Modeling the magnetospheric X-ray emission from solar wind charge exchange with verification from XMM-Newton observations

Ian C. Whittaker,[1] Steve Sembay,[1] Jennifer A. Carter,[1] Andrew M. Read,[1]

Steve E. Milan,[1] Minna Palmroth,[2]

Corresponding author: I. Whittaker, Department of Physics and Astronomy, University of Leicester, University Road, Leicester LE1 7RH, UK. (iw47@le.ac.uk)

[1]Department of Physics and Astronomy, University of Leicester, Leicester, United Kingdom.

[2]Finnish Meterological Institute, Helsinki, Finland.

Key Points.

- MHD based simulations provide a viable alternative to empirical modeling of X-ray emissivity.
- SWCX enhancement variance strongly depends on the accuracy of the heavy ion abundance.
- Modeled X-ray emissivity provides global imaging of the magnetosheath.

Abstract. An MHD based model of terrestrial solar wind charge exchange (SWCX) is created and compared to 19 case study observations in the 0.5-0.7 keV emission band taken from the EPIC cameras onboard XMM-Newton. This model incorporates the GUMICS-4 MHD code and produces an X-ray emission datacube from O^{7+} and O^{8+} emission lines around the Earth using in-situ solar wind parameters as the model input. This study details the modeling process and shows that fixing the oxygen abundances to a constant value reduces the variance when comparing to the observations, at the cost of a small accuracy decrease in some cases. Using the ACE oxygen data returns a wide ranging accuracy, providing excellent correlation in a few cases and poor/anti correlation in others. The sources of error for any user wishing to simulate terrestrial SWCX using an MHD model are described here and include mask position, hydrogen to oxygen ratio in the solar wind and charge state abundances. A dawn-dusk asymmetry is also found, similar to the results of empirical modeling. Using constant oxygen parameters, magnitudes approximately double that of the observed count rates are returned. A high accuracy is determined between the model and observations when comparing the count rate difference between enhanced SWCX and quiescent periods.

1. Introduction

The terrestrial solar wind charge exchange process involves the liberation and capture of an electron from a neutral species at the Earth (i.e., hydrogen) to a heavy, high charge state, ion in the solar wind. The electron can be captured in an excited state and transition to lower energy states via photon emission; which in the cases of X-ray photons, is detectable by X-ray telescopes [e.g. *Cravens et al.*, 2001; *Henley and Shelton*, 2008]. X-ray charge exchange is a non-thermal emission, which was first detected in Röntgen-Satellit (ROSAT) observations of comets, when the solar wind interacted with the neutral gas outflow [*Lisse et al.*, 1996; *Cravens*, 1997]. Quantification of the X-ray emission has focussed on highly ionised oxygen [e.g. *Koutroumpa*, 2012] as most space based X-ray observatories investigate photon energies around $\frac{3}{4}$ keV and have observed SWCX in this energy range, including XMM-Newton [e.g. *Snowden et al.*, 2004], Suzaku [*Ishikawa et al.*, 2013] and Chandra [*Slavin et al.*, 2013]. More recent attempts at quantifying charge exchange from other ions with emission lines around the $\frac{1}{4}$ keV band have also been performed, though the lack of cross sectional information for a number of faint transition lines causes a high uncertainty in the results [*Kuntz et al.*, 2015]. This previous research also showed a stronger correlation of the $\frac{1}{4}$ keV band ROSAT fluxes with solar wind flux than the $\frac{3}{4}$ keV band.

Charge exchange emission has the possibility of being used as a powerful global imaging tool [*Collier et al.*, 2012]. The peak charge exchange emission is expected to occur around the subsolar magnetopause boundary where the pressure balance between the solar wind and terrestrial atmosphere sits, thus allowing magnetopause, neutral hydrogen and plasma dynamics models to be tested using a global view rather than via traditional in-situ measurements [*Robertson et al.*, 2006; *Collier et al.*, 2010]. An example of the clear boundary definition can be seen in panel e) of Figure 1. The magnetopause can be seen at a subsolar distance of $9R_E$, while the bow shock sits at around $11R_E$. The charge exchange process has previously been used for magnetopause modeling, using the resultant energetic neutral atom emission from low charge state ions [e.g., *Collier et al.*, 2005; *Hosokawa et al.*, 2008; *Ogasawara et al.*, 2013]. Charge exchange X-ray emission has also been observed at Venus [*Dennerl*, 2008], Mars [*Holmström et al.*, 2001] and the Moon [*Collier et al.*, 2014]. This indicates that for comparisons between the induced magnetospheres of the unmagnetised planets and the Earth's magnetosheath, X-ray charge exchange emission could be a valuable tool. This is especially true with magnetopause modeling as the movement of the boundary layer provides a proxy for monitoring the transfer of solar wind energy into the magnetosphere [*Milan et al.*, 2004]. Hence, modeling and testing of the terrestrial charge exchange process is necessary for understanding future imaging studies.

The XMM-Newton observatory [*Jansen et al.*, 2001] was launched in 1999 and currently moves in a highly elliptical orbit with a perigee altitude of ∼7000 km and an apogee of ∼114,000 km, allowing long observation periods (∼48 hour orbital period with 42 hours of observations per orbit). The European Photon Imaging Camera (EPIC) onboard XMM-Newton, contains two MOS CCD cameras [*Turner et al.*, 2001] and a single pn CCD camera [*Strüder et al.*, 2001], which provides a spectral resolution of $\frac{\Delta E}{E} \sim 17$. We use observations from the EPIC-MOS cameras in this study and all mention of EPIC data refers to the MOS instruments. The EPIC-MOS cameras have a circular field of view with a 30 arcminute (0.5°) diameter. While this field of view provides a high spatial resolution at galactic distances, near the Earth this corresponds to ∼60 km across the camera. When we refer to SWCX we specifically mean the terrestrial emission. Our main aim is to determine the efficacy of using a magnetohydrodynamic simulation, in comparison to empirical models which have previously been used, to compare to observed SWCX X-ray emission. *Kuntz et al.* [2015] showed that the Spreiter magnetopause model [*Spreiter et al.*, 1966], typically used in empirical modeling of SWCX, underestimates the magnetopause position. We use an MHD model for this comparative study and a more recent magnetopause model. To determine the properties of the solar wind plasma throughout the magnetosheath we use the GUMICS-4 (Global Unified Magnetosphere-Ionosphere Coupling Simulation) MHD code [*Janhunen et al.*, 2012]. The process to acquire and convert the MHD grid into an X-ray emissivity datacube is described in Section 2. We then compare a line-of-sight integral through the datacube with the observations made by the EPIC-MOS cameras in Section 3. Section 4 looks at improving the correlation between observations and modeling, while Section 5 investigates the influence of the oxygen related variables. We then provide our conclusions in Section 6.

2. Method

2.1. XMM-Newton EPIC observation cases

A sytematic identification of observations affected by SWCX has been previously determined for XMM-Newton up to revolution 1773 in August 2009 [*Carter et al.*, 2011]. These cases were found by searching for variability in the 0.5 to 0.7 keV band which is primarily made up of O^{7+} and O^{8+} emission lines. Comparison of this variability in the oxygen $\frac{3}{4}$ keV band to the steady, source removed, continuum lightcurve of diffuse emission in the 2.5 to 5.0 keV band can indicate SWCX when the correlations between them are low [*Carter and Sembay*, 2008]. Table A.1 from *Carter et al.* [2011] lists 103 observations which are affected by SWCX in order of highest variability (χ^2_μ) between the steady continuum and oxygen band. We have chosen the top 30 observations (not including the comet cases), ranked by χ^2_μ from 27.2 to 3.4 as the basis for this study. These cases are indicated as having the highest deviance between the X-ray background flux and oxygen emission. Table 1 gives the revolution number and observation identifiers for each of the selected cases, and we also include the date and the duration in hours.

2.2. Creating an X-ray emissivity cube

To create an X-ray emissivity grid for each timestep in the MHD model we use Equation (1) [*Cravens*, 2000]. This requires the combination of the GUMICS-4 MHD simulation output with both the neutral hydrogen number density and the alpha value (α), a scale factor containing the cross section of the charge exchange interaction.

$$P_X = \eta_H \eta_{SW} v_{av} \alpha \tag{1}$$

where:

P_X = emissivity (eV cm^{-3} s^{-1})

η_{SW} = solar wind proton number density (cm^{-3})

η_H = neutral hydrogen number density (cm^{-3})

α = scale factor based on cross sectional data and oxygen abundance (eV cm^2)

$$v_{av} = \sqrt{v_{sw}^2 + \frac{3k_B T}{m_p}} \quad (\text{cm s}^{-1})$$

2.2.1. The GUMICS-4 MHD model

As a first step to running the MHD model we require the upstream solar wind conditions as an input to the GUMICS-4 code. These solar wind parameters are downloaded from NASA's Space Physics Data Facility in the form of OMNI [*King and Papitashvili*, 2005] one minute resolution averages, a dataset taken from a combination of ACE, WIND and IMP 8 satellite data, and timeshifted to the bow-shock. The required input variables for GUMICS-4 are time (s), proton number density (m^{-3}), temperature (°K), solar wind speed (v_x, v_y, v_z in m/s), and interplanetary magnetic field (B_x, B_y and B_z in T). In order to ensure a divergenceless solution, the IMF B_x component is kept constant. Missing data values in the OMNI data set have a linear interpolation applied to recover appropriate values for each timestep. We run the GUMICS-4 model using a four second timestep with a data output grid produced every five minutes (300s).

Once we have produced a set of MHD output datacubes for the duration of each XMM-Newton observation, we take a cuboid spatial subset covering the regions of interest (i.e., dayside magnetosheath). We define our irregular data grid with the highest spatial resolution closest to the planet, having limits of 0 to 15 R_E in GSE x and -18 to 18 R_E in GSE y and z in increasing intervals from 0.2 R_E up to 0.5 R_E. We use this grid in combination with the relevant GUMICS-4 output file to produce a datacube giving solar wind proton number density, velocity and temperature for each GSE x, y and z grid value.

2.2.2. The neutral hydrogen model

We use the Hodges neutral hydrogen model [*Hodges*, 1994] to create a grid of neutrals in the same grid format as the GUMICS-4 data. The Hodges study used a Monte Carlo simulation process to model the hydrogen exosphere as a function of spherics. The values were given for four different solar radio flux values at 10.7 cm wavelength ($F_{10.7}$) at equinox and solstice, while also being dependent upon radial distance. During each case study we take the daily $F_{10.7}$ average and using the date of the observation, we interpolate between the four given $F_{10.7}$ values and the temporal distance from summer solstice using day-of-year number. This interpolation process produces a unique neutral hydrogen grid for each case study.

2.2.3. Calculating α

The alpha value is a proportional factor based on a combination of the relative abundances and the cross-section of each possible interaction between a solar wind ion and a neutral particle causing an emission line in the relevant energy range [*Cravens*, 2000]. The general equation for the calculation of this value is shown in Equation (2), where X is the element required and q is the charge state.

$$\alpha_{X^{q+}} = \sigma E \left[\frac{X^{q+}}{O}\right]\left[\frac{O}{H}\right] \tag{2}$$

In the case of calculating the oxygen emission lines we need to know the emission line energy, cross section of the interaction, abundance of the relevant charge state and the ratio of oxygen to hydrogen (O/H) in the solar wind. We use two hour time resolution, O/H ratio and oxygen charge state abundance data from the ACE spacecraft, timeshifted to the bow shock in the same way as the OMNI data. If no solar wind composition data is available we use the values in *Schwadron and Cravens* [2000], which is discussed further in Section 4.2.

The cross section value and energy for each transition of O^{7+} (seven transitions) and O^{8+} (five transitions) with neutral hydrogen are based on experimental data taken from *Bodewits* [2007]. The cross section value for each transition is interpolated based on the input ion speed, from the five values given (200, 400, 600, 800, and 1000 kms^{-1}). The individual alpha values for each transition are then summed to produce a combined alpha value for all relevant oxygen transitions.

For comparison to empirical methods, the α value for the model timeframe shown in Figure 1 is 7.6×10^{-16} eV cm^2 . This has been calculated based on the input solar wind velocity of 438 km/s, an $\frac{O}{H}$ ratio of 1.1×10^{-3}, an O^{7+} abundance of 0.28, and an O^{8+} abundance of 0.05. This value compares favourably with empirical α values of 6×10^{-16} eV cm^2 [*Cravens et al.*, 2001; *Robertson and Cravens*, 2003].

2.2.4. Combination of the data

We now have all the requirements to produce an X-ray emissivity cube. The output of the GUMICS-4 simulation is combined with both the neutral hydrogen number density and the alpha value as shown in Equation 1. Example slices through the data grid for each of the components which are combined to form the X-ray emissivity grid are given in Figure 1. Each panel shows an example 2D slice through the 3D datacube in the ecliptic (x-y) plane at z = 0. Panels a) to c) of Figure 1 show GUMICS-4 data output of the solar wind proton number density, speed and temperature respectively. Panel d) shows a neutral hydrogen data slice from the Hodges model. Panel e) shows the result of combining the model output, neutral hydrogen data and cross section values together to create the model X-ray emissivity.

2.2.5. Applying a mask

As a one fluid simulation, the GUMICS-4 MHD code does not identify the difference between solar wind plasma and plasma of terrestrial origin. As a consequence we can observe high terrestrial plasma densities near the Earth which in turn produce unphysical X-ray emissivity. The terrestrial plasma does not contain the same ratio of highly ionized oxygen species and hence cannot produce the same level of charge exchange emission. We can clearly see this effect by comparing panels f) and g) of Figure 1 which are slices of panel e). In panel e) we present the x-y plane of X-ray emissivity which shows some emission enhancement very close to the Earth. When we examine the y-z plane with a cut taken at $x = 3.9$ R_E, in panel f), we see the extent of the terrestrial plasma. Panel g) shows a cut at $x = 9$ R_E where the emission is not affected by the terrestrial plasma. The Earth size and position is also included in panels e), f) and g) for comparison. We assume the boundary between the terrestrial and solar wind plasmas is at the magnetopause [*Spreiter et al.*, 1966]. The magnetopause model given in *Shue et al.* [1998] defines this boundary, with all proton densities (panel a) within this region set to zero. The relative merits of the empirical and MHD based magnetopause are discussed in Section 4.1.

2.3. Calculating an estimated instrument view

We now have a three dimensional X-ray emissivity product in an irregular grid. To simulate what XMM-Newton would see we integrate along the viewing path from the satellite location. To determine the amount of X-ray emission directed along the line of sight we integrate along the look vector, $\int P_X dS$, from Equation (1).

At every 0.5 R_E step distance through the data cube from the satellite location we take an interpolated emissivity value. The nearest neighbor emissivity data points within a 0.5 R_E radial distance are taken and averaged, with each neighbor weighting dependent upon the distance from the required point. This interpolated emissivity value is multiplied by the step distance and totaled to create the integral column flux. While this integral energy collection value can be used for comparison, as a final step we pass both the energy and appropriate spectrum of the oxygen transitions through the EPIC instrument response matrix to provide a counts per second (c/s) value in the 0.5 to 0.7 keV band. The start and end times of each EPIC timestep (at 1000s resolution) can then be determined and an appropriate average count rate for that specific time period returned. Shorter step distances were also trialed, resulting in negligible flux differences due to the weighted averaging method.

3. Results

An initial investigation of the 30 case studies showed that the EPIC camera suffered from sparse data in six cases, which were removed from the study. Of the remaining 24 cases another five had XMM-Newton at a negative x value, i.e. antisunward, with an instrument view direction that did not intersect our datacube and these cases were also removed, leaving 19 case studies with data. Each case was assigned an identifying code, comprising the year and number of the case within that year in date order (e.g. YY-C). These identifiers have been included in Table 1.

3.1. Background removal

Processing of the raw EPIC data to produce the lightcurves used in this study is described in detail in *Carter and Sembay* [2008]. This includes the methodology for identifying and removing astrophysical point sources from the data and cleaning the data of soft-proton flares which can produce a strongly variable diffuse background.

The residual diffuse signal is dominated by the variable foreground SWCX component and background components which are non-variable on the timescale of individual observations. This background is a combination of an X-ray component and the residual particle background in the EPIC detectors. The background X-ray component is a combination of the astrophysical X-ray background arising from emission from our Galaxy and unresolved point sources (extragalactic active galactic nuclei), and from SWCX in the wider heliosphere.

The particle background in each observation can be estimated by a well establised procedure [*Carter and Read*, 2007]. We have estimated the X-ray background from the ROSAT all sky survey [*Voges et al.*, 1999]. For each look direction on the sky appropriate to the EPIC data we have used existing procedures within NASA's HEASARC toolkit to derive an estimated count rate in the EPIC instrument 0.5 to 0.7 keV energy band from the observed count rate in the ROSAT R4 band, which has an energy range of between

0.44 and 1.01 keV. To make the conversion from one instrument to the other requires the assumption of a spectral model. Formally the diffuse X-ray background spectrum is well represented by a two-component thermal APEC model for the Galactic emission and a power law for the unresolved power law [*Kuntz and Snowden*, 2008]. We have used the spectral parameters from a deep analysis of case 01-3 previously studied by *Carter et al.* [2010] modified by the appropriate absorption in the light-of-sight which is provided by the HEASARC toolkit and derived from the Leiden/Argentine/Bohn (LAB) neutral hydrogen survey [*Kalberla et al.*, 2005]. Technically the spectral parameters will vary according to sky position, however, the ROSAT to EPIC conversion in these bands is not very sensitive to plausible variations in the parameters and the resultant uncertainty is comparable to the uncertainty due to the inter-calibration between the instruments.

The total estimated background for each observation is listed in Table 1. In comparison with the median observed count rate, also listed in Table 1, we can see that the background (i.e. non local SWCX components) represents between ∼10-40% of the observed 0.5 to 0.7 keV signal in these cases.

3.2. Observation to model comparison

For each case study we produce a full set of plots including the GUMICS-4 integral energy and estimated count rate output, the solar wind conditions, the EPIC lightcurve, and satellite positional information. These plots allow us to check for errors and notice patterns in large count rate differences. An example of this type of plot is shown for case 01-1 in Figure 2. The left hand panels show; the GUMICS-4 integral energy output (a), the GUMICS-4 estimated count rate (b), the velocity and number density of solar wind protons (c), B_z and dynamic pressure (d), the oxygen to hydrogen ratio (e) and the oxygen state abundances (f). The right panels show; the observational data (g), a comparison between GUMICS-4 and observations before background removal (h), a comparison between GUMICS-4 and observations after background removal (i), a normalized comparison (j), the radial distance of the satellite and magnetopause (k) and the final smaller panels show the position, orbit path and look direction for the case.

The 01-1 case is of interest as it covers a long time period (26.5 hours), shows a wide range of features, and was previously examined in detail by *Snowden et al.* [2004]. We see good agreement between the background removed observations and the GUMICS-4 count rates (panel i) from ∼15:00 onwards, including a gradual decline in magnitude starting around 22:00. The start of the case study suffers from an integral emissivity which is several orders of magnitude in error (panel b). This is discussed further in Section 4.1 but results from the mask not accurately removing all the plasma of terrestrial origin. Panel f) also shows a large variation of O^{7+} abundances, ranging from 10% to almost 50%, which is discussed in Section 4.2.

The GUMICS-4 count rate estimation comparison to the background reduced EPIC observations, i.e. panel i), for all 19 cases are shown in Figures 3 and 4, with the respective identification number from Table 1. To provide an initial comparison between the cases we determine the median count rate of both the modeled and observated lightcurve and take a ratio of the two. The magnitude ratios for each case varied between 0.11 and 20.9 with a median value of 1.65. This magnitude difference average is reasonable for comparison although it is highly variable within cases, as observed in Figures 3 and 4. The correlation between the two lightcurves was also calculated for each case, the average for all cases was 0.07 with a standard deviation of 0.52. We discuss the importance and large variance of these values in Section 4. The correlation, based on a zero time lag cross correlation and normalized lightcurves, rather than covariance is used due to the large magnitude differences indicated by the magnitude ratio limits.

4. Discussion

The average magnitude ratio between modeled and observed count rates of 1.65 indicates the model count rates are comparable to the observations although, as previously mentioned, this comes with a large variability. We can compare this ratio to the empirical study of *Carter et al.* [2011] who found less than a factor of ∼2 in magnitude difference for 50% of their cases. In our study we find only 6 of 19 cases (32%) within a factor of two greater or smaller (i.e., a ratio between 0.5 and 2). This difference suggests that the modeled process does a poorer job of magnitude modeling than the empirical study. The correlation values are of concern with a mean value close to zero, due to 8 of the 19 cases (42%) returning a negative correlation. These correlations can have high values and the average absolute correlation value is returned as 0.44. In an attempt to determine why the simulation of the X-ray emission is not reproducing the observations accurately we investigate each of the model components, as set out in Section 2.2.

4.1. A re-examination of the method

From Equation (1), there are four important possible sources of error in our modeled data; the MHD model, the neutral hydrogen model, the mask to remove the cold terrestrial plasma, and the alpha value.

The GUMICS-4 model code has been verified in a one year study [*Gordeev et al.*, 2013] and used in a range of other studies [e.g., *Hubert et al.*, 2006; *Palmroth et al.*, 2013]. The requirements for the magnetospheric plasma simulation are well within the boundaries set by GUMICS-4 of ± 64 R_E in y and z and up to 32 R_E in x. The main limitations of the model, as described in *Janhunen et al.* [2012], are magnetotail reconnection and near Earth plasma modeling (<3.7 R_E). The first limitation is not relevant as we only generate X-ray emission data at $x > 0$ and the second limitation is taken care of by use of a magnetopause position mask. It is also important to note that by its fluid nature, MHD models have difficulty accurately simulating the physics in regions where details of the plasma distribution function are important, such as areas where kinetic effects are dominant. As we are focussing on the magnetosheath emission this is less of an issue than if we were to be looking at the cusp regions.

The Hodges neutral hydrogen model has been used in this study. A comparison of other neutral hydrogen models was performed for equinox and solstice at high and low $F_{10.7}$ values, included as supplementary material to this manuscript. The first compared model was the Bonn model [*Nass et al.*, 2006] using the coefficients from the TWINS-1 LAD data [*Bailey and Gruntman*, 2011]. We also compared the *Østgaard et al.* [2003] IMAGE model; while only designed to be used on the nightside, the returned values are comparable to the Hodges model. The last comparison was a simple r^{-3} model with a 25 cm^{-3} number density at a distance of 10 R_E [*Cravens et al.*, 2001]. The results of each comparison show very minor differences in shape and magnitude, certainly not enough for the Hodges model to be the cause of the variations between the modeled and observed lightcurves. It should also be noted that when comparing SWCX through different parts of the magnetosheath, *Kuntz and Snowden* [2008] demonstrate that the solar wind flux is a more important factor than the magnetosheath density along the line of sight. This indicates that small differences in the Hodges number density are unlikely to make any significant differences.

We next investigate the mask used to remove the cold terrestrial plasma. The Shue model is a commonly used magnetopause positional model [e.g. *Liemohn et al.*, 1999; *Dimmock et al.*, 2015] providing a subsolar distance and flaring value based on solar wind conditions. The position of the magnetopause has been extensively tested with our model output. In terms of subsolar stand off distance, the position appears reasonable most of the time but as the model has no historical knowledge of the conditions it can change position rapidly while the plasma simulation suggests a slower movement. This swift movement results in the mask occasionally being placed within the plasmasphere as described by the GUMICS-4 model, allowing the dense terrestrial plasma to be included in the X-ray emission grid increasing the integral line emission by several orders of magnitude as seen in the modeled emission in Figure 2a). It is apparent these large magnitude increases are due to poor masking by looking at the normalized data of case 01-1, panel j) of Figure 2. In this normalisation panel we have included emission along the x-axis without any masking in red, by providing no mask we can determine whether large increases are due to higher emission or errors in mask position. At 12:00UT in panel j) we see a very large increase in integral magnitude, yet there is only a small increase in the non-masked subsolar emissivity. At this time B_z turns negative which will instantly move the Shue mask position Earthward, while the MHD model will take time for this change to have an effect. *Gordeev et al.* [2013] generally found good agreement between the empirical Shue magnetopause and the fluopause [*Palmroth et al.*, 2003] defined from GUMICS-4 simulations. The greatest differences were found in strong southward B_z conditions near the subsolar point, where the simulated magnetopause position can be up to 15-20% more distant than the Shue model. This masking issue is also discussed more fully in *Kuntz et al.* [2015] who use a closed field line model to place the mask on their BATS-R-US MHD model [*Powell et al.*, 1999]. The closed field line approach was not used in this study as the field model will respond in a similar instantaneous movement to the Shue magnetopause model, resulting in similar errors.

An MHD model based magnetopause has been trialed for our case studies, constructed by applying a gaussian fit to the proton number density along the sub-solar line. The magnetopause can then be taken as a full width half maximum distance from the central location, an example is shown in panel k) of Figure 2 as the black dashed line. The positional difference between the Shue and proton defined boundaries is small but the proton boundary is much smoother. This model defined magnetopause produces excellent subsolar distances as defined, but with no angular data the magnetospheric flanks are poorly determined. During the testing process we also attempted a region threshold detection method, which failed regularly due to the low intensity of the flanks compared to the nose. As it is clear where the Shue method differs from the MHD model (by the dramatic increase in integral emission), it is simple to remove these times by applying a magnitude upper limit of 10 counts per second to the model count rate data.

This leaves only the alpha value as the main source of variation error. While the cross sectional data for all possible solar wind ions in the 0.5 - 0.7 keV range is limited, it is assumed that O^{7+} and O^{8+} are the major contributors and other ion species line spectra in this range will be negligible. Hence, we are left with the upstream data inputs on the oxygen charge state abundance and total oxygen number density. These values are highly variable over each case and so we investigate whether the variance in oxygen data from ACE is responsible for the primary variation in the simulated X-ray emission.

4.2. Oxygen composition data

Previously utilised empirical models have been run using constant values for the oxygen to hydrogen ratio and the charge state abundances. Whilst these values have been used as a back up for missing compositional data bins during the analysis process, the variability from the ACE data to the constant values has been quite high. This in turn could be a source of error either in the observed oxygen data or for models using the constant values. As part of this study we have included not only the nineteen cases with data analysed in the results section, but also the five cases where the XMM-Newton pointing direction did not intersect the datacube. These extra cases are labelled in Table 1.

4.2.1. Oxygen to hydrogen ratio

The model constant values of the oxygen to hydrogen ratio (O/H) are 6.45×10^{-4} for fast solar wind and 5.62×10^{-4} for slow solar wind [*Schwadron and Cravens*, 2000]. Investigating the data values given by ACE, suitably time delayed to the bow shock, we note some wide variation from these constant values. Across all 24 cases the median and mean O/H ratios are 3.11×10^{-4} and 3.94×10^{-4} respectively with a standard deviation of 3.01×10^{-4}. While there is likely to be a reasonable amount of error in the ACE data values due to limited instrument sensitivity, viewing angle and resolution, this should still produce an average value close to the Schwadron and Craven (hereafter referred to as S&C) constant values if both are representative.

Panel a) of Figure 5 shows the O/H ratio across all 24 case studies in date order, with each data point taken at a 300s resolution. The solid orange background shows the extent of the mean value for each case ±1 standard deviation. The solid blue line within each region shows the mean value, while the data are shown in black. The S&C values of the O/H ratio are indicated by the dashed red and blue lines for fast and slow solar wind respectively. This plot is complemented by a histogram of the ratio distribution in panel b) with a bin size of 5×10^{-5}. This histogram shows a skewed normal distribution, with the S&C constant values intersecting at a ratio greater than the FWHM value.

When we compare the O/H ratio to solar wind speed, using a 500 km/s cutoff between fast and slow solar wind, the fast solar wind shows a correlation between speed and ratio. We also observe that a comparison of the O/H ratio to the solar wind proton density shows a correlation, indicating that the dynamic pressure should provide a correlation too (as it is based on speed and density). While the speed and density plots are not included for space, we have plotted the O/H ratio against the solar wind dynamic pressure in panel c) of Figure 5 with the slow solar wind data points in black (dashed red fit line) and the fast solar wind ratio values in blue (solid red fit line). We observe that as expected, the fast solar wind correlates very well ($r^2 = 0.78$), producing the power law fit shown below.

$$\frac{O}{H} = 6.42 \times 10^{-4} \, P_{dyn}^{-3.19} \tag{3}$$

The slow wind fit correlates poorly ($r^2 = 0.21$) which is an expected result from the lack of correlation with both speed and density. The dynamic pressure relation could simply be symptomatic of the ACE measurements increasing in signal to noise ratio as the total solar wind content increases, producing more accurate results. It should be noted that a constant ratio value defined by the median of the slow wind data ($\sim 3.17 \times 10^{-4}$) is more appropriate in these cases than using the S&C value of 5.62×10^{-4}.

4.2.2. Oxygen charge state ratios

The other data observations required in the α value determination are the O^{7+} and O^{8+} abundances as a fraction of the total solar wind oxygen. The values taken from Table 1 of *Schwadron and Cravens* [2000] give abundances of 0.2 and 0.07 for O^{7+} and O^{8+} respectively for the solar wind. These slow wind values are used as a replacement for missing observational data for all cases. The S&C abundance values for fast wind are 0.03 and 0.00 for O^{7+} and O^{8+}, indicating that only 3% of the oxygen is available to produce SWCX in the 0.5 to 0.7 energy range. Using these values are likely to result in undetectable count rates, which from Figures 3 and 4, is clearly not the case and so the slow wind abundances are used for fast wind cases as well. Figure 5 shows the mean abundance taken from ACE of O^{7+} (panel d) and O^{8+} (panel e) for each of the 24 cases with the expected value shown as the red dashed line. As the abundance values are given on a 2 hour resolution most cases have fewer than 5 measurements and so a standard deviation is not appropriate for visualising the variance of the charge states. The error bars on the plots in panels d) and e) show the maximum and minimum values in each case. We note that looking at panel d), 14 of the 24 O^{7+} cases (58%) have error bars that do not cross the expected value at all. The O^{8+} abundance ranges show a similar result with 13 of 24 cases (54%) where the error bars do not cross the equivalent expected value. The mean and median values for all O^{7+} cases are 0.28 and 0.31, while the equivalent averages for O^{8+} are 0.05 and 0.03. The mean values suggest the S&C constant abundance value is acceptable for both charge states, however, this does not take into account the high variability. The final panel (panel f) of Figure 5 shows the O^{7+} to O^{8+} abundances, with the red dashed lines showing the appropriate expected values. The correlation between O^{7+} and O^{8+} abundance is to be expected and we can fit a power law to the data (shown in blue), given by $O^{7+} = 0.78\ O^{8+\ 0.32}$.

It should be noted that during our analysis we noticed that in certain cases the O^{8+} abundance value from ACE reached exceptionally high values. Therefore one of the conditions put in place during our analysis was that the O^{8+} abundance was not allowed to exceed 0.2, any values which did were set to the *Schwadron and Cravens* [2000] value of 0.07. This limit was an arbitrary value based on the expected O^{7+} abundance and was only enforced due to a few exceptionally high abundances causing non-physical variances in the GUMICS-4 simulated count rate. Three cases required this adjustment to one of the O^{8+} data points; 02-3, 03-4, and 03-6. The relation between O^{7+} and O^{8+} abundances in panel f) of Figure 5 suggests the O^{8+} abundance at $\sim$0.15 may also be artifically high.

The same is true for the O^{7+} abundance at ∼0.52, without a reference for the range of values that this abundance can take we did not limit the O^{7+} value in this study.

5. Reanalysis with a constant O/H ratio

5.1. Removing the oxygen variation

We have observed from the rapid changes in oxygen composition and number density, that the oxygen variances observed in each case are high. To attempt to determine if the oxygen variance is causing strong model emission variances we have reanalysed two case studies with a range of O/H and charge state abundance values. The important difference is that we do not allow the oxygen values to vary over the cases. The two cases (00-2 and 01-3) were chosen because of their difference from the modal oxygen value in Figure 5. Case 01-3 is also the observation used in *Carter et al.* [2010] for observing SWCX enhancement during a coronal mass ejection interaction with the magnetosheath. Figure 6 shows the two cases with 00-2 in the left panels and 01-3 in the right panels. The top row shows the original model result using ACE oxygen composition data, the blue line shows the model counts and the black line is the observational data points included with the appropriate errorbars. The second row of Figure 6 shows the *Schwadron and Cravens* [2000] O/H ratio of $\frac{1}{1780}$ for slow solar wind and O^{7+} and O^{8+} abundances of 0.2 and 0.07 applied respectively. The third row of Figure 6 shows the results from using the modal O/H ratio from Figure 5 of 2×10^{-4} and using the case mean O^{7+} and O^{8+} abundances. The final row of Figure 6 shows the results from using the mean O/H ratio, O^{7+}, and O^{8+} abundances from each case. These mean values for case 00-2 are 7.26×10^{-4} for the O/H ratio, 0.19 for the O^{7+} abundance and 0.027 for the O^{8+} abundance. The equivalent values for case 01-3 are 5.03×10^{-5} for the O/H ratio, 0.31 for the O^{7+} abundance and

0.14 for the O^{8+} abundance.

Figure 6 clearly shows that in both cases setting the O/H ratio to a constant results in a simulated lightcurve variance that is a lot closer to that of the observed variance. The magnitude of the ratio can be seen to directly affect the magnitude of the output X-ray emission. In terms of matching the magnitudes as closely as possible, the modal O/H ratio from Figure 5 provides the closest match for both examples in Figure 6. We therefore recalculate all our simulated X-ray emission lightcurves using a constant O/H ratio of 2×10^{-4} and the mean O^{7+} and O^{8+} abundances for each case. Each case comparison is shown once again in Figures 7 and 8, which can be directly compared to Figures 3 and 4.

5.2. Accuracy of the modeled to observed magnitudes

A comparison of the cases with a fixed O/H ratio and those using ACE in Figures 3, 4, 7 and 8 allows an intial quality check by eye. Of the 19 cases, 11 show visible improvement (58%), 5 show little to no improvement (26%), and only 3 cases show a decline in both magnitude and variance matching (16%). This basic check confirms that we should continue the analysis with these new simulations.

Figure 9 shows the comparison of these newly calculated model count rates to the background removed EPIC observations. The top panel shows both the mean and median magnitude ratio between the modeled and observed counts for each case. Taking the median ratios, the minimum value is 0.38, the maximum is 12.45 and the median value is 2.23. This range shows that while the median magnitude is slightly higher in ratio than the ACE varying modeled count rates, the range of the spread is much smaller. This can be seen by the fact that 8 of the 19 case averages (42%) now sit within a factor of 2 higher or lower of the observation magnitude average, two cases greater than the ACE O/H varying results (Section 3.2). The second panel of Figure 9 shows the correlation value of the modeled and observed count rates. In comparison to the ACE varying data, the median correlation is now 0.35 (compared to 0.07) with a standard deviation of 0.48 (compared to 0.52) and 5 of the cases show negative correlation. This indicates that by removing the oxygen variation we obtain much better correlations between the model and observations. The median of the absolute value of correlation is 0.57 (compared to 0.44), indicating that whether the case is positively or negatively correlated the variances are more closely related with the O/H ratio kept constant.

The lower panel of Figure 9 shows a scatter plot of each observed count rate bin against the respective modeled count rate for all cases. The scatter plot is accompanied by histograms of each count rate distribution. The solid black line indicates an exact count rate match between observation and modeled count rates and, as expected by the case average magnitude ratio of 2.23, most of the data points sit above this line. This is illustrated further by the red dashed lines which indicate the modal count rate bins, the EPIC modal value of 0.087 counts is approximately half the modal GUMICS-4 count rate of 0.199. This approximate factor of two is duplicated in the median of all data points (blue dashed line) where the EPIC value is 0.091 compared to the GUMICS-4 median value of 0.218. The actual factor of 2.4 is slightly different from the case average value of 2.23 due to the fact that each observation ranges in length from 3 hours to 26.5

hours. It should be noted that the histogram of varying O/H modeled count rates (not shown) resulted in a bi-modal distribution at 0.16 and 0.63 counts per second which is not repeated when the O/H ratio is kept constant.

5.3. Comparing SWCX quiet and enhanced times

As well as looking at the correlation value we can also compare the model and observed count rates to ensure that both are seeing more generalised X-ray emission enhancement at the same time. We determine the periods of enhancement during each case from the observed count rate lightcurve, given in *Carter et al.* [2011]. Each case shows a definitive period of enhancement which can be either from the start of the observation, near the end of the observation or sometime between the start and end. By determining these enhancement cut off times we separate out the observed and model count rates for each case into quiet and enhanced categories. Taking the mean count rate of both the quiet and enhanced periods, we can create both a ratio and difference value between the two for each case.

Figure 10 shows the values for the ratio between enhanced to quiet count rate. The top panel shows the cases in data order with each symbol representing; the EPIC observed ratio (green circle), the modeled ratio with an ACE varying O/H value (black star), and the modeled ratio with a constant O/H value (red plus). Placing all the case values in date order shows there are no temporal trends such as a decrease in accuracy with solar cycle or instrument degradation. The dashed line shows the 1:1 ratio with any points falling below the line indicating that the enhanced time is producing less flux. By definition, the EPIC observations all fall above the ratio line with a mean increase of 48% and median increase of 22% in c/s during SWCX enhancement times. The O/H varying cases have four cases where the ratio is less than 1, indicating that the SWCX enhanced period is returning less X-ray emission. Whereas, for the constant O/H there are only two cases where this occurs. The mean and median count rate increases for the O/H varying model are 370% and 53% respectively. The equivalent values for the constant O/H model are 116% and 96% respectively. While the varying O/H data provides a median increase between quiet and enhanced times similar to that seen in the observed data, the extremely high mean value indicates this is subject to high variation. The constant O/H ratio enhancement again shows a factor of two in both the mean and median enhancement rates. We investigate this further by plotting out each modeled enhancement ratio against the observed values in the lower two panels of Figure 10. The left panel shows the ACE varying enhancement ratio and the right panel shows the constant O/H enhancement ratio values, to be able to show both data sets on the same scale we have plotted these on a log x axis. The solid red line indicates the y=x line for ease of comparison. The variability of the results can once again be seen in the ACE varying model data although around the ratio of 1.5 the observations match up to the model extremely well. The constant O/H enhanced ratio values show a tighter spread but a reduced accuracy in the cases which matched well in the varying O/H plot.

The ratio between enhanced and quiet times will be very sensitive to the quiet time magnitudes, which in turn will be heavily influenced by the calculated background values. As a complement to the ratio calculation we have also determined the magnitude difference between enhanced and quiet times for each case, shown in Figure 11. The top panel shows the difference values in time order, again showing no temporal pattern between observations and model results. The mean and median enhancements are 0.08 and 0.06 counts per second for the observed differences, 0.05 and 0.04 for the ACE varying model, and 0.15 and 0.09 for the constant O/H model. The lower panels of Figure 11 show the scatter plot between observed and model differences, the solid red line in each plot shows the line of unity. The lower left plot also indicates the position of an outlying point at a model difference value of -0.63 c/s, this has been shown in blue. The ACE varying data shows a similar result to Figure 10 with a few cases correlating very well to the observed differences but the spread is wider than the constant O/H model data. The data from both the difference and ratio between enhanced and quiet times agrees, in some cases the ACE varying data does an excellent job while setting the O/H ratio to constant produces a more reliable result but reduces accuracy.

The case of April 17 2002 (02-2) has no enhanced to quiet ratio for either O/H value as the enhancement occurs in the final two bins of the observation and neither simulation returns counts for this period.

5.4. Positional accuracy

As a final piece of analysis we have also displayed the spatial position of the model data using a constant O/H ratio in Figure 12. The left panel shows the data in a cylindrical coordinate system (x-r) with the r axis signed by whether the y value is positve or negative. This view gives us positional values projected onto a 2D plane with a 0.5 R_E by 1 R_E resolution. We have binned all the data points and taken the average integral count rate for each bin, with our limited number of case studies this leaves a large proportion of the grid without any data but does show that the higher modeled count rates occur when the satellite is looking in the positive y direction (dusk). The top right panel of Figure 12 shows the data binned in the y-z plane, with no dependence upon the x value. We can again observe the asymmetry in y but the highest count rates occur at the z values closest to zero, as these values are likely to be closest to the nose of the magnetosheath it could simply be a proximity relation to the highest emission rates. To determine whether distance from the magnetopause is significant we plot each model count rate against the radial distance from the Shue magnetopause during the specific data point conditions. This scatter plot is shown in the middle right panel of Figure 12, with the data points split by y position. The positive y values are shown by black + signs and the negative y values are shown by blue * symbols. When looking at all the data points combined we can see that the count rate increases with distance from the magnetopause. This result is initially counterintuitive as we would expect the count rate to be higher the deeper in the magnetosheath the satellite is. What must be considered is the pointing direction and case selection bias. A case where the satellite is far from the magnetopause would only have shown initial significant SWCX if the pointing direction intersected a significant fraction of the magnetosheath. As the satellite comes closer to the Earth the integral path through the data grid includes fewer bins. If we took a sample of spacecraft positions with the spacecraft pointing in random directions then the opposite relation should be true. This magnitude plot shows, in a similar manner to the binned grid plots, that the count rates when y is positive are generally higher. The final plot of Figure 12, in the lower right panel, shows average distance from the magnetopause with correlation between the observed and modeled lightcurves. These data points are again split by whether y is positive or negative. There appears to be no general pattern between lightcurve variance and magnetopause distance, although the highest correlations occur in the negative y value (dawn) data values. This asymmetry was also mentioned in *Carter et al.* [2011], where they found that the empirical model fitted better in the dawnside. This dawn-dusk asymmetry could be related to the known asymmetries in either the magnetosheath plasma conditions [e.g. *Walsh et al.*, 2012] or magnetopause position [e.g. *Dmitriev et al.*, 2004], indicating that this asymmetry needs to be considered during the modeling process.

6. Conclusions

In this study we have taken the data from 19 case studies using the EPIC-MOS instruments on XMM-Newton to examine the accuracy of MHD modeling when describing solar wind charge exchange from the Earth's magnetosheath. We found that a large amount of variation in the modeled lightcurve was caused by variations in the oxygen to hydrogen ratio and abundances of oxygen charge states. In a large number of these cases setting the oxygen to hydrogen ratio to a constant improved the variance matching. These modeled data values with a constant O/H ratio and mean charge state abundances were then compared to the observed lightcurves, providing an average correlation value of 0.35. This correlation has been reduced by the fact that five of the nineteen cases are anti-correlated. The average magnitude ratio is a factor of 2.4 when averaging across all data points, giving 42% of the cases having an average magnitude within a factor of two of the observed data values, a slight decrease on the empirical method used in *Carter et al.* [2011]. The highest modeled count rates occur when the satellite is in the positive GSE y, with the highest correlations arising in negative y (dawn).

It is clear from Sections 4.2 and 5 that the oxygen data inputs to the MHD model include substantial errors. The O/H variances cause large changes in the modeled lightcurves which are simply not seen in the observed lightcurves for a significant number of cases. The longer (temporally) the case is, the more likely that a constant O/H ratio is inappropriate, yet accurate data is needed. The same applies to the oxygen charge state abundances, the two hour resolution of this data is low for modeling that runs at a four second calculation resolution and a five minute grid output. The absolute abundance values themselves are also an issue, it is unknown what the upper and lower limits of O^{7+} and O^{8+} should be. We observed in Figure 5 that the O^{7+} abundance can take a wide range of values, up to 52% which is likely unphysical. It is certain that the values given in *Schwadron and Cravens* [2000], while of the right order of magnitude, are of limited use for this particular modeling, especially as they describe almost no highly charged states in the fast wind. To improve on model accuracy we either require more accurate and numerous solar wind oxygen observations closer to the Earth, or an accurate proxy such as an extension of the proton entropy correlation work by *Pagel et al.* [2004] to include the O^{8+}/O^{7+} ratio. Using a constant value for the O/H ratio of 2×10^{-4}, and mean oxygen charge states for each case we have removed a large amount of this variation at the cost of a small accuracy loss (e.g., Figures 10 and 11).

The accuracy of the MHD modeling ranges from anti-correlated to an excellent correlation. We can link several of the errors in both magnitude and variance to the oxygen data, and the disparity of the MHD magnetopause position to the Shue model. The other data inputs to the MHD model behave well and we have some excellent comparisons as seen in Figures 7 and 8. When comparing the MHD model to the empirical model used in *Carter et al.* [2011] we can say that it performs equally well. The slight decrease of magnitude matching, 42% rather than 50%, of cases within a factor of two, could easily be due to the background removal. Decreasing the background removal values by 0.03 c/s actually increases the magnitude comparison accuracy to 63%, hence showing the importance of the background removal when we look at comparing the magnitudes. The background removal does not affect the correlation or the enhanced to quiet differencing comparison however. Examining the magnitude difference between quiet and enhanced periods, we see very similar results between the observed values and the MHD model. The difference in the dawn-dusk correlations, also seen in *Carter et al.* [2011], suggests that there could be an asymmetrical process affecting the charge exchange emission magnitude, which is missing from both models.

Users wishing to estimate the near-Earth SWCX values are advised that using either the empirical model or an MHD model with constant solar wind oxygen parameters are equally likely to produce a useable value. When comparing enhanced to quiet times, i.e., taking an average over a longer time period, using the variable O/H data is likely to be a valid approach. For those interested in a more in depth view of what is happening in terms of global SWCX around the Earth, the MHD based model with a constant oxygen ratio and abundances, will produce a more accurate result, including matching short time scale emissivity variation. This study also acts as a validation of the model methodology for global imaging of the magnetosheath using SWCX, by providing similar emissivities to observed values. However, the relative inaccuracy of using a far upstream monitor for the solar wind conditions can affect the model results considerably. This modeling will be especially important for future missions involving wide angle X-ray imaging of the Earth's magnetosheath.

Acknowledgments. This research used the ALICE/SPECTRE High Performance Computing Facility at the University of Leicester. The GUMICS-4 MHD code is provided by the Finnish Meteorological Institute, while the simulation runs are stored on the HPC facilities at the University of Leicester. The OMNI solar wind data was taken from NASA's GSFC/SPDF repository (*omniweb.gsfc.nasa.gov*). The ACE solar wind composition data was taken from the ACE Science Center *www.srl.caltech.edu/ACE/ASC/*. The background count rates were calculated using the WebPIMMS service (*heasarc.gsfc.nasa.gov/cgi-bin/Tools/w3pimms/w3pimms.pl*). Effort sponsored by the Air Force Office of Scientific Research, Air Force Material Command, USAF, under grant number FA9550-14-1-0200. The U.S. government is authourized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. JAC and SEM gratefully acknowledge support from the STFC consolidated grant ST/K001000/1.

References

Bailey, J. and M. Gruntman. (2011), Experimental study of exospheric hydrogen atom distributions by Lyman-alpha detectors on the TWINS mission, *J. Geophys. Res.*, **116** A09302, doi:10.1029/2011JA016531

Bodewits, D. (2007), PhD thesis, University of Groningen, The Netherlands.

Carter, J. A., and A. M. Read (2007), The XMM-Newton EPIC background and the production of background blank sky event files, *A&A*, **464**, doi:10.1051/0004-6361:20065882

Carter, J. A. and S, Sembay (2008), Identifying XMM-Newton observations affected by solar wind charge exchange - Part I, *A&A*, **489** 2, doi:10.1051/0004-6361/200809997

Carter, J. A., S, Sembay and A. Read (2010), A high charge state coronal mass ejection seen through solar wind charge exchange emission as detected by XMM-Newton. *Mon. Not. R. Astron. Soc.*, **402**, doi:10.1111/j.1365-2966.2009.15985.x

Carter, J. A., S, Sembay and A. Read (2011), Identifying XMM-Newton observations affected by solar wind charge exchange - Part II, *A&A*, **527** A115, doi:10.1051/0004-6361/201015817

Collier, M. R., T. E. Moore, M.-C. Fok, B. Pilkerton, S. Boardsen and H. Khan. (2005), Low-energy neutral atom signatures of magnetopause motion in response to southward Bz, *J. Geophys. Res.*, **110** A2, doi:10.1029/2004JA010626

Collier, M. R., D. G. Siebeck, T. E. Cravens, I. P. Robertson and N. Omidi. (2010), Astrophysics Noise: A Space Weather Signal, *Eos, Transactions American Geophysical Union*, **91** 24, doi:10.1029/2010EO240001

Collier, M. R., F. S. Porter, D. G. Sibeck, J. A. Carter, M. P. Chiao, D. J. Chornay, T. E. Cravens, M. Galeazzi, J. W. Keller, D. Koutroumpa, K. Kuntz, A. M. Read, I. P. Robertson, S. Sembay, S. Snowden and N. Thomas. (2012), Prototyping a global soft X-ray imaging instrument for heliophysics, planetary science, and astrophysics science, *Astron. Nachr.*, **333** 4, doi:10.1002/asna.201211662

Collier, M. R., S. L. Snowden, M. Sarantos, M. Benna, J. A. Carter, T. E. Cravens, W. M. Farrell, S. Fatemi, H. Kent-Hills, R. R. Hodges, M. Holmström, K. D. Kuntz, F. Scott-Porter, A. Read, I. P. Robertson, S. F. Sembay, D. G. Sibeck, T. J. Stubbs, P. Travnicek and B. M. Walsh. (2014), On lunar exospheric column densities and solar wind access beyond the terminator from ROSAT soft X-ray observations of solar wind charge exchange. *J. Geophys. Res.*, **119** 7, doi:10.1002/2014JE004628

Cravens, T. E. (1997), Comet Hyakutake X-ray source: Charge transfer of solar wind heavy ions. *Geophys Res. Lett.*, **24** 105109, doi:10.1029/96GL03780

Cravens, T. E. (2000), Heliospheric X-ray emission associated with charge transfer of the solar wind with interstellar neutrals. *ApJ*, **532** 2, doi:10.1086/312574

Cravens, T. E., I. P. Robertson and S. L. Snowden (2001), Temporal variations of geocoronal and heliospheric X-ray emission associated with the solar wind interaction with neutrals. *J. Geophys. Res.*, **106** A11, doi:10.1029/2000JA000461

Dennerl, K. (2008), X-rays from Venus observed with Chandra. *PSS*, **56** 10, doi:10.1016/j.pss.2008.03.008

Dimmock, A. P., K. Nykyri, H. Karimabadi, A. Osmane, and T. I. Pulkkinen. (2015), A statistical study into the spatial distribution and dawn-dusk asymmetry of dayside magnetosheath ion temperatures as a function of upstream solar wind conditions, *J. Geophys. Res.*, **120** 4, doi:10.1002/2014JA020734

Dmitriev, A. V., A. V. Suvorova, J. K. Chao and Y.-H. Yang (2004), Dawn-dusk asymmetry of geosynchronous magnetopause crossings, *J. Geophys. Res.*, **109** A5, doi:10.1029/2003JA010171

Geiss, J., G. Gloeckner and R. von Steiger (1999), Origin of the solar wind from composition data, *Space Sci. Rev.*, **72** 49-60

Gordeev, E., G. Facskó, V. Sergeev, I. Honkonen, M. Palmroth, P. Janhunen and S. Milan. (2013), Verification of the GUMICS-4 global MHD code using empirical relationships, *J. Geophys. Res.*, **118** 6, doi:10.1002/jgra.50359

Henley, D. B. and R. L. Shelton (2008), Comparing Suzaku and XMM-Newton Observations of the soft X-ray background: Evidence for solar wind charge exchange emission. *ApJ*, **335** doi:10.1086/528924

Hodges, R. R. Jr. (1994), Monte Carlo simulation of the terrestrial hydrogen exosphere, *J. Geophys. Res.*, **99** A12, doi:10.1029/94JA02183

Holmström, M., S. Barabash and E. Kallio (2001), X-ray imaging of the solar windMars interaction, *GRL*, **28** 7, doi:10.1029/2000GL012381

Hosokawa, K., S. Taguchi, S. Suzuki, M. R. Collier, T. E. Moore and M. F. Thomsen. (2008), Estimation of magnetopause motion from low-energy neutral atom emission, *J. Geophys. Res.*, **113** A10, doi:10.1029/2008JA013124

Hubert, B., M. Palmroth, T. V. Laitinen, P. Janhunen, S. E. Milan, A. Grocott, S. W. H. Cowley, T. Pulkkinen and J-C. Gérard. (2006), Compression of the Earth's magnetotail by interplanetary shocks directly drives transient magnetic flux closure, *Geophys. Res. Lett.*, **33** 10, doi:10.1029/2006GL026008

Ishikawa, K., Y. Ezoe, Y. Miyoshi, N. Terada, K. Mitsuda and T. Ohashi (2013), Suzaku Observation of Strong Solar-Wind Charge-Exchange Emission from the Terrestrial Exosphere during a Geomagnetic Storm, *Publ. Astron. Soc. Jpn.*, **65** 3, doi:10.1093/pasj/65.3.63

Janhunen, P., M. Palmroth, T. Latinen, I. Honkonen, L. Juusola, G. Facskó and T. I. Pulkkinen. (2012), The GUMICS-4 global MHD magnetosphere-ionosphere coupling simulation, *J. Atmos. Sol. Terr. Phys.*, **80**, doi:10.1016/j.jastp.2012.03.006

Jansen, F., D. Lumb, B. Altieri, J. Clavel, M. Ehle, C. Erd, C. Gabriel, M. Guainazzi, P. Gondoin, R. Much, R. Munoz, M. Santos, N. Schartel, D. Texier and G. Vacanti. (2001), XMM-Newton observatory: I. The spacecraft and operations, *A&A*, **365** L1-L6, doi:10.1051/0004-6361:20000036

Kalberla, P. M. W., W. B. Burton, D. Hartmann, E. M. Arnal, E. Bajaja, R. Morras and W. G. L. Pöppel. (2005), The Leiden/Argentine/Bonn (LAB) Survey of Galactic HI. Final data release of the combined LDS and IAR surveys with improved stray-radiation corrections, *A&A*, **440** 2, doi:10.1051/0004-6361:20041864

King, J. H., and N. E. Papitashvili (2005), Solar wind spatial scales in and comparisons of hourly Wind and ACE plasma and magnetic field data *J. Geophys. Res.*, **110**, doi:10.1029/2004JA010649

Koutroumpa, D. (2012), Update on modeling and data analysis of heliospheric solar wind charge exchange X-ray emission, *Astron. Nachr.*, **333** 341346, doi:10.1002/asna.201211666

Kuntz, K. D. and S. L. Snowden (2008), The EPIC-MOS particle induced background spectra *A&A*, **478** 2, doi:10.1051/0004-6361:20077912

Kuntz, K. D., Y. M. Collado-Vega, M. R. Collier, H. K. Connor, T. E. Cravens, D. Koutroumpa, F. S. Porter, I. P. Robertson, D. G. Sibeck and S. L. Snowden (2015), The solar wind charge-exchange production factor for hydrogen. *ApJ*, **808** 2, doi:10.1088/0004-637X/808/2/143

Liemohn, M. W., J. U. Kozyra, V. K. Jordanova, G. V. Khazanov, M. F. Thomsen and T. E. Cayton. (1999), Analysis of early phase ring current recovery mechanisms during geomagnetic storms, *Geophys. Res. Lett.*, **26** 18, doi:10.1029/1999GL900611

Lisse, C.M., K. Dennerl, J. Englhauser, M. Harden, F. E. Marshall, M. J. Mumma, R. Petre, J. P. Pye, M. J. Ricketts, J. Schmitt, J. Trümper, R. G. West (1996), Discovery of X-ray and Extreme Ultraviolet Emission from Comet C/Hyakutake 1996 B2 *Science*, **274** 5285, doi: 10.1126/science.274.5285.205

Milan, S. E., S. W. H. Cowley, M. Lester, D. M. Wright, J. A. Slavin, M. Fillingim, C. W. Carlson, and H. J. Singer (2004), Response of the magnetotail to changes in the open flux content of the magnetosphere, *J. Geophys. Res.*, **109** A04220, doi:10.1029/2003JA010350

Nass, H. U., J. H. Zoennchen, G. Lay and H. J. Fahr. (2006), The TWINS-LAD mission: Observations of terrestrial Lyman-α fluxes, *Astrophys. Space. Sci. Trans.*, **2** 27-31, doi:10.5194/94JA02183

Ogasawara, K., V. Angelopoulos, M. A. Dayeh, S. A. Fuselier, G. Livadiotis, D. J. McComas and J. P. MCFadden. (2013), Characterizing the sayside magnetosheath using energetic neutral atoms: IBEX and THEMIS observations, *J. Geophys. Res.*, **118** 6, doi:10.1002/jgra.50353

Østgaard, N., S. B. Mende, H. U. Frey, G. R. Gladstone and H. Lauche. (2003), Neutral hydrogen density profiles derived from geocoronal imaging, *J. Geophys. Res.*, **108** A7, doi:10.1029/2002JA009749

Pagel, A. C., N. U. Crooker, T. H. Zurbuchen and J. T. Gosling. (2004), Correlation of solar wind entropy and oxygen ion charge state ratio, *J. Geophys. Res.*, **109** A1, doi:10.1029/2003JA010010

Palmroth, M., I. Honkonen, A. Sandroos, Y. Kempf, S. von Alfthan and D. Pokhotelov. (2013), Preliminary testing of global hybrid-Vlasov simulation: Magnetosheath and cusps under northward interplanetary magnetic field. *J. Atmos. Sol. Terr. Phys.*, **99**, doi:10.1016/j.jastp.2012.09.013

Palmroth, M., T. I. Pulkkinen, P. Janhunen and C.-C Wu. (2003), Stormtime energy transfer in global MHD simulation. *J. Geophys. Res.*, **108**, A1, doi:10.1029/2002JA009446

Powell, K. G., P. L. Roe, T. J. Linde, T. I. Gambosi and D. L. de Zeeuw. (1999), A Solution-Adaptive Upwind Scheme for Ideal Magnetohydrodynamics. *J. Comp. Phys.*, **154** 2, doi:10.1006/jcph.1999.6299

Robertson, I. P. and T. E. Cravens. (2003), Spatial maps of heliospheric and geocoronal X-ray intensities due to the charge exchange of the solar wind with neutrals, *J. Geophys. Res.*, **108** A10, doi:10.1029/2003JA009873

Robertson, I. P., M. R. Collier, T. E. Cravens and M.-C. Fok. (2006), X-ray emission from the terrestrial magnetosheath including the cusps, *J. Geophys. Res.*, **111** A12105, doi:10.1029/2006JA011672

Schwadron, N. A. and T. E. Cravens (2000), Implications of solar wind composition for cometary X-rays, *ApJ*, **554** 558:566

Shue, J. H., P. Song, C. T. Russell, J. T. Steinberg, J. K. Chao, G. Zastenker, O. L. Vaisberg, S. Kokubun, H. J. Singer, T. R. Detman and H. Kawano (1998), Magnetopause location under extreme solar wind conditions, *J. Geophys. Res.*, **103** A8, doi:10.1029/98JA01103

Slavin, J. D., B. J. Wargelin and D. Koutroumpa (2013), Solar wind charge exchange emission in the Chandra deep field north, *ApJ*, **779** 1, doi:10.1088/0004-637X/779/1/13

Snowden, S. L., M. R. Collier and K. D. Kuntz (2004), XMM-Newton Observation of Solar Wind Charge Exchange Emission, *ApJ*, **610** 2 1182-1190, doi:10.1086/421841

Spreiter, J. R., A. L. Summers and A. Y. Alksne (1966), Hydromagnetic flow around the magnetosphere, *PSS*, **14** 223, doi:10.1016/0032-0633(66)90124-3

Strüder, L., U. Briel, K. Dennerl, R. Hartmann, E. Kendziorra, N. Meidinger, E. Pfeffermann, C. Reppin, B. Aschenbach, W. Bornemann, H. Bräuninger, W. Burkert, M. Elender, M. Freyberg, F. Haberl, G. Hartner, F. Heuschmann, H. Hippmann, E. Kastelic, S. Kemmer, G. Kettenring, W. Kink, N. Krause, S. Müller, A. Oppitz, W. Pietsch, M. Popp, P. Predehl, A. Read, K. H. Stephan, D. Stötter, J. Trümper, P. Holl, J. Kemmer, H. Soltau, R. Stötter, U. Weber, U. Weichert, C. von Zanthier, D. Carathanassis, G. Lutz, R. H. Richter, P. Solc, H. Böttcher, M. Kuster, R. Staubert, A. Abbey, A. Holland, M. Turner, M. Balasini, G. F. Bignami, N. La Palombara, G. Villa, W. Buttler, F. Gianini, R. Lainé, D. Lumb and P. Dhez. (2001), The European Photon Imaging Camera on XMM-Newton: The pn-CCD camera, *A&A*, **365** L18-L26, doi:10.1051/0004-6361:20000066

Turner, M. J. L., A. Abbey, M. Arnaud, M. Balasini, M. Barbera, E. Belsole, P. J. Bennie, J. P. Bernard, G. F. Bignami, M. Boer, U. Briel, I. Butler, C. Cara, C. Chabaud, R. Cole, A. Collura, M. Conte, A. Cros, M. Denby, P. Dhez, G. Di Coco, J. Dowson, P. Ferrando, S. Ghizzardi, F. Gianotti, C. V. Goodall, L. Gretton, R. G. Griffiths, O. Hainaut, J. F. Hochedez, A. D. Holland, E. Jourdain, E. Kendziorra, A. Lagostina, R. Laine, N. La Palombara, M. Lortholary, D. Lumb, P. Marty, S. Molendi, C. Pigot, E.

Poindron, K. A. Pounds, J. N. Reeves, C. Reppin, R. Rothenflug, P. Salvetat, J. L. Sauvageot, D. Schmitt, S. Sembay, A. D. T. Short, J. Spragg, J. Stephen, L. Strüder, A. Tiengo, M. Trifoglio, J. Trümper, S. Vercellone, L. Vigroux, G. Villa, M. J. Ward, S. Whitehead and E. Zonca. (2001), The European Photon Imaging Camera on XMM-Newton: The MOS cameras, *A&A*, **365** L27-L35, doi:10.1051/0004-6361:20000087

Voges, W., B. Aschenbach, T. Boller, H. Bräuninger, U. Briel, W. Burkert, K. Dennerl, J. Englhauser, R. Gruber, F. Haberl, G. Hartner, G. Hasinger, M. Kürster, E. Pfeffermann, W. Pietsch, P. Predehl, C. Rosso, J. H. M. M. Schmitt, J. Trümper and H. U. Zimmermann (1999), The ROSAT all-sky survey bright source catalogue, *A&A*, **349** 2, 389-405

Walsh, B. M., D. G. Sibeck, Y. Wang and D. H. Fairfield (2012), Dawn-dusk asymmetries in the Earth's magnetosheath, *J. Geophys. Res.*, **117** A12, doi:10.1029/2012JA018240

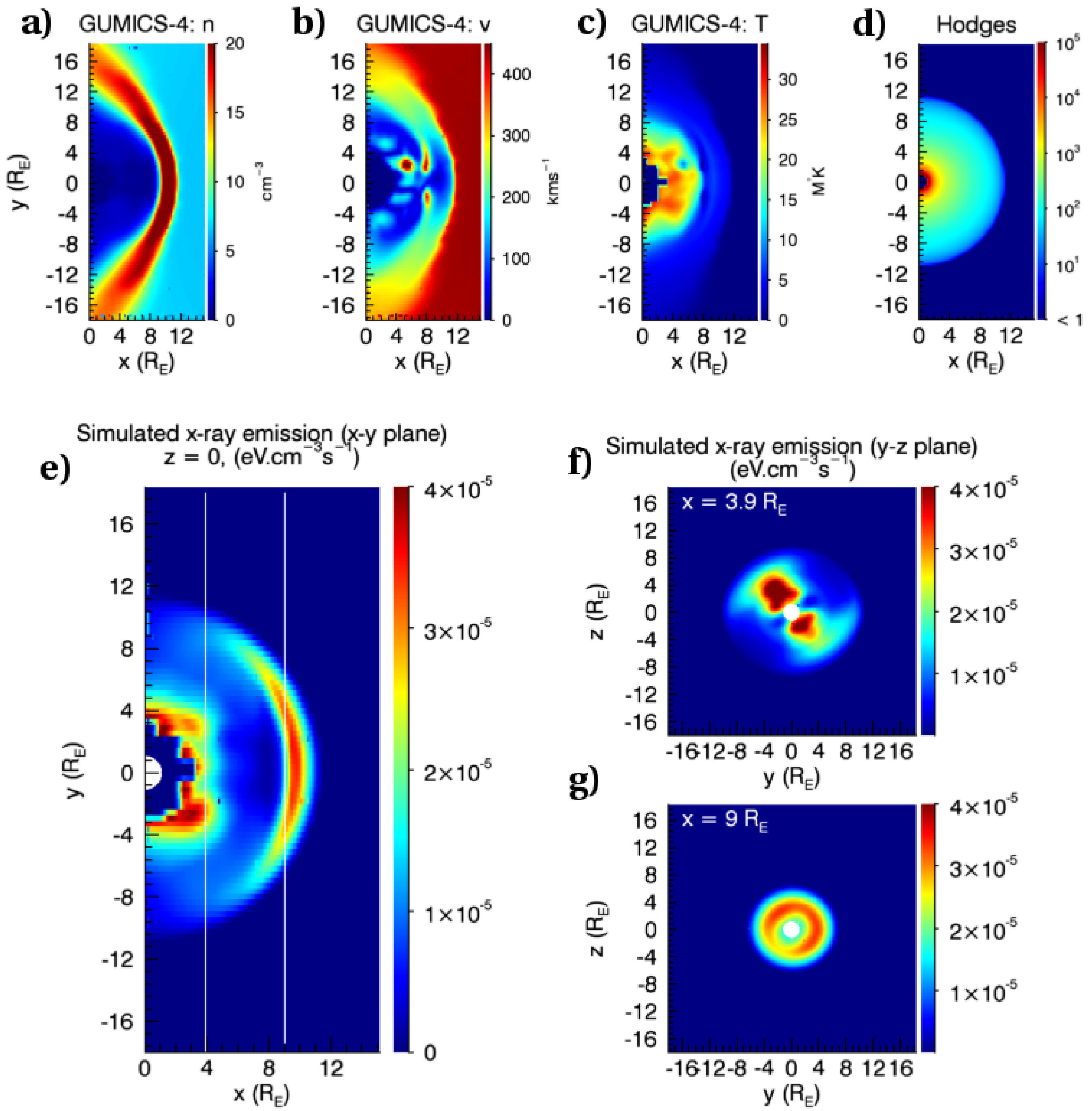

Figure 1. **a) - d)** The GUMICS-4 output for a single timestep showing solar wind proton number density, bulk flow speed and temperature with the final panel showing the equivalent Hodges neutral hydrogen density. Each of these panels shows a slice through the data cube in the x-y plane at $z = 0$. **e)** The calculated X-ray emissivity in the x-y plane, with cuts taken to show the y-z plane at $x = 3.9\mathrm{R}_E$, the magnetopause sits around a subsolar distance of $9\mathrm{R}_E$ and the bow shock is at approximately $11\mathrm{R}_E$.

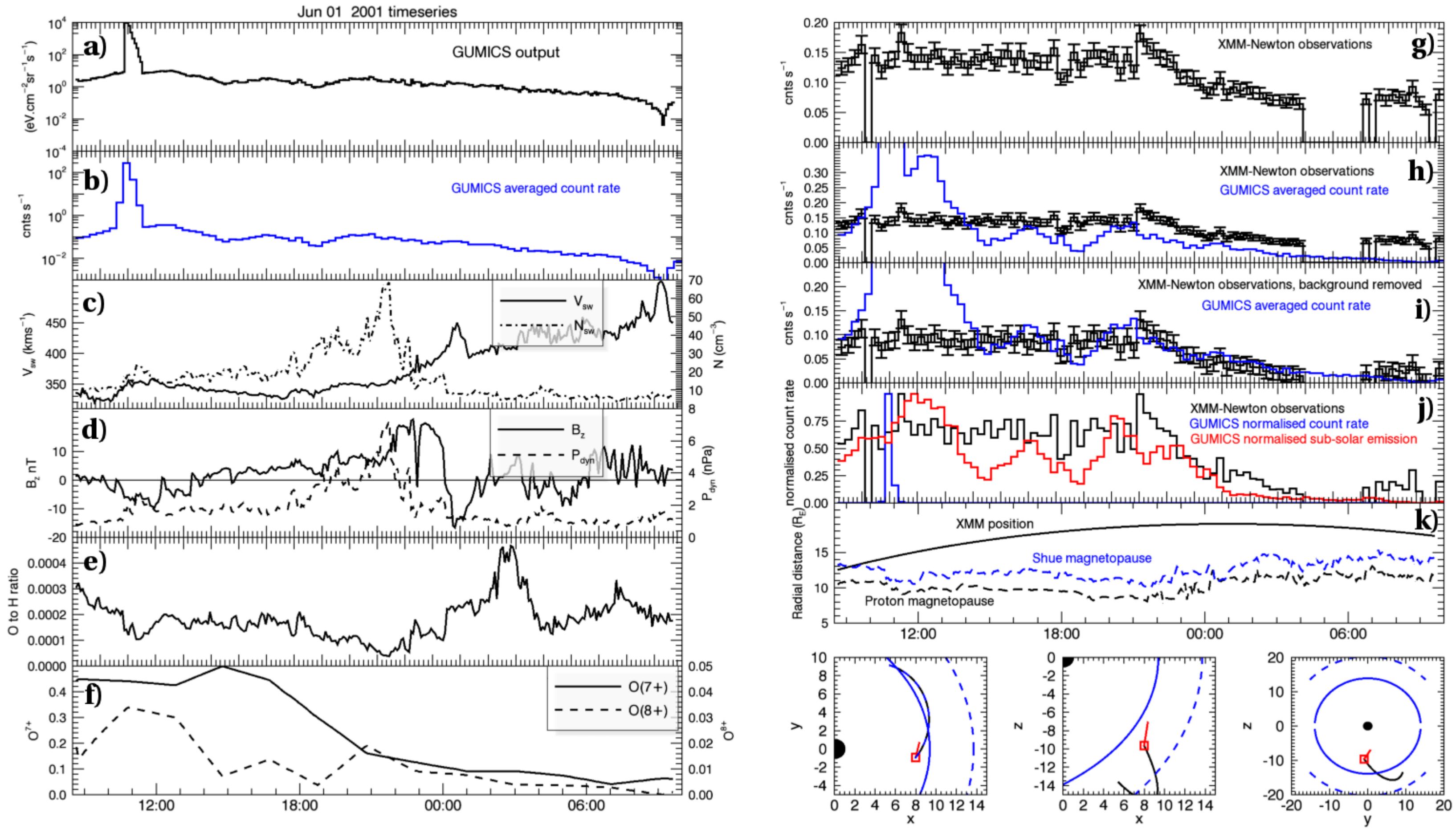

Figure 2. A set of plots showing all the specific details for an example case (01-1). The panels show; **a)** modeled X-ray integral emission, **b)** the equivalent count rates, **c)** solar wind density and speed, **d)** B_z and dynamic pressure, **e)** O/H ratio, **f)** oxygen charge state abundances, **g)** EPIC observation data, **h)** comparison of observed and modeled count rates including background, **i)** comparison of observed and modeled count rates with the background removal applied, **j)** normalized lightcurves, **k)** XMM position relative to the magnetopause, and the final panels show the orbital position.

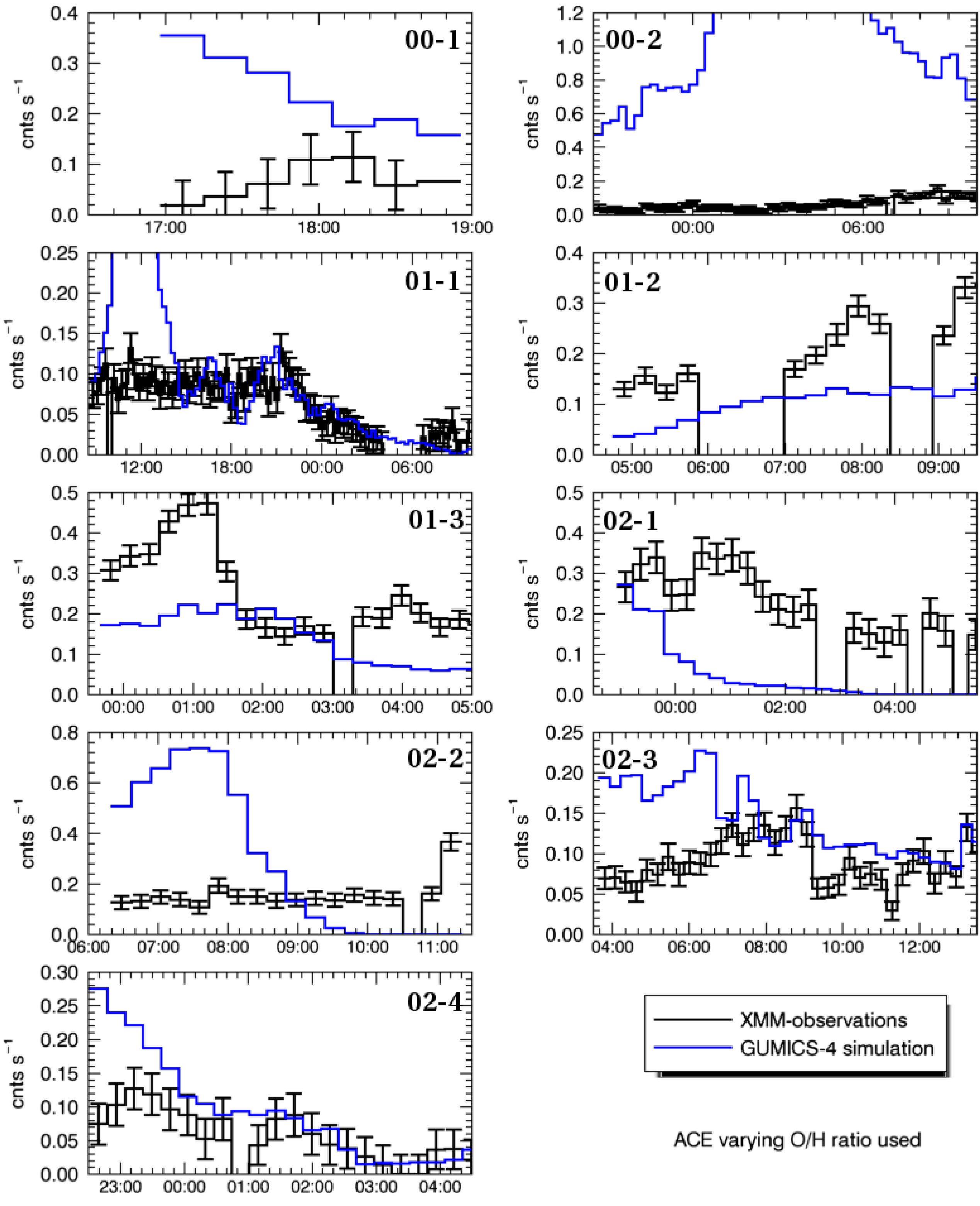

Figure 3. Cases 1-9 of the study, each panel shows the GUMCIS-4 estimated count rate in blue and the EPIC observations in black with the combined observational and background error bars included.

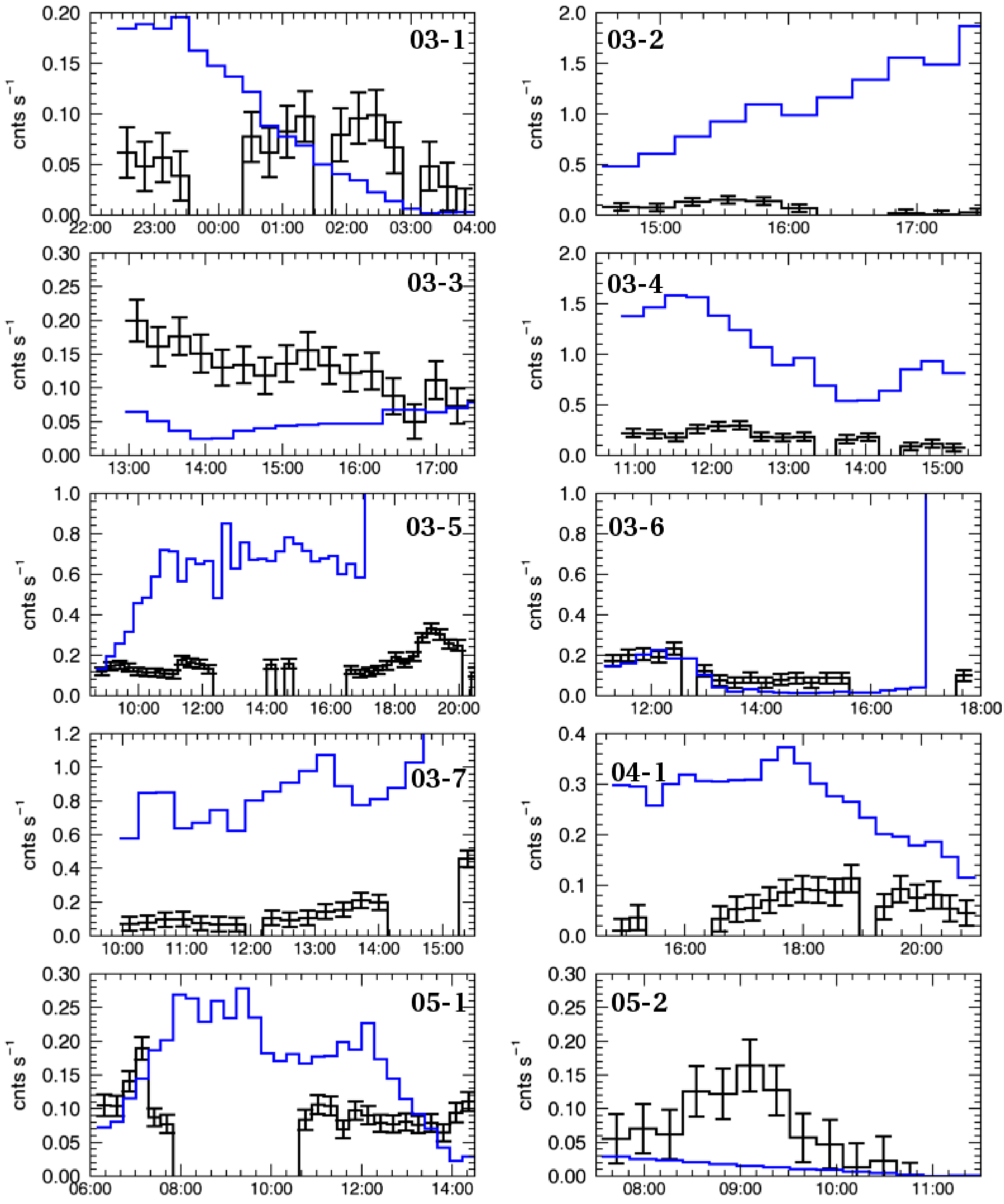

Figure 4. Cases 10-19 of the study, each panel shows the GUMCIS-4 estimated count rate in blue and the EPIC observations in black with the combined observational and background error bars included.

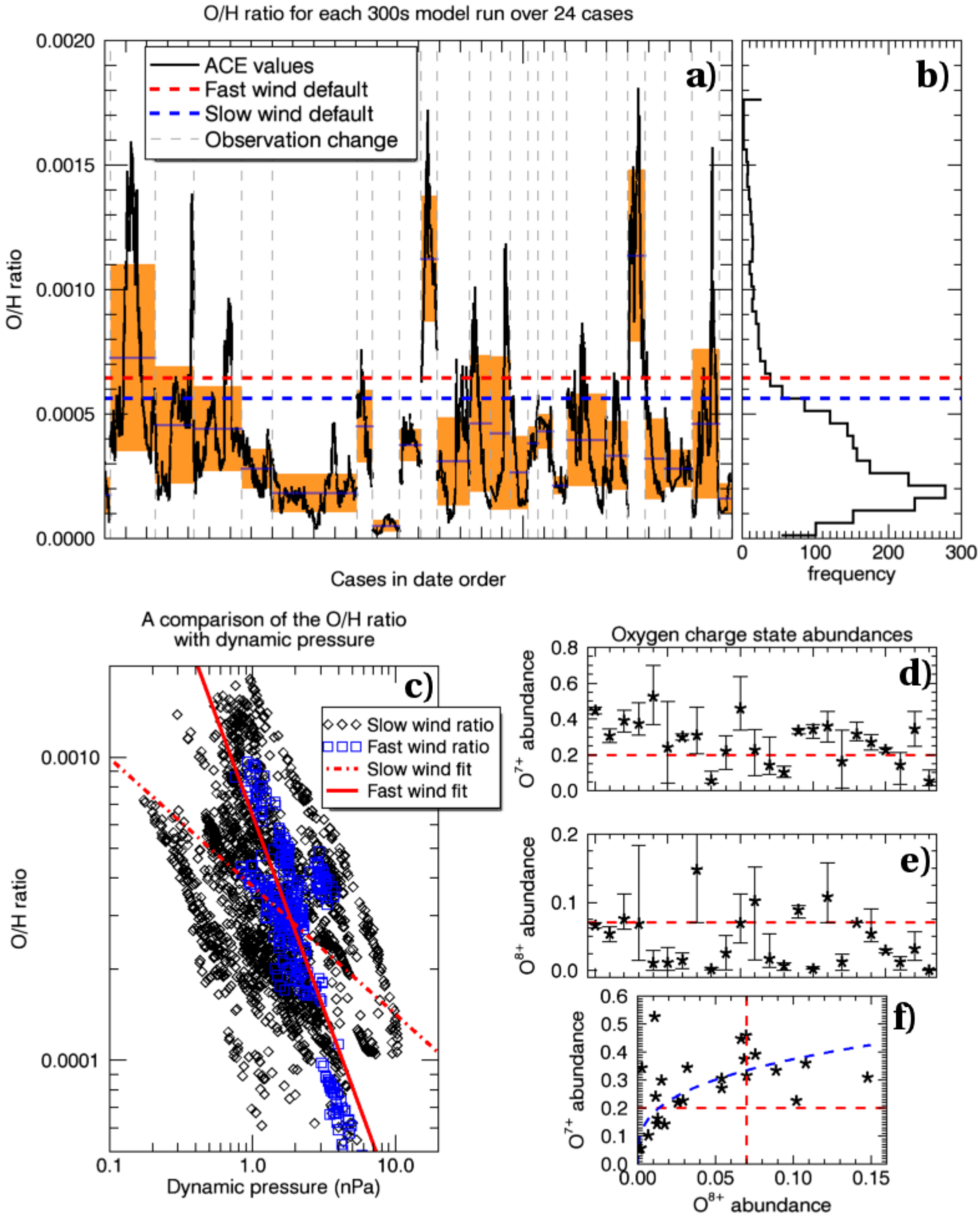

Figure 5. A selection of plots showing the variation in oxygen to hydrogen ratio and the oxygen charge state abundance. **a)** The O/H ratio for every OMNI data point used in the 24 case studies with data, **b)** the ratio in a histogram format. **c)** The slow and fast O/H ratio against solar wind dynamic pressure. **d)** and **e)** The average O^{7+} and O^{8+} charge state abundance for each case. **f)** The O^{7+} charge state abundance plotted against the O^{8+} abundance.

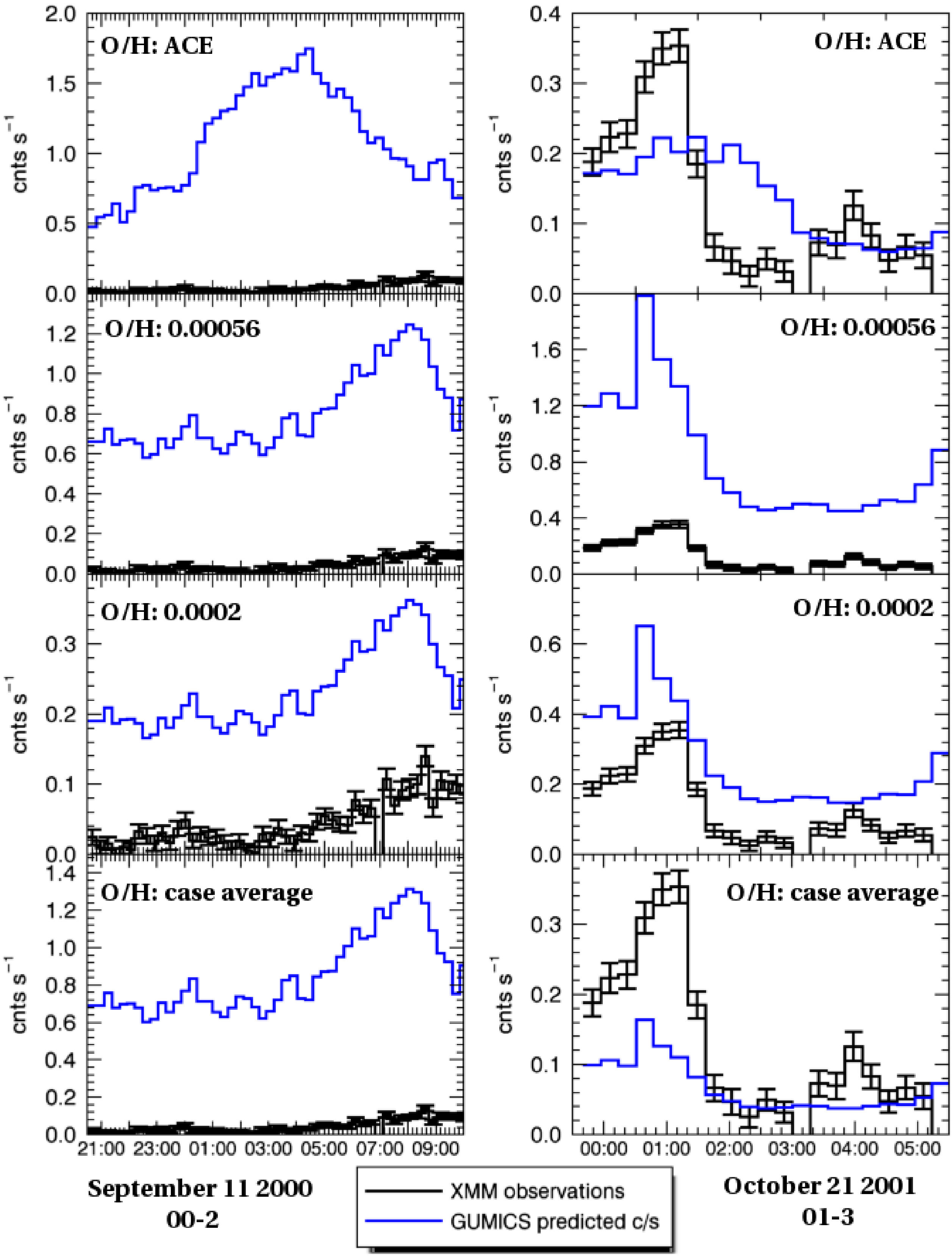

Figure 6. A comparison of changing the oxygen to hydrogen ratio. Each panel shows the XMM-Newton observations for cases 00-2 and 01-3 (in black with errorbars) as well as the GUMICS simulation in blue. Keeping the O/H ratio constant produces a variation in the simulated light curve closer to the observed values, in comparison to the ACE varying O/H ratio in the top panels.

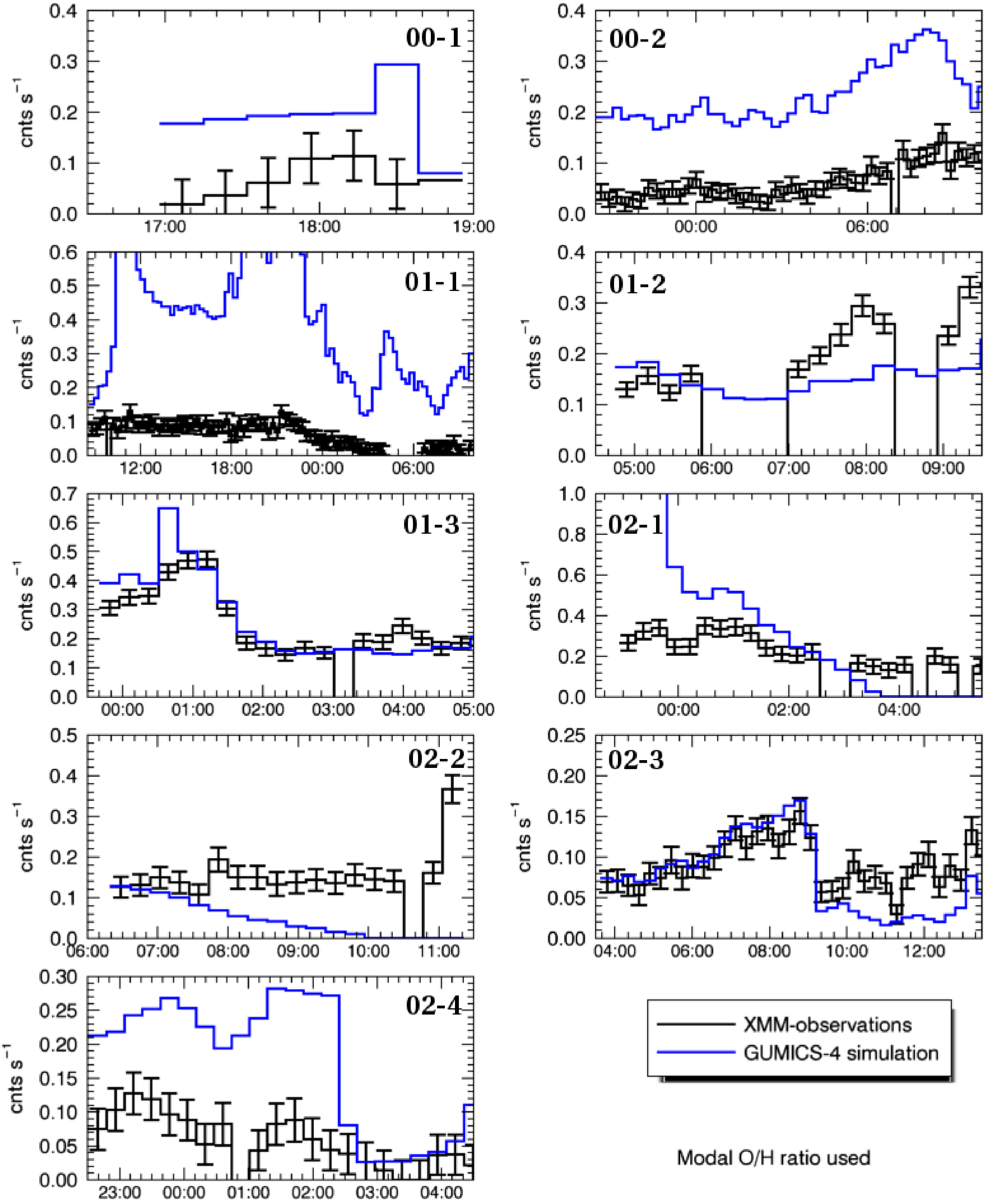

Figure 7. Cases 1-9 of the study using a constant O/H ratio, each panel shows the GUMCIS-4 estimated count rate in blue and the EPIC observations in black with the combined observational and background error bars included.

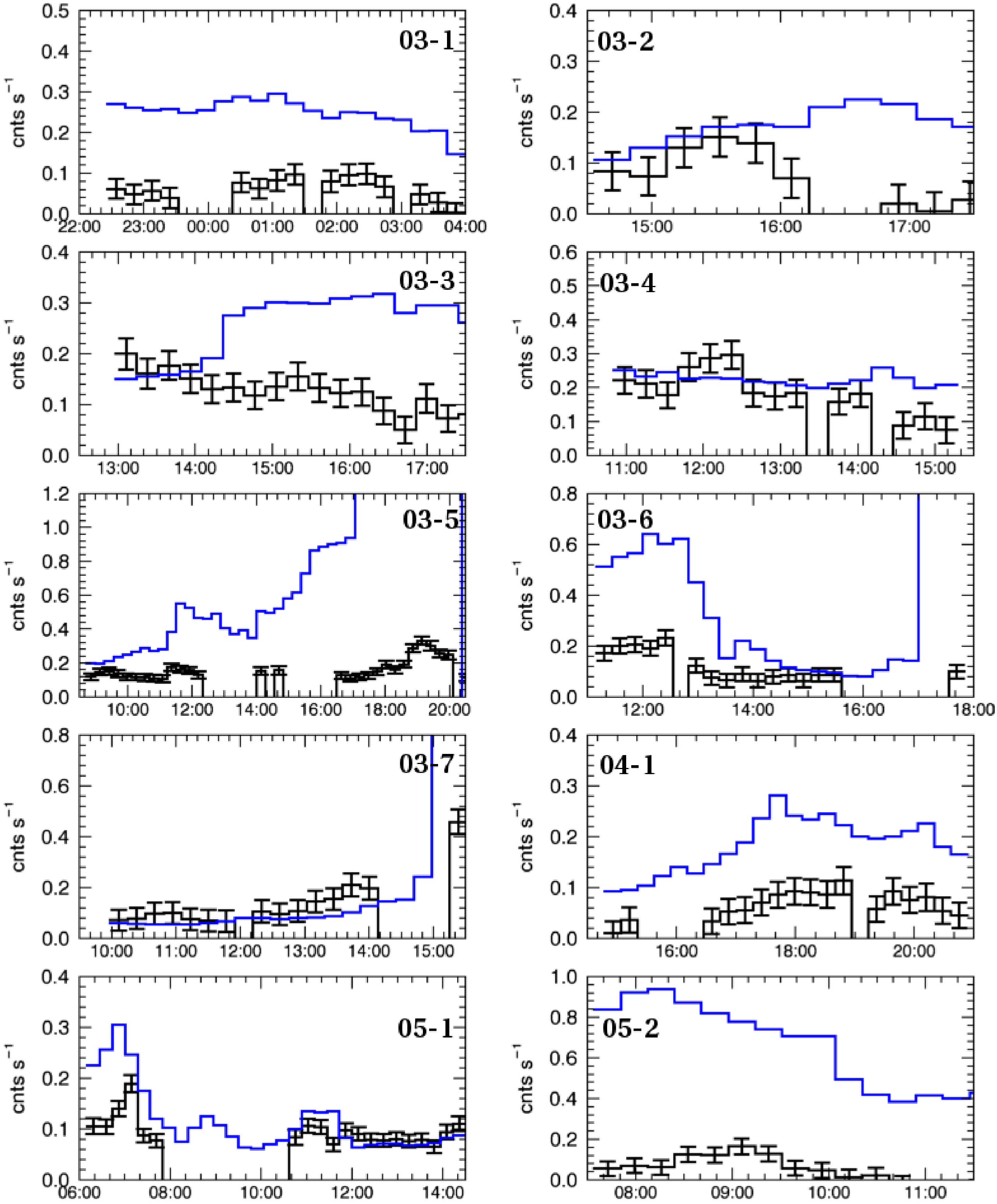

Figure 8. Cases 10-19 of the study using a constant O/H ratio, each panel shows the GUMCIS-4 estimated count rate in blue and the EPIC observations in black with the combined observational and background error bars included.

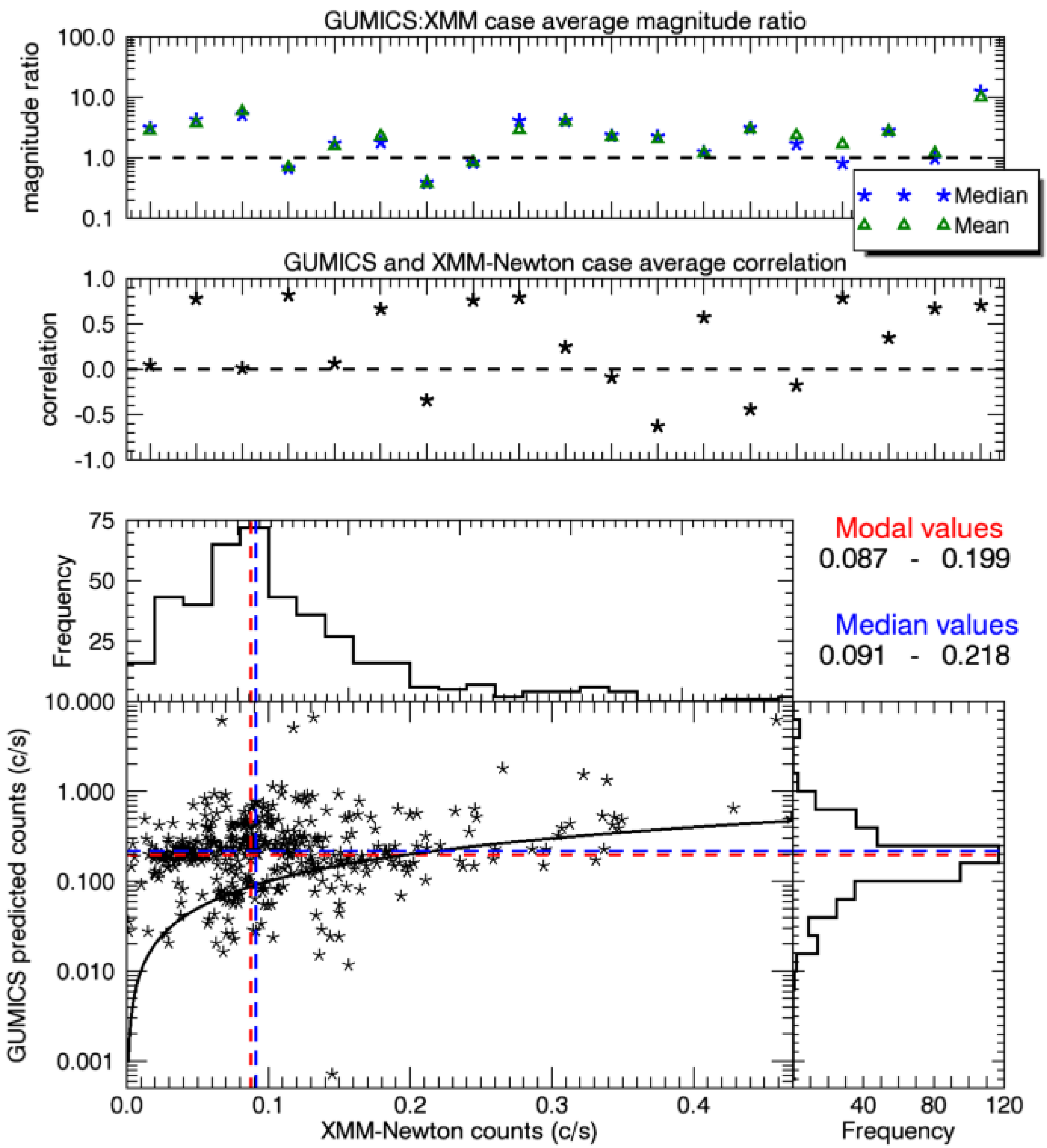

Figure 9. A comparison of the observed and modeled count rates for each case, using the constant modal O/H value from Figure 5. The top panel uses the mean and median magnitude for each case and takes the ratio of the modeled to the observed magnitudes, with the dashed line showing the equal magnitude line. Data points below this line show a lower modeled than observed average magnitude. The middle panel shows the correlation between modeled and observed counts. The lower panel is a scatter point of all data points, with histograms included for both the modeled and observed count distributions. The red and blue dashed lines show the modal and median lines respectively, the solid black dashed line shows the y=x line.

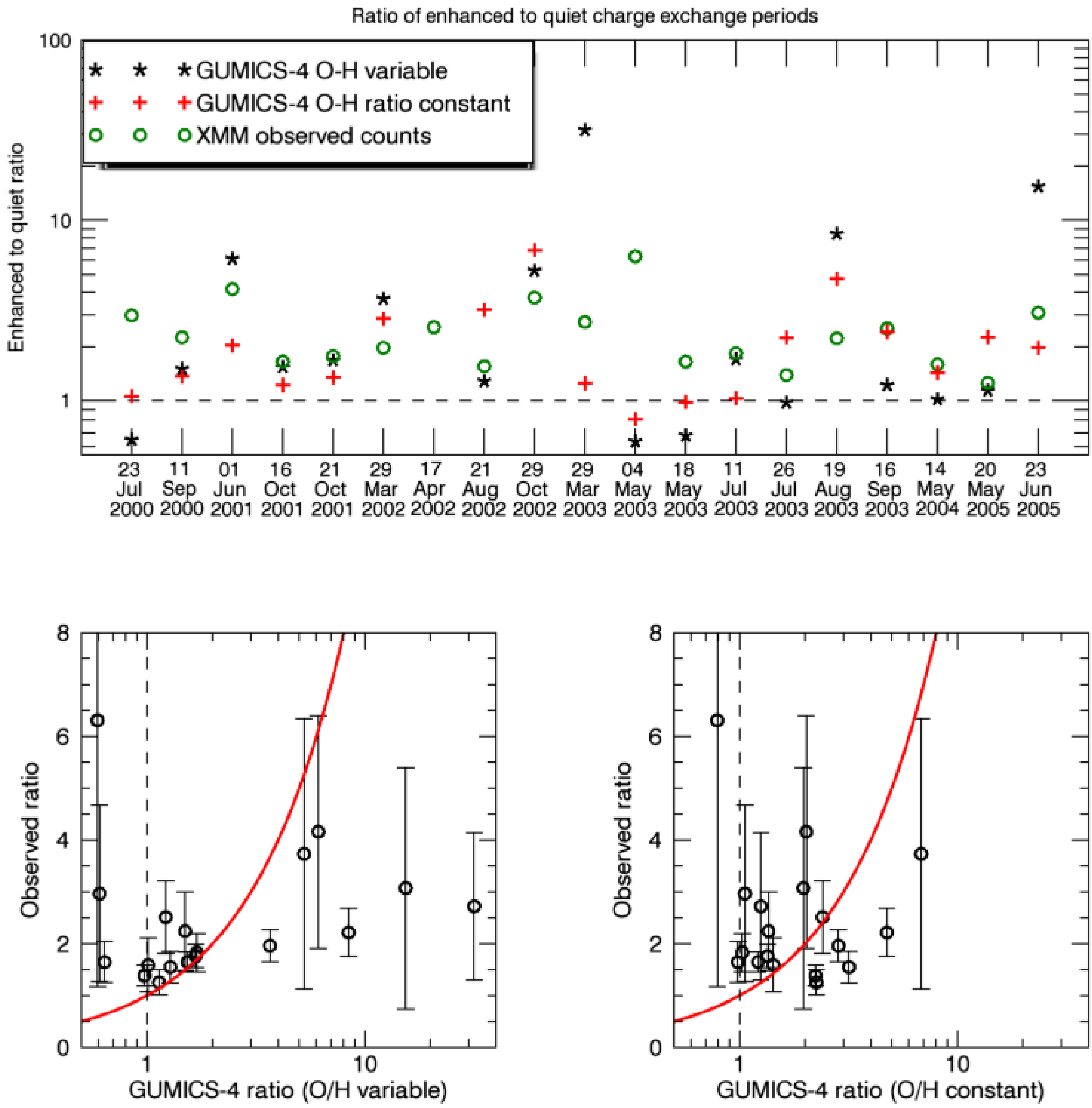

Figure 10. The ratio between the enhanced and quiet charge exchange periods of the lightcurve. The top panel shows the ratio for the O/H varying model data (*), O/H constant model data (+) and the observed data (o). These are in date order to determine any temporal bias. The lower left panel shows a scatter plot of O/H varying model data against the observed data, with the red solid line showing the y=x line. The error bars on the observed data have been propogated from the background and observational data. The lower right panel shows the same plot but with the O/H constant data points.

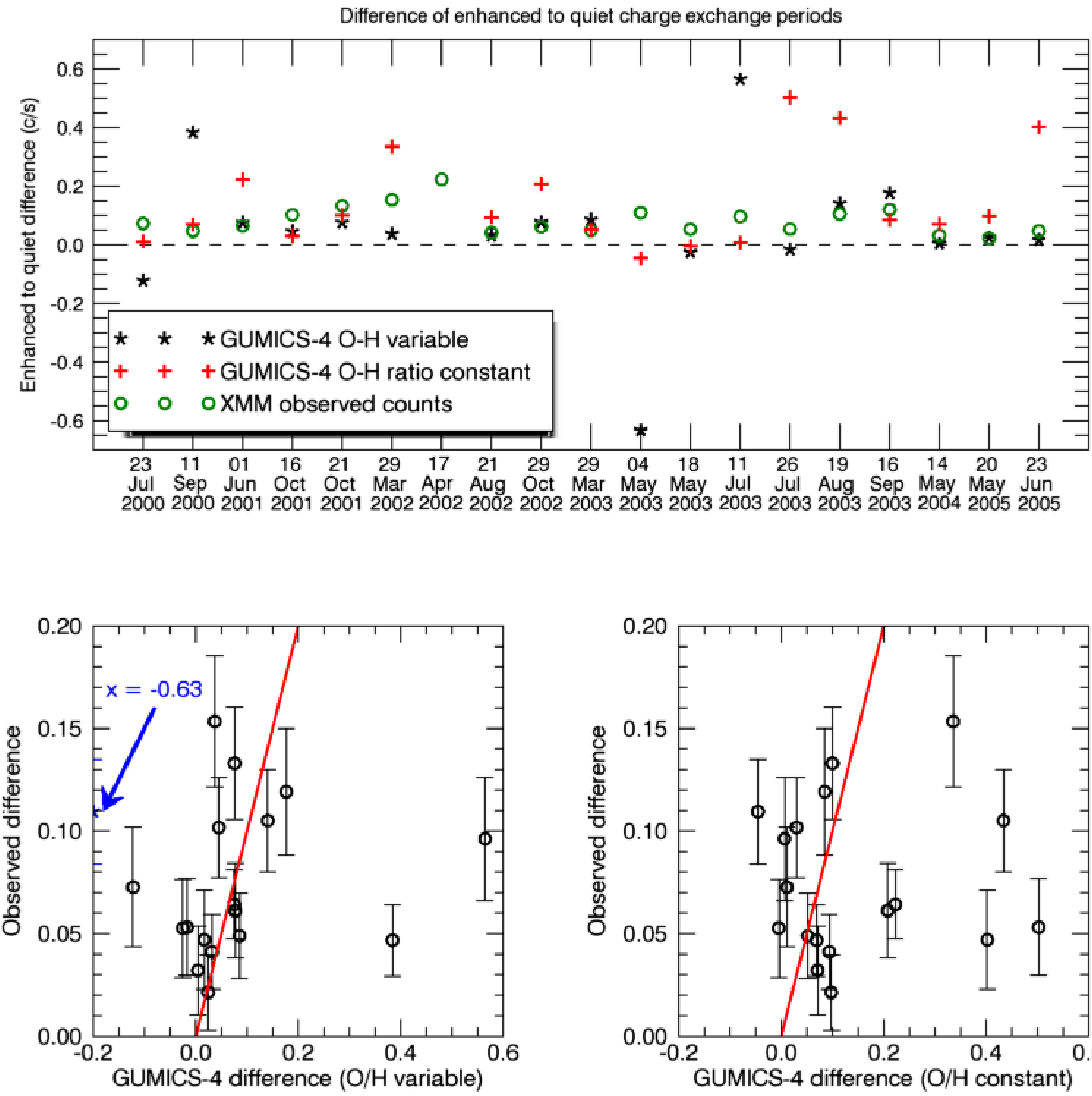

Figure 11. The difference between the enhanced and quiet charge exchange periods of the lightcurve. The top panel shows the ratio for the O/H varying model data (*), O/H constant model data (+) and the observed data (o). These are in date order to determine any temporal bias. The lower left panel shows a scatter plot of O/H varying model data against the observed data, with the red solid line showing the y=x line. The error bars on the observed data have been propagated from the background and observational data. The lower right panel shows the same plot but with the O/H constant data points.

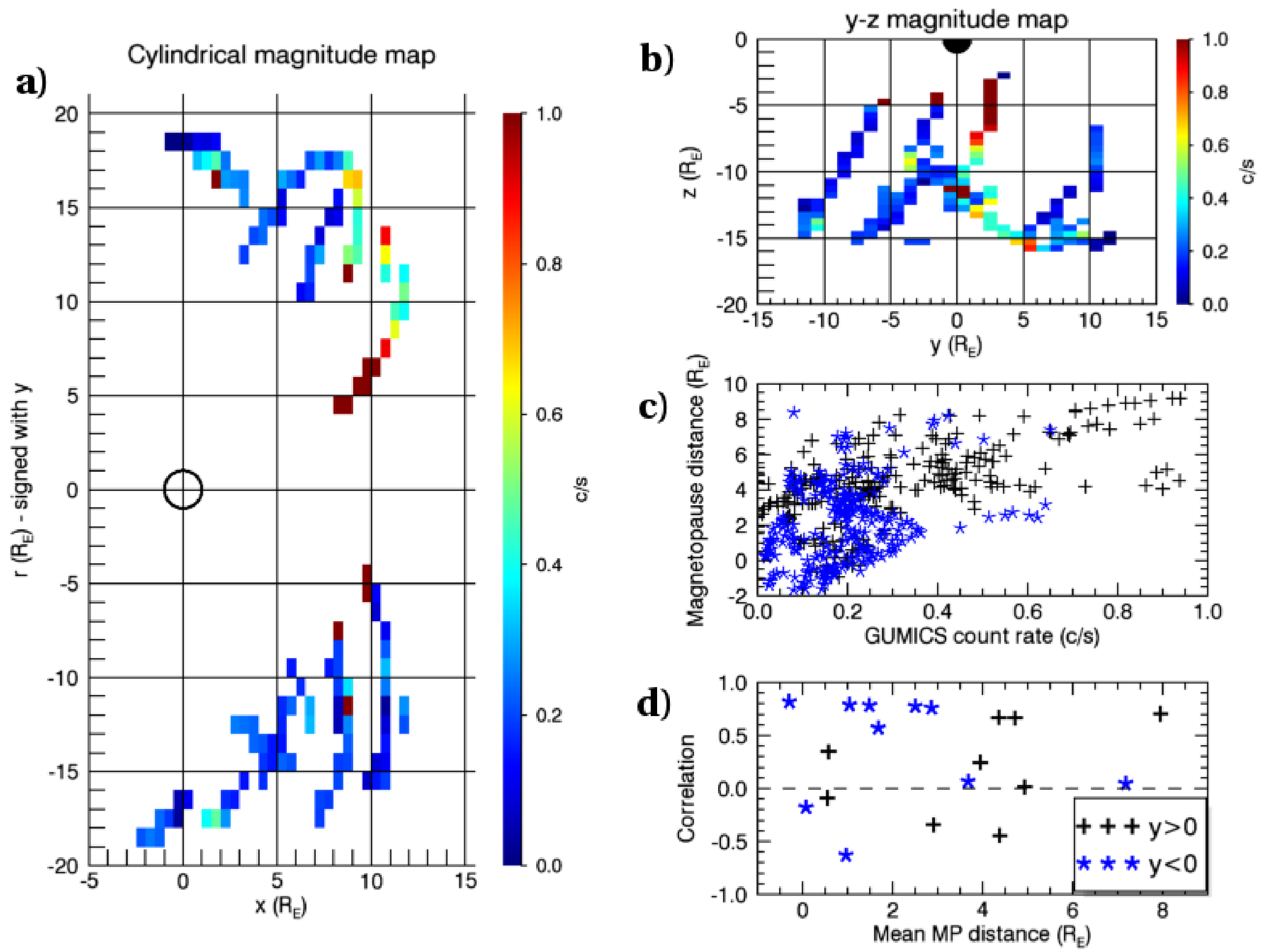

Figure 12. The distribution of GUMICS-4 count rate distribution. **a)** A cylindrical plot in the x-r plane with r signed by y showing the average modeled count rate in each bin, based on the position of XMM-Newton. **b)** The count rates binned in the y-z plane. **c)** A scatter plot showing the GUMICS-4 count rate values against distance from the magnetopause. **d)** The correlation between GUMICS-4 and observations for each case plotted against the average distance from the magnetopause.

Table 1. List of the XMM-Newton observation cases used in this study. The mean case magnitude, pre background removal, is included as well as the calculated background value, both in c/s and as a percentage of the case magnitude. Rejected cases are due to the instrument pointing into the nightside of the Earth where we don't use the MHD model.

Case	Revolution	Obs ID	Date	Duration (hrs)	Magnitude[1] (c/s)	Background (c/s)
	2000					
00-1	0114	0127921101	23 Jul	2.5	0.319	0.129 (40.4%)
00-2	0139	0109060101	11 Sep	14.5	0.146	0.048 (33.2%)
	2001					
01-1	0271	0111550401	01 Jun	26.5	0.179	0.050 (28.1%)
01-2	0339	0054540501	16 Oct	6.5	0.300	0.032 (10.9%)
01-3	0342	0085150301	21 Oct	9	0.306	0.059 (19.2%)
	2002					
02-1	0422	0113050401	29 Mar	7.5	0.444	0.101 (22.8%)
02-2	0431	0136000101	17 Apr	6	0.298	0.077 (25.9%)
02-3	0494	0109120101	21 Aug	11	0.165	0.038 (23.0%)
02-4	0529	0147540101	29 Oct	7	0.232	0.089 (38.5%)
	2003					
03-1	0605	0146390201	29 Mar	6.5	0.215	0.077 (35.6%)
03-2	0623	0150610101	04 May	3	0.288	0.107 (37.2%)
03-3	0630	0143150601	18 May	5.5	0.291	0.080 (27.7%)
03-4	0657	0141980201	11 Jul	6	0.398	0.107 (27.0%)
03-5	0664	0150680101	26 Jul	13	0.268	0.059 (22.0%)
03-6	0676	0049540401	19 Aug	7	0.254	0.082 (32.2%)
03-7	0690	0149630301	16 Sep	5.5	0.331	0.116 (35.0%)
	2004					
04-1	0811	0202100301	14 May	7	0.228	0.079 (34.6%)
	2005					
05-1	0997	0303260501	20 May	9.5	0.164	0.037 (22.4%)
05-2	1014	0305920601	23 Jun	7	0.267	0.105 (39.4%)
	Rejected cases					
	0151	0094800201	05 Oct	2000		
	0163	0100640201	29 Oct	2000	[2]	
	0178	0110980101	27 Nov	2000	[2]	
	0178	0101040301	28 Nov	2000	[2]	
	0209	0093552701	28 Jan	2001		
	0279	0070340501	18 Jun	2001		
	0505	0153752201	11 Sep	2002	[2]	
	0645	0150320201	17 Jun	2003		
	0906	0203361501	19 Nov	2004	[2]	
	0982	0306700301	19 Apr	2005		
	1199	0402250201	27 Jun	2006		

[1] Mean case magnitude before background removal.

[2] The pointing direction is away from the datacube, providing no data. However, the case can still be used for solar wind information.