SPACE WEATHER, VOL. ???, XXXX, DOI:10.1029/,

Reanalysis of Plasma Measurements at Geosynchronous Orbit

T.P. O'Brien, C.L. Lemon

Space Science Department, The Aerospace Corporation, El Segundo,

California, USA

C.L. Lemon, T.P. O'Brien, M2-259, PO Box 92957, Space Science Department, The Aerospace Corporation, El Segundo, CA 90009-2957, USA. (paul.obrien@aero.org)

Abstract. Retrospective analyses of spacecraft anomalies, especially surface charging, require information about the local plasma enviornment. Such information is usually not available at geosynchronous orbit because most vehicles lack appropriate sensors. We have constructed a nearly continuous hourly database covering 4 plasma moments at geosynchronous orbit: density and temperature for hot protons and electrons. The 4 moments are provided for 24 hourly fiducial local time bins and cover the years 1990-2005. We apply an ensemble of simple models to "fill in" the local time gaps between plasma measurements at geosynchronous orbit made by Los Alamos sensors. Validation on artificial data gaps shows that the reconstruction has rank order correlations ranging from 0.65 to 0.97, depending on which bulk moment and local time. As a demonstration, we include a comparison of our reanalysis electron temperature to on-orbit surface charging measurements at a vehicle that lacks a plasma sensor. In additon to anomaly analyses, the reanalysis database we have produced can be used to provide geosynchronous boundary conditions for numerical simulations of the inner magnetosphere, and temporal variations at a single location for statistical studies.

1. Introduction

When evaluating spacecraft anomalies at geosynchronous orbit, one must often consider the possibility of surface charging, which can lead to electrostatic discharges, which in turn can lead to a variety of problems, including phantom commands in spacecraft electronics [*Purvis et al.*, 1984]. Retrospective analysis of spacecraft anomalies requires information about the local plasma environment (e.g., electron temperature for surface charging anomalies), which is not usually available since most geosynchronous vehicles lack plasma sensors. Thus, there is need for some capability to exploit available plasma measurements at geosynchronous orbit to determine conditions elsewhere along the orbit.

For scientific purposes, emprical and simulation studies of the formation of the ring current and, more generally, of the transport of charged particles earthward from the plasma sheet often require or benefit from plasma observations at geosynchronous orbit. For example, numerous ring current simulations employ geosynchronous plasma measurements as a boundary condition [e.g., *Chen et al.*, 2005; *Ebihara et al.*, 2005; *Jordanova and Miyoshi*, 2005; *Liemohn et al.*, 2005, and references therein]. Similarly, statistical studies often prefer or require geosynchronous plasma measurements at a fixed local time rather than in the rotating frame of the Earth [e.g., *Borovsky et al.*, 1998; *Thomsen et al.*, 1998].

Sensors built by Los Alamos National Lab provide in situ plasma measurements on a series of geostationary vehicles first lanched in 1989 [*Bame et al.*, 1993; *McComas et al.*, 1993]. The Magnetospheric Plasma Analyzer (MPA) consists of 6 particle detectors behind spherical section electrostatic analyzer optics and is capable of measuring energy per charge for electrons and ions with energies from $\sim$ 1 eV to $\sim$ 40 keV in 40 logarithmically-

spaced channels. From these, bulk plasma moments are computed [*Thomsen et al.*, 1999]. Here we concern ourselves only with the "hot" plasma, i.e., >100 eV, and with density and temperature moments for electrons and protons (acually total ions). There is no explicit ion species resolution, but the vast majority of the ions are thought to be protons. Corrections to the bulk moments have been attempted [e.g., *Denton et al.*, 2005] to account for Oxygen ions, but these have yet to be validated by comparison to a more capable instrument during conjunction. We denote density and temperature by n and T respectively, and the "hot" proton and electron moments are denoted by subscripts hp and he, respectively; thus n_{hp} is the hot proton density.

With measurements from up to 6 vehicles available at any time, the MPA data constitute one of the most extensive plasma measurement databases in the entire magnetosphere. Nonetheless, for many applications, the presence of 6 or fewer measurements is insufficient, and some method must be found for interpolating between the sensors.

At present, statistical maps exist that can provide geosynchronous plasma conditions as a function of Kp [*Korth et al.*, 1999], Dst [*Denton et al.*, 2005], super dense plasma sheet event epoch time [*Lavraud et al.*, 2005], storm epoch time [*Denton et al.*, 2005], or solar wind structure type [*Denton et al.*, 2006]. Other efforts, in progress now, will provide more detailed specifications of plasma conditions at geosynchronous orbit as a function of recent solar wind measurements (*Denton et al.* and *Lemon and O'Brien* manuscripts in preparation). We describe here a database of historical geosynchronous plasma parameters interpolated from the MPA observations to produce a temporally and spatially continuous "reanalysis" of plasma moments in 24 fiducial local time bins at hourly cadence for years 1990-2005.

The term "reanalysis" is taken from the field of climate research. The "analysis" refers to the output of some kind of climate or weather model. The "re-" prefix stems from the fact that it is often the case that an initial analysis is performed near real-time, and a later analysis, the reanalysis, is performed with better models and data as they become available. For example, the NCAR/NCEP Climate Reanalysis Project [*Kalnay et al.*, 1996] uses a standardized numerical climate model and archived weather observations to construct a 40 year history of global terrestrial weather. A preliminary effort has begun for space weather, with emphasis on the ionosphere and upper atmosphere [*Kihn et al.*, 2002]. Ours may be the first attempt at reanalysis for any high altitude space weather.

2. Method

In order to "fill in" the local time gaps in the MPA measurements, we considered various methods. One approach is to use the Kp or Dst models described in the previous section, or a more sophisticated multi-variate method that combines Kp, Dst, and possibly other magnetic indices [e.g., *Fung*, 1996]. However, these methods do not directly take advantange of the concurrent MPA measurements availale at up to 6 local times.

2.1. Model Ensemble

We have chosen an ensemble approach. For a given target bulk moment and local time, we construct an ensemble of simple empirical regression models. The regression models are linear in log-log space, with the dependent variable being the target bulk moment at a particular fiducial local time (e.g., n_{hp} at local midnight). We try a variety of independent variables for each model: Kp, $F10.7$, and the very moment we are predicting at a different hourly local time with a time offset up to 4 hours forward or backward. We thus have

$24 \times 9 + 1 = 217$ models for each target moment and local time. We would have 218 models but we exclude the degenerate model built from the moment itself at the same local time at zero time offset. Considering 4 target moments and 24 target local times, we have a total of $(24 \times 9 + 1) \times 24 \times 4 = 20,832$ models altogether. Needless to say, we will not be providing details of each model here.

To construct an individual model, we select those times when the target y and the independent variable x are both available. We then perform a least-squares linear regression of the logs of x and y, leaving out pairs with $x = 0$ or $y = 0$. If there are not at least 365×2 valid pairs (i.e., the equivalent of 2 years of daily samples), we reject the model as underspecified. Otherwise we apply the individual model to construct estimated $\hat{y}$ for all the cases when x is available but y is not.

The Kp- and $F10.7$-based models are available essentially all of the time. However, the models driven by the plasma moments are only sparsely available. To determine which model to use, we compute an in-sample linear correlation in log-log space on the data used to build the model. For each time bin, for each target moment, for each target local time, we determine the 3 available models with the best in-sample prediction efficiency (r^2). From these we compute a mean $\bar{y}$ and standard error $\bar{\sigma}$, where the mean is weighted by the prediction efficiency.

2.2. Regression toward the Mean

Most models, ours included, tend to distort the statistical distribution of data by clipping the extreme values in favor of values closer to the median. This resembles the so-called "regression toward the mean" observed among human and plant populations by *Galton* [1886], in which the offspring of individuals extreme in some attribute tend to be

closer to the population mean in that same attribute. In our case, the "regression" occurs because linear regression tends to treat extreme values as noise in favor of more likely values near the population mean. As an example, Figure 1 compares the distribution of the observations to those of the preliminary (raw) reanalysis for hot electron density at 0530 Local Time. The plot depicts the quantiles of the ordinate versus the quantiles of the abscissa. The clockwise rotation away from the one-to-one line indicates suppression of extreme values. In order for the reanalysis to accurately reflect the observed statistical distribution of the data, the QQ plot must be corrected back to the red line.

To correct for the regression phenomenon, we consider the $\bar{y}$ values as preliminary, with each local time k having cumulative distribution $\bar{F}_k(\bar{y}^{(k)})$. We determine such distributions empirically by assigning $F(y_i)$ for the i^{th} smallest y as:

$$F(y_i) = \frac{i}{N_y + 1}, \tag{1}$$

Where N_y is the number of valid samples of y. We adjust $\bar{y}$ to obtain a final estimate $\hat{y}$ which has the same distribution as the measurements $F_k(y^{(k)})$:

$$\hat{y}_i^{(k)} = \left(F_k^{(k)}\right)^{-1} \left(\bar{F}_k(\bar{y}_i^{(k)})\right). \tag{2}$$

Evaluation of the distribution functions F and $\bar{F}$ away from actual samples is achieved by linear interpolation. Extreme values of $\hat{y}^{(k)}$ are truncated to be within the observed range of $y^{(k)}$. The correction from $\bar{y}$ to $\hat{y}$ introduces an adjustment error $|\bar{y} - \hat{y}|$. We combine the adjustment error with the standard error by taking the larger of the two, and report a final error estimate $\hat{\sigma}$:

$$\hat{\sigma}_i^{(k)} = \max\left(|\bar{y}_i^{(k)} - \hat{y}_i^{(k)}|, \bar{\sigma}_i^{(k)}\right). \tag{3}$$

3. Validation

To validate the reanalysis we sequester data for out-of-sample testing. Before constructing the ensembles of models, we artificially create data gaps in 5% of our measurements, selected at random. We carry through our entire reanalysis with these artificial gaps, and then compare the values produced by our reanalysis with the sequestered 5%. Figure 2 shows the out-of-sample rank-order correlation coefficient between the sequestered data and the reanalysis. Rank-order correlation is computed by converting each value in a sample to its rank order (or percentile) and computing the correlation between the rank orders. This type of correlation measure is not affected by monotonic transformations, like logarithm or equation (2).

The correlations are generally in the range 0.80-0.95, with hot proton density usually having the best correlation and hot electron temperature usually having the worst. Correlations worse than 0.80 do occur, especially for electron moments near dusk. An examination of these data indicates that the Kp and $F10.7$ models, which are used extensively for other local times and other plasma moments, have unusually poor performance for electrons near dusk. Panels (a) and (b) of Figure 3 show scatter plots for representative best and worst cases. We see in panel (a) that the high correlation model has good performance, over most of the range of the data, with greater scatter at large values. In panel (b) we see that the data is clustered in two regimes. The histograms, panels (c) and (d), show that both data sets make extensive use of Kp and $F10.7$ models as well as recent and local measurements. The poor performance in panel (b) can only be attributed to poor correlation between the available data and the hot electron temperature near dusk.

4. Results

The final product of our efforts is a nearly continuous specification of the 4 plasma moments for 24 fiducial local times every hour from 1990 to 2005. Figures 4-7 show sample intervals for each moment (Panels a-b). This interval included a high speed stream (500 km/s on 28-29 Nov) and two magnetic storms (Dst -50 nT on 27 Nov, and -60 nT on 1 Dec). Panels c-d depict the spatial and temporal correlations computed from the entire observation data set and from the reanalysis.

In all of the Panels labeled (a) in Figures 4-7 there is some evidence of diagonal striation associated with actual measurements influencing the reanalysis along the trajectory of the vehicles, which are diagonal in local time versus time coordinates. There is also some residual diurnal variation, most of which is attributable to actual variations in the observations, as follows. In all of the sample intervals, there is a striking amount of temporal variability in panel (b), even though the moments are being plotted at a fixed location. To put this variability in context, actual measurements at the fixed location are marked in red, and the minimum and maximum observations at any location for each time bin are indicated in green. The measurements shown in red were those included in the construction of the reanalysis, so the reanalysis is forced to match them. However, the variability in the reanalysis (blue) whenever there is no local measurement is within the range of the concurrent observations at other locations (green). In a statistical context, we are concerned with the temporal and spatial correlation functions. The lower panels (c)-(d) in each figure show that the reanalysis captures the temporal correlations with somewhat more fidelity than the spatial correlations.

We have thus produced a reanalysis that reproduces with high fidelity the frequency distribution and temporal and spatial correlation structure. The quantitative accuracy of the reanalysis ranges from 0.65 to 0.97, depending on the bulk moment and local time of interest, with proton moments being generally more accurate than electron moments, which experience their worst performance near dusk.

5. Discussion

There are many things our reanalysis cannot do. Foremost among them is alleviating the need for multiple in-situ measurements of the plasma environment. The short spatial decorrelation lengths of the plasma moments (and likely even shorter lengths for the underlying fluxes), suggest that continued and vigilant monitoring is the only option for the foreseeable future. Further, by construction, our reanalysis cannot be used to extract new details (e.g., extreme values) from the observed frequency distribution of the plasma moments.

What can our reanalysis do that the observations cannot do? The relative value of the reanalysis is greatest when (1) a complete specification at all local times is required, or (2) a continuous specification at a fixed location is required, or (3) a specification is needed at some longitude away from the LANL measurements. In the first case, we find the boundary conditions for numerical simulations. In the second case we find statistical analyses where we wish to separate spatial and temporal variability from the observation record. Investigations from both categories abound in the literature. In the third case, we find operational space weather applications.

What is the operational space weather application of our reanalysis? One obvious application mentioned above is spacecraft charging: while individual spacecraft charge

under different conditions, surface charging is thought to be well-correlated with the ambient electron temperature [e.g., using the LANL MPA data, *Thomsen et al.*, 1999; *Davis et al.*, 2003]. Our reanalysis can be used to infer the electron temperature at a geosynchronous vehicle that lacks its own plasma or charging measurements. This can be valuable for retrospective studies of suspected surface charging anomalies; although, caution is warranted because actual surface charging has been shown to have spatial decorrelation lengths even shorter than those reported in Figure 7 [*Koons et al.*, 2006]. For example, INTELSAT 802, located at about 174°E is about 3 hours of local time (45 degrees) away from the nearest LANL vehicle. It carries a set of charge plate analyzers (CPA) [*Bogorad et al.*, 1995], of which the zenith sensor is known to best reflect the presence of surface charging conditions on the night side (1800-0600 LT) [*Ozkul et al.*, 2001]. We have interpolated our hourly temperature reanalysis to the 1-minute time series of CPA voltages, using the nearest hourly local time and universal time value. Figure 8 shows the cumulative distribution of CPA voltage when the local reanalysis hot electron temperature is above and below 1000 eV (night side only). We see that when $T_{he} > 1000$ eV, the surface potential exceeds -100 V more than half the time, whereas when $T_{he} < 1000$ eV, the surface potential exceeds -100 V less than 7% of the time. At -500 V, percentages are $\sim 12\%$ and $\sim 0.4\%$, respectively. Thus, the reanalysis electron temperature provides a reasonable indication of an elevated surface charging hazard.

Finally, we would like to explain why we have emphasized reanalysis, which is inherently retrospective, as opposed to prediction. Most of the methods we have discussed here are not explicitly constrained to retrospective analyses, and could be developed for near-real-time uses, such as situational awareness. However, before an environment forecast can be

actionable, it is usually necessary to establish a strong connection between the environment and anomalous behavior on the vehicle. Such connections nearly always arise from retrospective studies of one or more anomalies experienced on the vehicle in question. Thus, retrospective studies, and hence reanalysis, are an essential step in turning environmental forecasts into actionable information. Also, for gaining physical understanding of space weather processes, scientists primarily use retrospective analysis. In either case, it is necessary to extract as much information as possible about the state of the magnetosphere from the limited spacecraft data available, and we would like to emphasize that there is more to the study of space weather than forecasting and nowcasting.

Acknowledgments. We would like to thank M. Thomsen and M. Denton as well as the entire Los Alamos space science group for insightful discussions of the plasma measurements and reanalysis. We thank INTELSAT, New Skies Satellites, and Lockheed Martin Commercial Space Systems for providing the surface charging data. We also thank our colleagues at The Aerospace Corporation for helpful comments. This research was funded in part by NSF GEM grant ATM-0302467, NASA LWS TR&T grant NNG05GM22G, and The Aerospace Corporation's Independent Research and Development Program. ATR-2006(5072)-2.

References

Bame, S.J., D.J. McComas, M.F. Thomsen, B.L. Barraclough, R.C. Elphic, and J.P. Glore, Magnetospheric plasma analyzer for spacecraft with constrained resources, *Rev. Sci. Instrum., 64*, 1026-1033, 1993.

Bogorad, A., et al., Integrated environmental monitoring system for spacecraft, *IEEE Trans. Nucl. Sci., 42*(6), 2051-2057, 1995.

Borovsky, J. E., M. F. Thomsen, and R. C. Elphic, The driving of the plasma sheet by the solar wind, *J. Geophys. Res., 103*, 17,617, 1998.

Chen, M. W., M. Schulz, P. C. Anderson, G. Lu, G. Germany, and M. Wüest (2005), Storm time distributions of diffuse auroral electron energy and X-ray flux: Comparison of drift-loss simulations with observations, *J. Geophys. Res., 110*, A03210, doi:10.1029/2004JA010725.

Davis, V. A.; Mandell, M. J.; Thomsen, M. F., Characterization of Magnetospheric Spacecraft Charging Environments Using the LANL Magnetospheric Plasma Analyzer Data Set, NASA/CR-2003-212745, 2003.

Denton, M. H., M. F. Thomsen, H. Korth, S. Lynch, J. C. Zhang, and M. W. Liemohn (2005), Bulk plasma properties at geosynchronous orbit, *J. Geophys. Res., 110*, A07223, doi:10.1029/2004JA010861.

Denton, M. H., J. E. Borovsky, R. M. Skoug, M. F. Thomsen, B. Lavraud, M. G. Henderson, R. L. McPherron, J. C. Zhang, and M. W. Liemohn (2006), Geomagnetic storms driven by ICME- and CIR-dominated solar wind, *J. Geophys. Res., 111*, A07S07, doi:10.1029/2005JA011436.

Ebihara, Y., M.-C. Fok, S. Sazykin, M. F. Thomsen, M. R. Hairston, D. S. Evans, F. J. Rich, and M. Ejiri (2005), Ring current and the magnetosphere-ionosphere coupling during the superstorm of 20 November 2003, *J. Geophys. Res., 110*, A09S22, doi:10.1029/2004JA010924.

Fung, S.F., Recent development in the NASA trapped radiation models, in *Radiation Belts, Models and Standards, Geophys. Monogr. Ser. 97*, AGU, Washington, DC, 79-90, 1996.

Korth, H., M. F. Thomsen, J. E. Borovsky, and D. J. McComas (1999), Plasma sheet access to geosynchronous orbit, *J. Geophys. Res., 104*, 25,047-25,061.

Galton, F., Regression Towards Mediocrity in Hereditary Stature, *Journal of the Anthropological Institute of Great Britain and Ireland, 15* 246-263, 1886.

Jordanova, V. K., and Y. S. Miyoshi (2005), Relativistic model of ring current and radiation belt ions and electrons: Initial results, *Geophys. Res. Lett., 32*, L14104, doi:10.1029/2005GL023020.

Kalnay, E. et al. The NCEP/NCAR 40-Year Reanalysis Project, *Bull. Amer. Meteor. Soc., 77*(3), 437-471, 1996.

Kihn, E.A., A.J. Ridley, M. Zhizhin, The Space Weather Reanalysis, (abstract), *Eos Trans. AGU, 83*(47), Fall Meet. Suppl., SH51A-0429, 2002.

Koons, H., J. Mazur, A. Lopatin, D. Pitchford, A. Bogorad, and R. Herschitz, Spatial and Temporal Correlation of Spacecraft Surface Charging in Geosynchronous Orbit, *J. Spacecr. Rockets, 43*(1), 178-185, 2006.

Lavraud, B., M. H. Denton, M. F. Thomsen, J. E. Borovsky, and R. H. Friedel (2005), Superposed epoch analysis of cold, dense plasma sheet access to geosynchronous orbit, *Ann. Geophys., 23*, 1-11.

Liemohn, M. W., A. J. Ridley, P. C. Brandt, D. L. Gallagher, J. U. Kozyra, D. M. Ober, D. G. Mitchell, E. C. Roelof, and R. DeMajistre (2005), Parametric analysis of nightside conductance effects on inner magnetospheric dynamics for the 17 April 2002 storm, *J.*

Geophys. Res., 110, A12S22, doi:10.1029/2005JA011109.

McComas, D.J., S.J. Bame, B.L. Barraclough, J.R. Donart, R.C. Elphic, J.T. Gosling, M.B. Moldwin, K.R. Moore, and M.F. Thomsen, Magnetospheric plasma analyzer: Initial 3-spacecraft observations from geosynchronous orbit, *J. Geophys. Res., 98*, 13,453-13,465, 1993.

Ozkul, A., et al., Initial correlation results of charge sensor data from six INTELSAT VIII class satellites with other space and ground-based measurements, *Proceedings of the 7th Spacecraft Charging Technology Conference,* SP-476, ESA, Noordwijk, The Netherlands, 293-298, 2001.

Purvis, C.K., H.B. Garrett, A.C. Wittlesey, and N.J. Stevens, Design Guidelines for Assessing and Controlling Spacecraft Charging Effects NASA Technical Paper 2361, 1984. http://powerweb.grc.nasa.gov/pvsee/publications/geoguide.pdf

Thomsen, M. F., J. E. Borovsky, D. J. McComas, and M. R. Collier, Variability of the ring current source population, *Geophys. Res. Lett., 25*, 3481, 1998.

Thomsen, M.F., E. Noveroske, J.E. Borovsky, D.J. McComas, Calculation of moments from measurements by the Los Alamos magnetospheric plasma analyzer, Los Alamos National Lab Technical Report LA-13566-MS, 1999.

Ukhorskiy, A. Y., M. I. Sitnov, A. S. Sharma, B. J. Anderson, S. Ohtani, and A. T. Y. Lui (2004), Data-derived forecasting model for relativistic electron intensity at geosynchronous orbit, *Geophys. Res. Lett., 31*, L09806, doi:10.1029/2004GL019616.

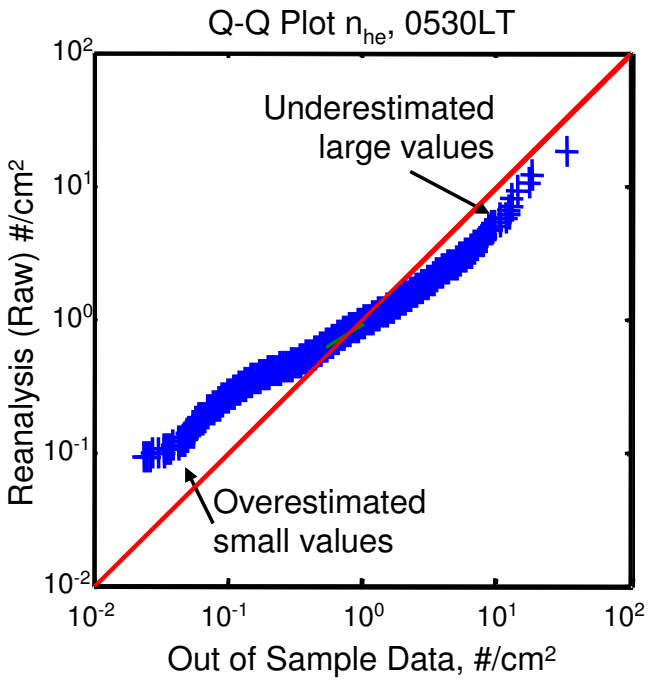

Figure 1. A quantile-quantile (QQ) plot comparing the distribution of unadjusted reanalysis to observations for hot electron density at 0530 Local Time. The QQ plot depicts the quantiles of the ordinate versus the quantiles of the abscissa. The clockwise rotation away from the red line indicates suppression of extreme values.

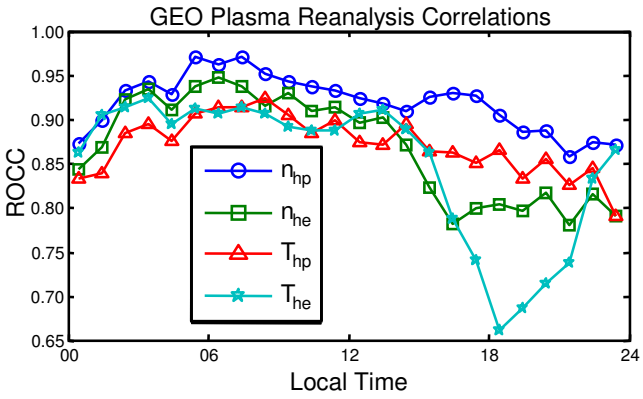

Figure 2. Rank order correlation coefficient for out-of-sample reanalysis as a function of local time for all 4 bulk moments.

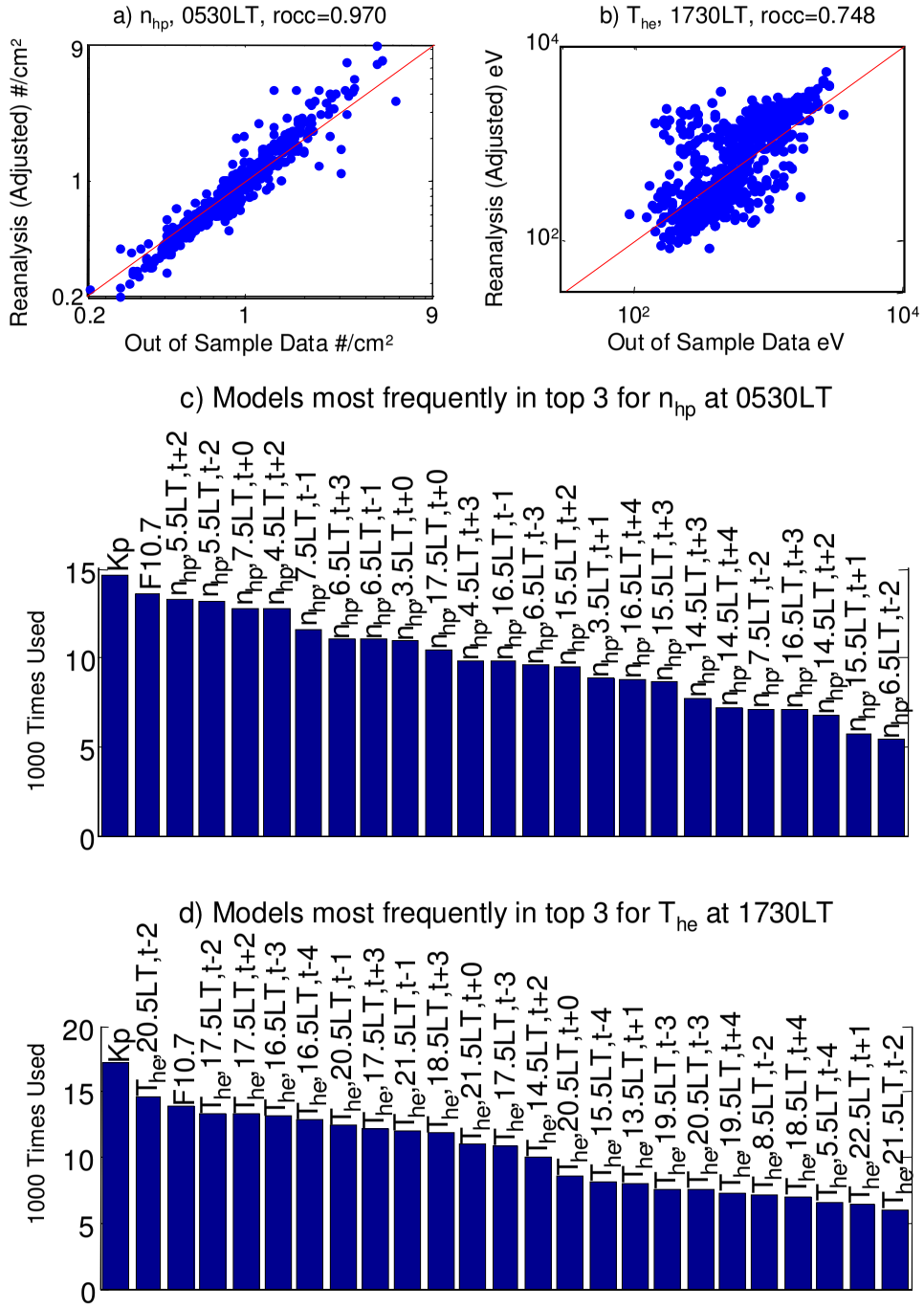

Figure 3. Panels a and b: scatter plots for sequestered data against adjusted reanalysis. Panels c and d: frequency with which most used individual models occur in the top 3; $t-\tau$ indicates that the quantity used in the model was measured τ hours before the moment being modeled.

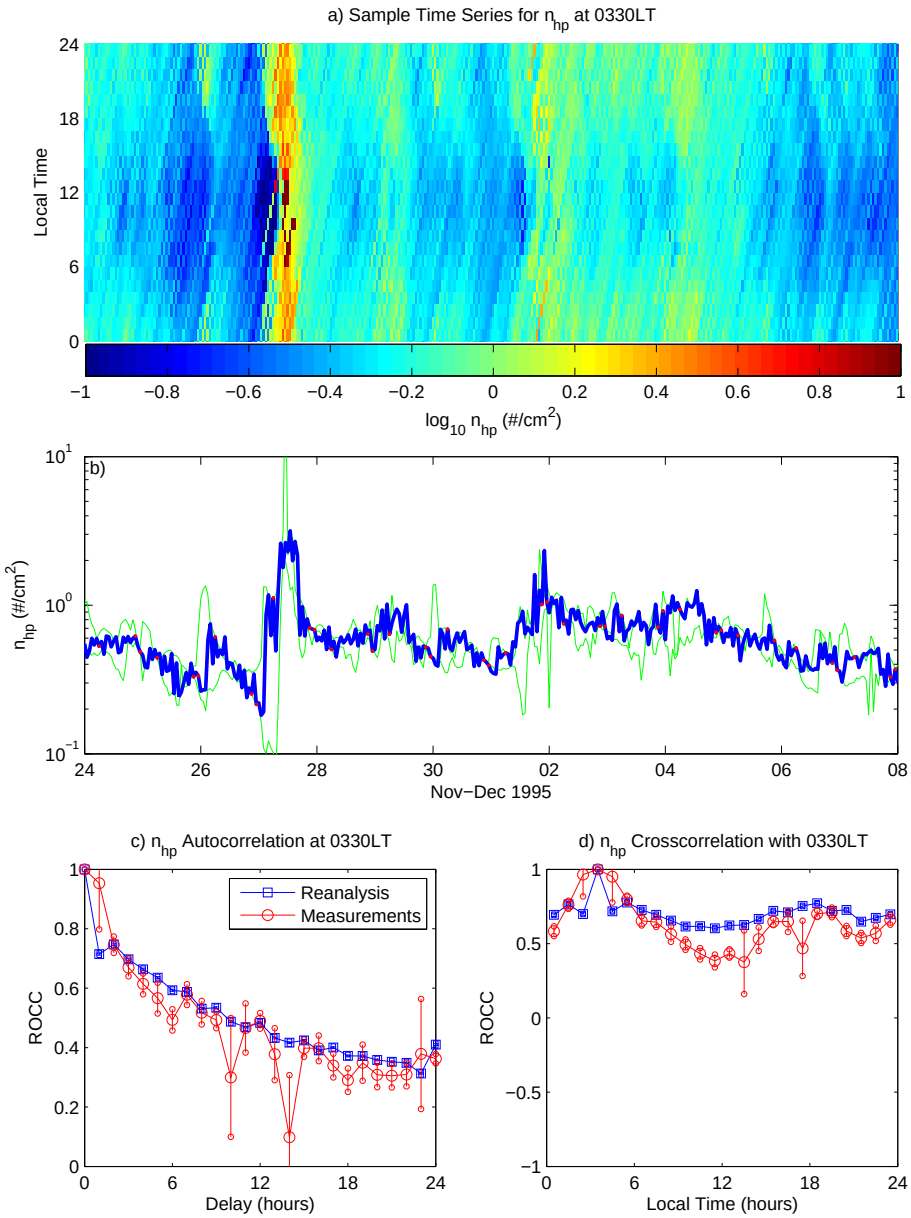

Figure 4. A section of time showing the n_{hp} reanalysis. Panel (a): reanalysis at all local times. Panel (b) in-sample measurements (red) and the reanalysis (blue) at 0330LT on the same time axis as (a). The green trace indicates the minimum/maximum observation at any local time for each time bin. Panels (c) and (d) correlation versus n_{hp} at time lags and at other local times.

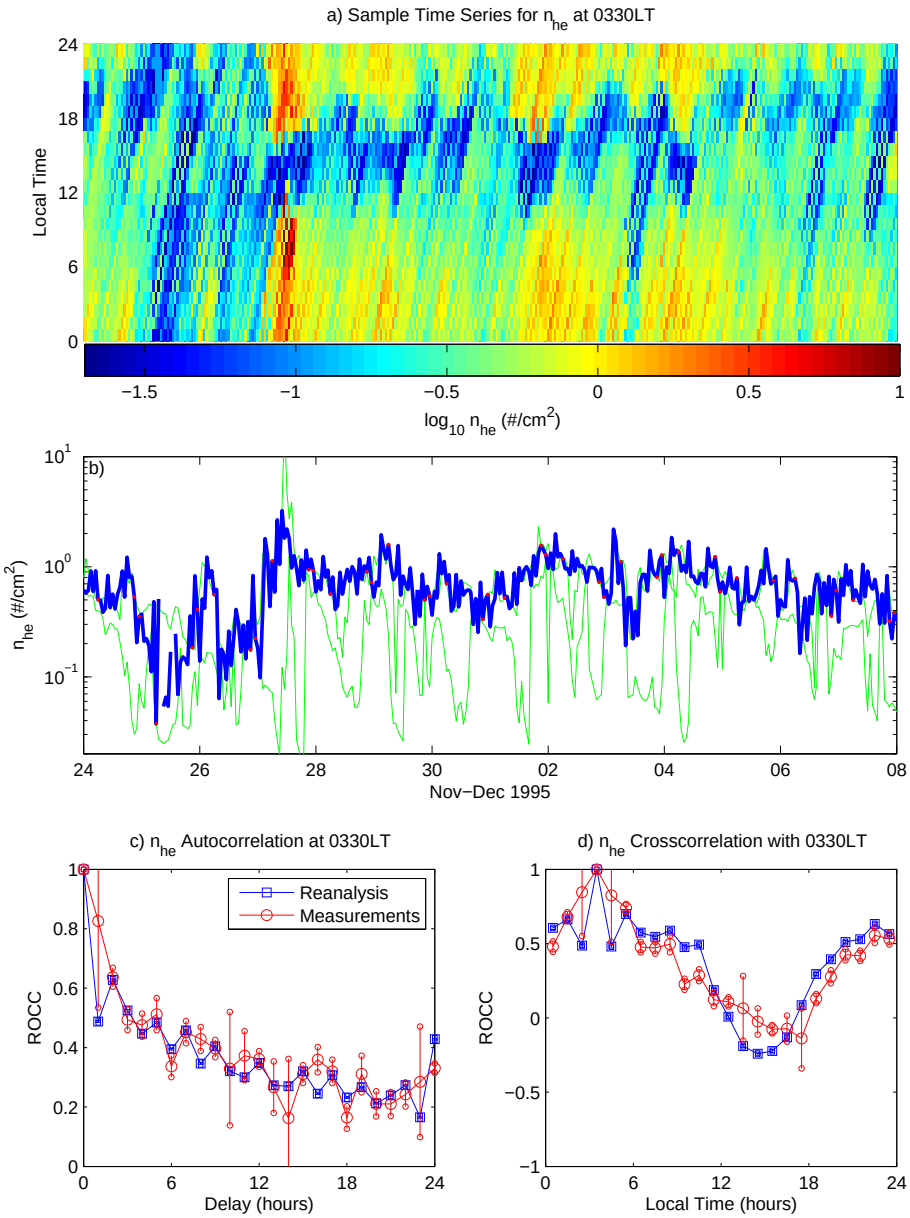

Figure 5. Reanalysis of n_{he} in the format of Figure 4.

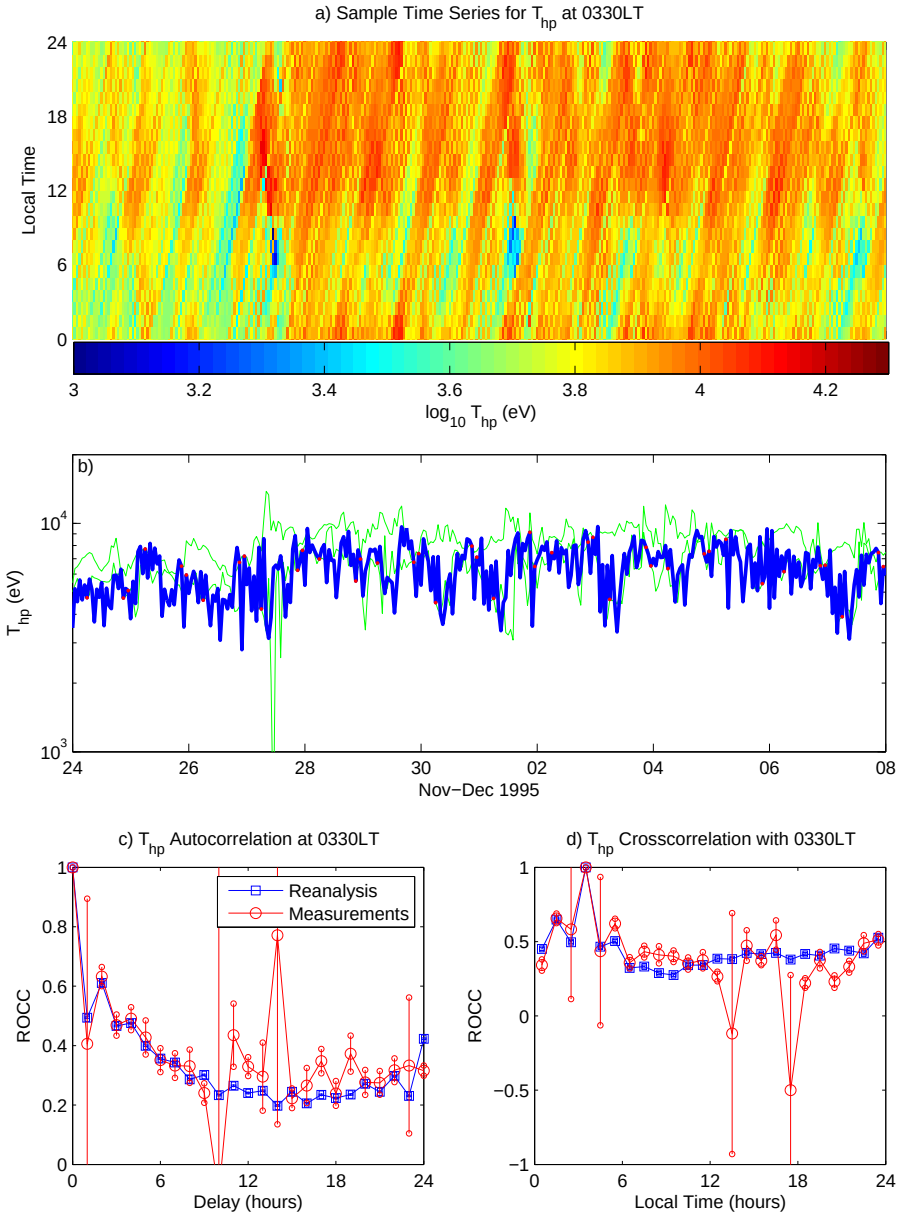

Figure 6. Reanalysis of T_{hp} in the format of Figure 4.

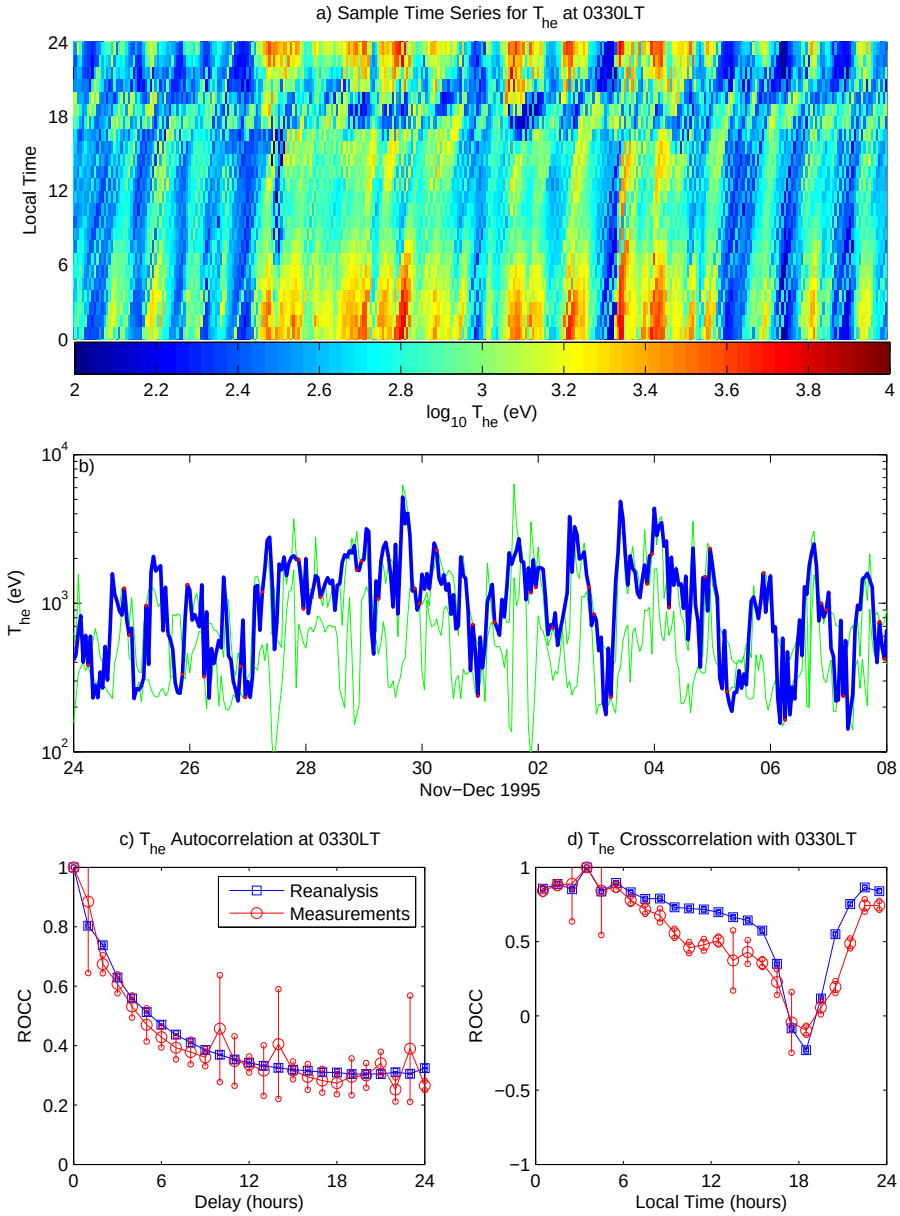

Figure 7. Reanalysis of T_{he} in the format of Figure 4.

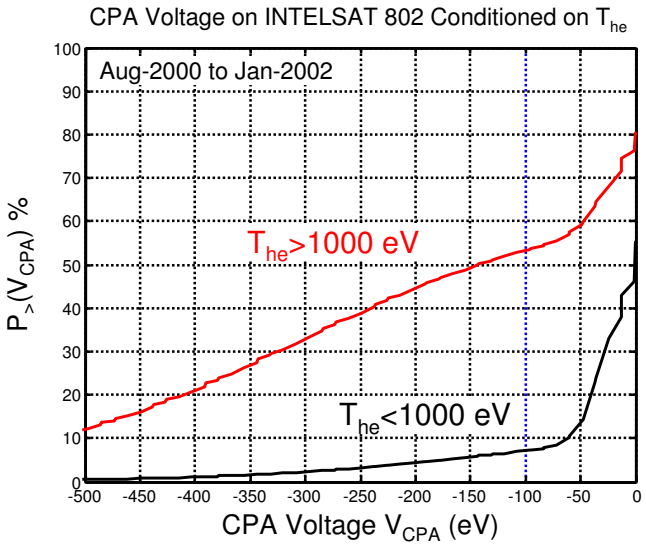

Figure 8. Cumulative distribution functions for surface potential conditioned on local electron temperature from the reanalysis. The nearest LANL measurement is 3 hours (45 degrees) away. (Night side data only.)