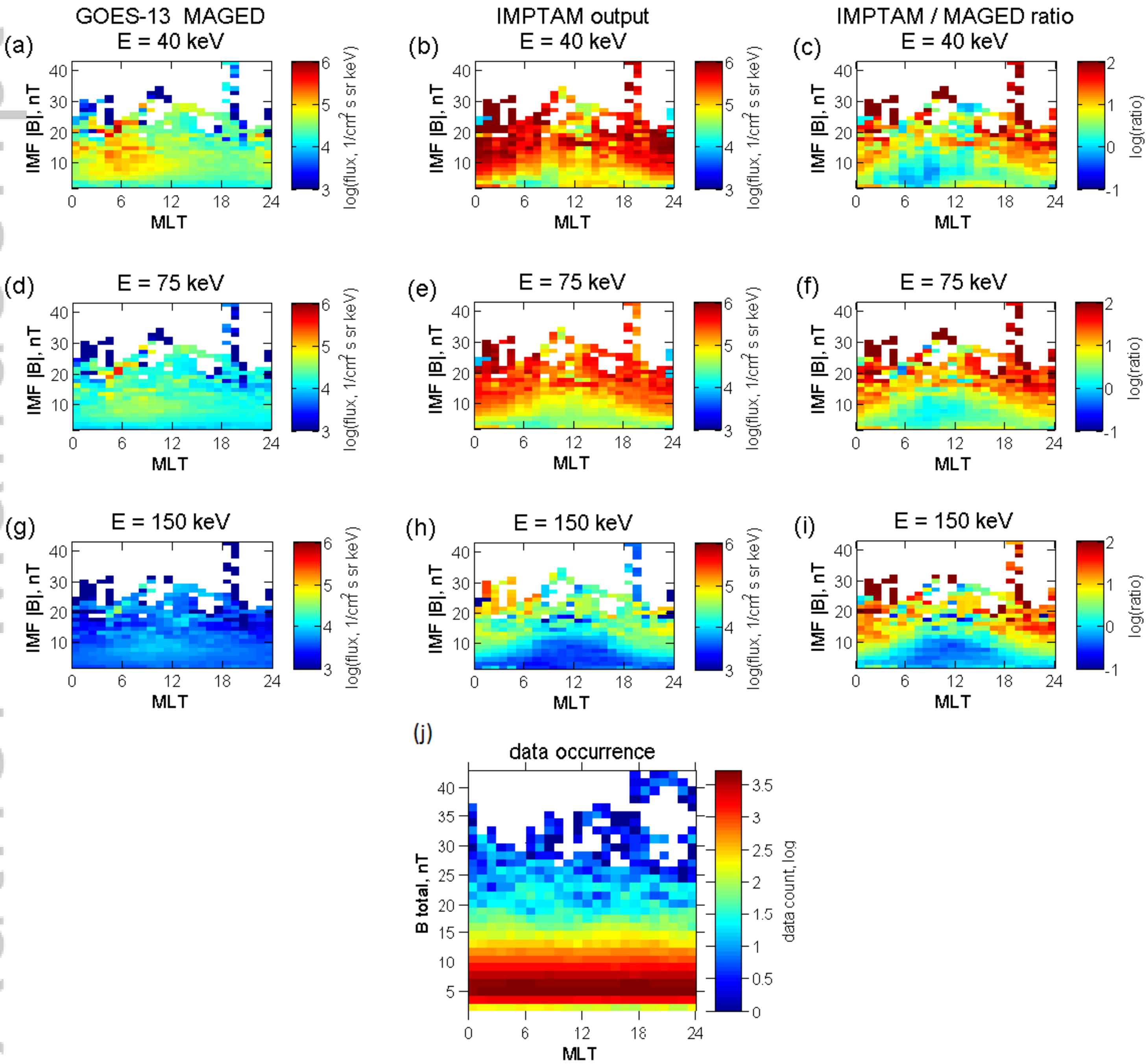
2018SW002028-f01-z-.png



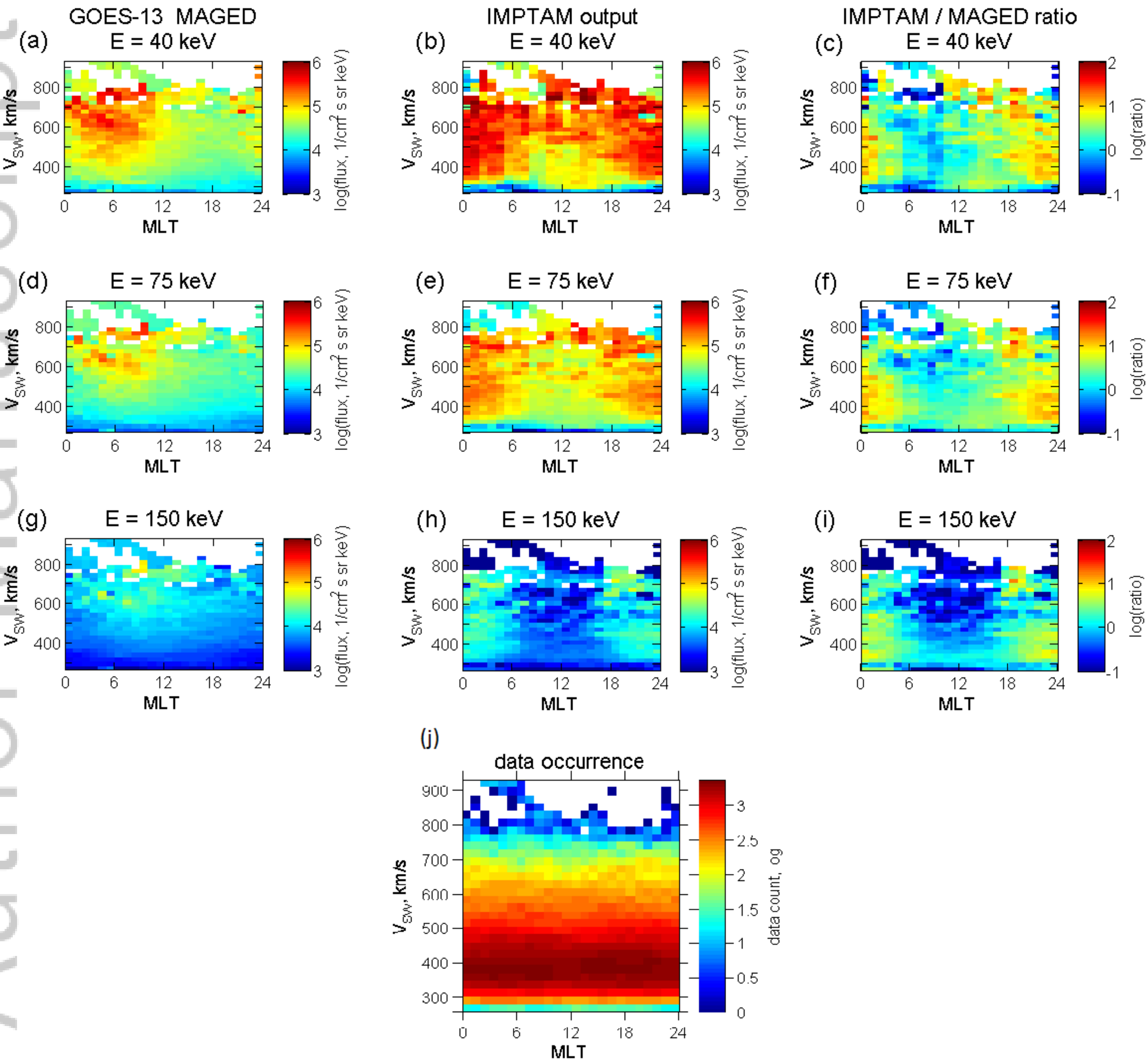
2018SW002028-f02-z-.png



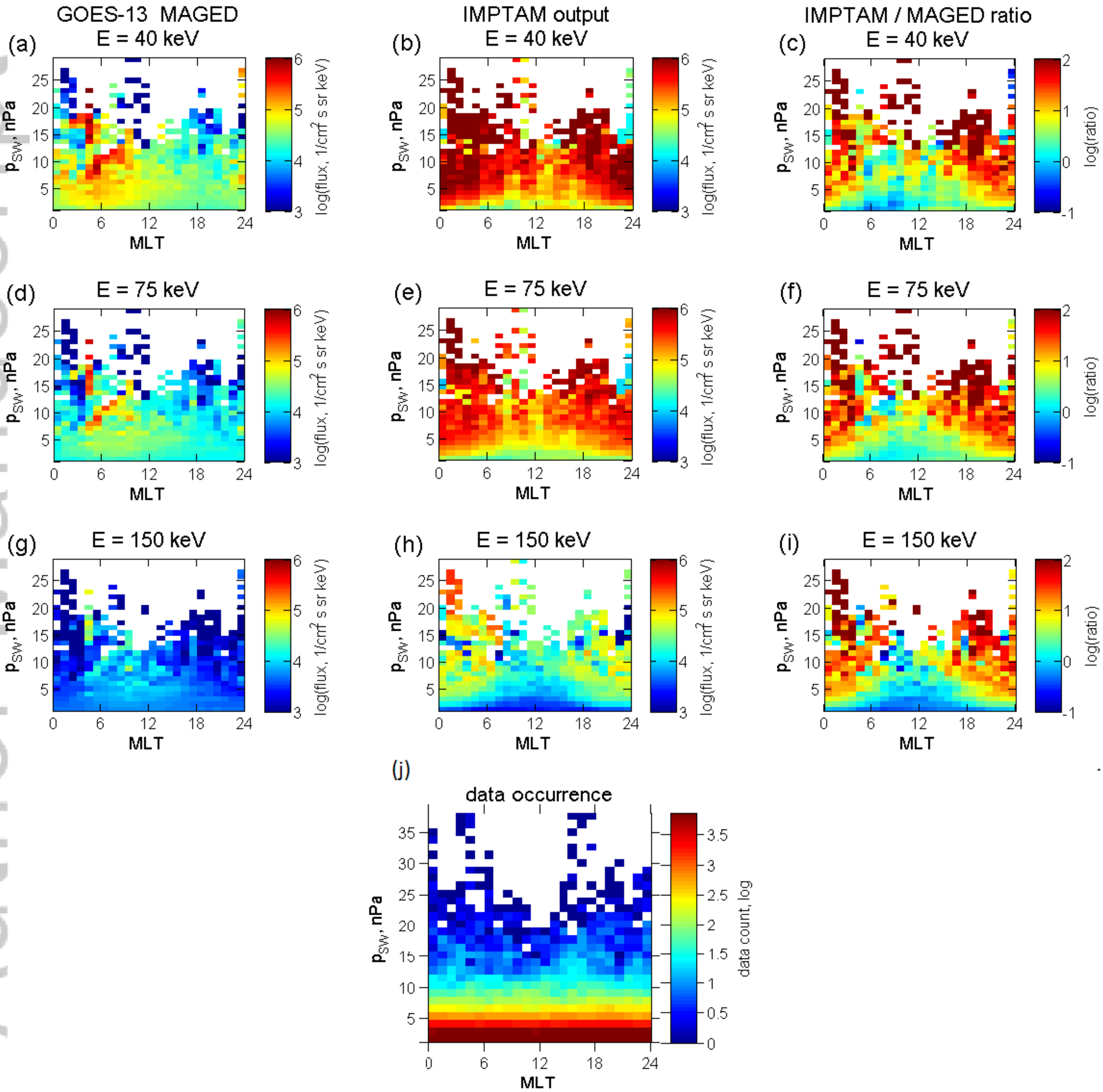
2018SW002028-f03-z-.png



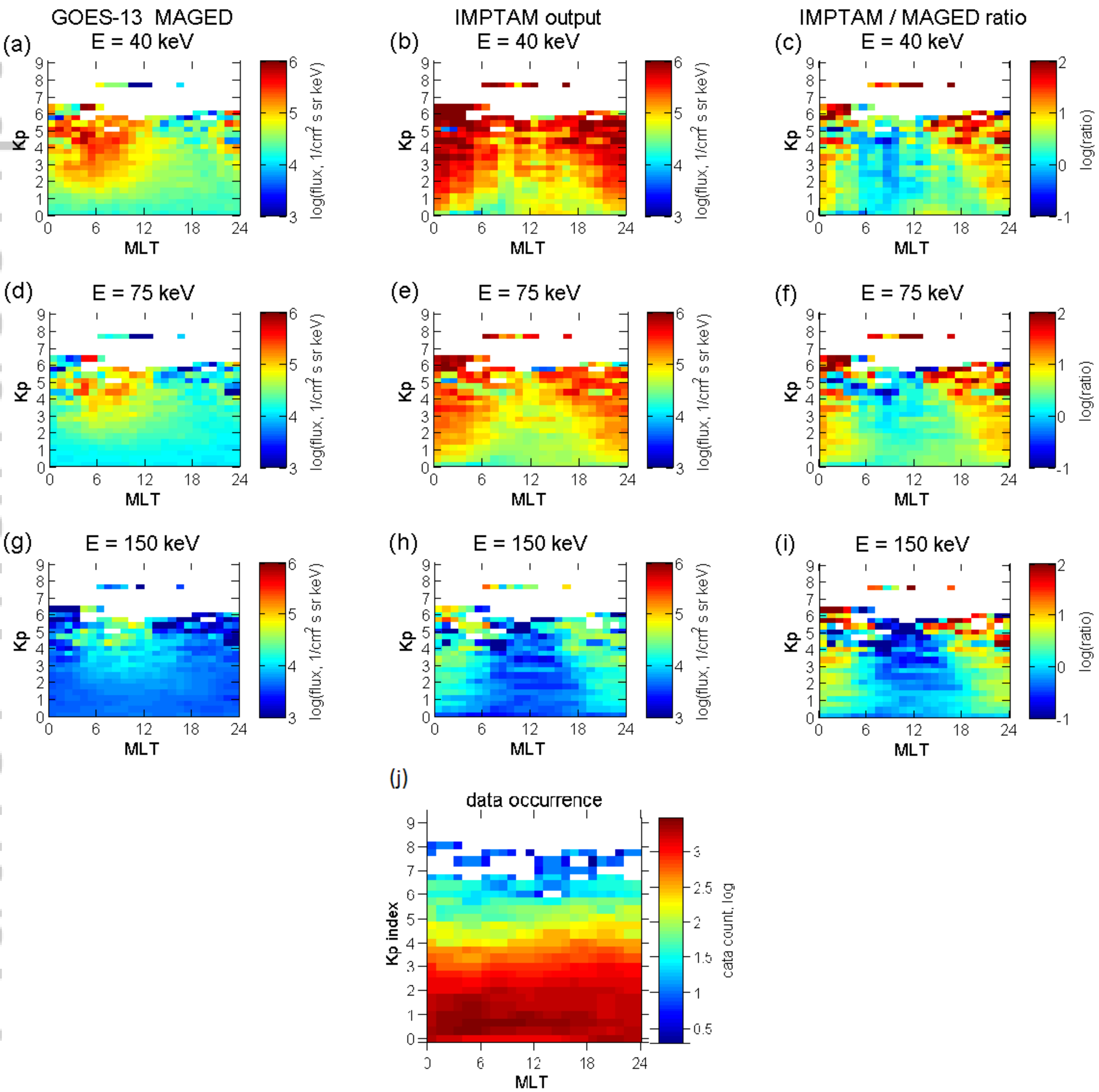
2018SW002028-f04-z-.png



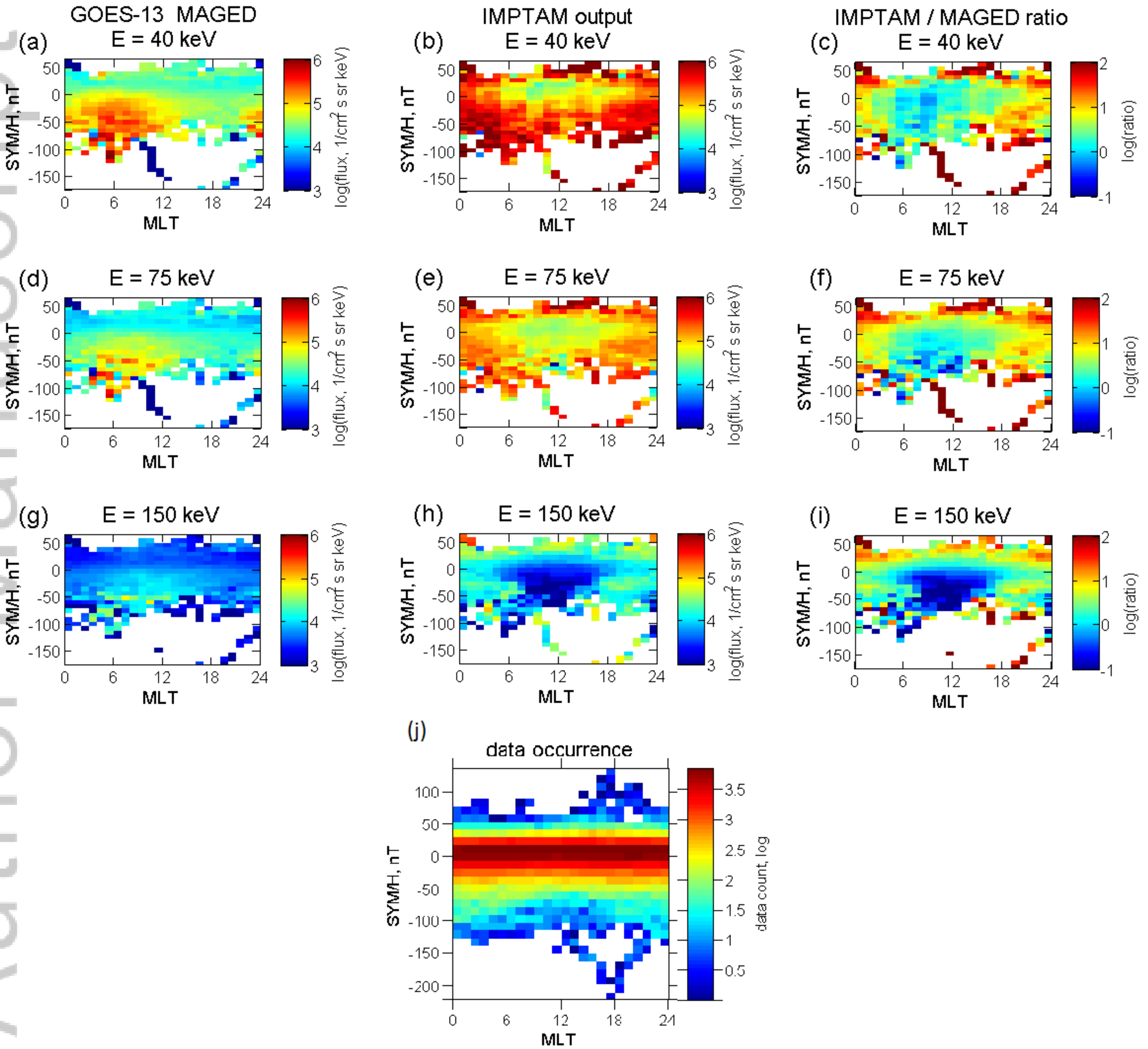
2018SW002028-f05-z-.png



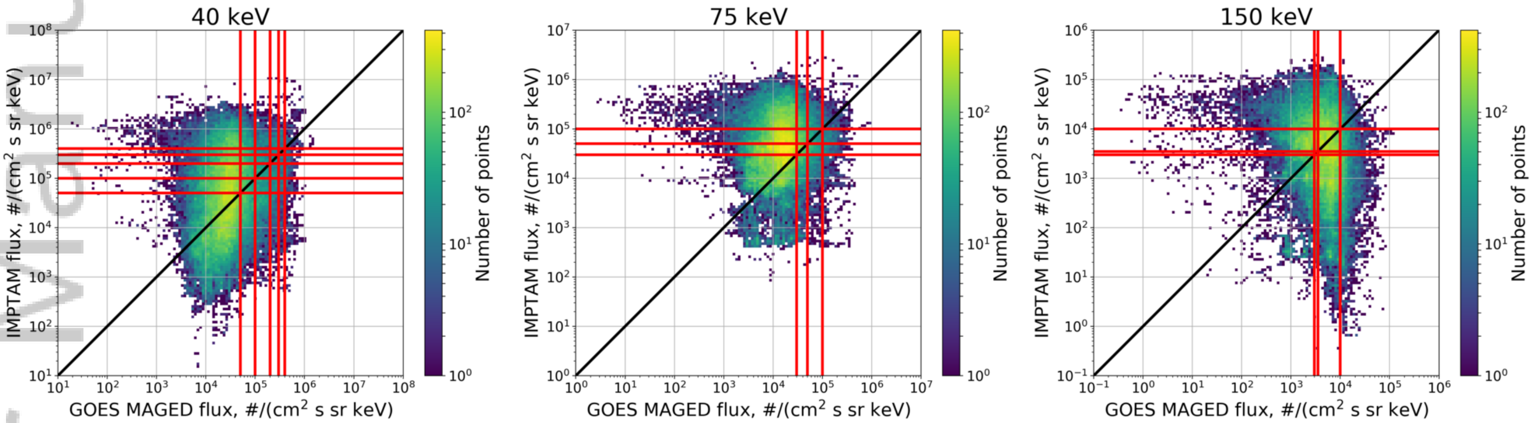
2018SW002028-f06-z-.png



2018SW002028-f07-z-.png



2018SW002028-f08-z-.png



2018SW002028-f09-z-.png