F region climatology during the SUNDIAL/ATLAS 1 campaign of March 1992: Model-measurement comparisons and cause-effect relationships


Abstract. We present the first joint comparison of global measurements of F region characteristics with three models used widely in the specification of the ionospheric-thermospheric system. The models, the International Reference Ionosphere (IRI), the field line interhemispheric plasma (FLIP) model, and the Thermospheric-Ionospheric General Circulation Model (TIGCM), represent a unique set of capabilities with major differences in approaches to the prevailing physics and different levels of computational complexity. The database was developed by a global network of 53 ionosonde stations operating around-the-clock for the period March 22 through April 4, 1992 in collaboration with the ATLAS 1 mission. The emphasis is on the F region characteristics of peak heights (NmF2) and densities (NmF2), their climatological (i.e., average) behavior during the ATLAS 1 period, and associated cause-effect relationships. We explore latitudinal and local time variations with attention to the influences of meridional winds and plasmaspheric fluxes in the maintenance of different domains in the ionospheric-thermospheric system. We find that all three models tend to underestimate the values of NmF2 and NmF2 with the largest discrepancies in NmF2, resulting in the FLIP and TIGCM representations at night. These discrepancies can grow to levels as large as 110% near 0400 LT, a "rediscovery" of the old but unsettled issue of maintenance of the nighttime ionosphere. This nighttime discrepancy is traceable in first order to model underestimates of prevailing meridional winds. The contributions of plasmaspheric fluxes are also considered, with the conclusion that they are of secondary importance, but substantial uncertainty is necessary to uniquely quantify their role. In contrast, to their nighttime characteristics, the FLIP and TIGCM generally have excellent agreement (i.e., ±5%) with daytime observations of NmF2, and the IRI tends to underestimate the observed values of NmF2 by a nominally LT-insensitive level of 28±6%. Other campaign results are reviewed in this issue, with a focus on regional responses to the prevailing conditions and their characterization in terms of latitudinal and longitudinal distributions of F region heights and densities.

1. Introduction

1.1. Perspectives on the Investigation

There are a number of trends evolving in experimental and theoretical aspects of solar-terrestrial physics. These may be considered to include the following: (1) the organization of characteristic responses within the framework of space weather and climatology; (2) an increased emphasis on remote sensing techniques, and (3) dependence on large-scale numerical models for predictions or deductions about cause-effect relationships that are not readily measured.

Within the concept of "space weather and climatology" [Smith and Behnke, 1994; St. evee et al., 1994; SzuSzczeWicz, 1995], climatology deals with "average" conditions that result under a set of large-scale slowly changing (i.e., weeks, to months, to years) forcing functions. On the other hand, weather addresses drivers and system responses framed within their hour-to-hour and day-to-day variability. Embodied in these concepts is the idea of "predictions" that encompass the scientifically rigorous position that accurate predictions of weather and climatology are the ultimate test of understanding. As space science matures, the logic of such a focus looms as increasingly important.

The trend toward an ever-increasing reliance on large-scale numerical models stems in part from the premise that no measurement program will ever be able to define completely the input and output parameters in a given cause-effect chain of geophysical events. Thus theory, and in particular comprehensive first-principle models, must be relied upon to make deductions where data are not available and to provide
insight into the physical processes of importance. However, theoretical models are accurate only to the extent that the fundamental processes are adequately understood, mathematically represented, and empirically bounded. Therefore there is the need to use experimental data, in particular global-scale data, to validate individual models, quantify their accuracies, impose self-consistency, and identify regions for improvement.

The trend toward an increased emphasis on remote sensing techniques is motivated in part by the number of geophysical or solar-terrestrial domains that are difficult or impossible to routinely access with "in situ" techniques. This trend also appears to be motivated by a perspective that remote sensing techniques can provide more global views of phenomenological domains and therefore perform (arguably) more effectively than satellite- or rocket-borne "in situ" diagnostics. Such has been the thrust of the ATLAS 1 mission [Torr et al., 1993; Fennelly et al., 1994], the current series of shuttle-borne GLO investigations [Broadfoot et al., 1992; Gardner et al., 1995], and the planned NASA/TIMED satellite program. Many of the remote sensing techniques however are integrated line-of-sight measurements that require inversion procedures that are forced to make assumptions about the physical processes and local particle distributions giving rise to their respective observables. Therefore there is often need for coincident ground-truth measurements that provide benchmarks for the remote sensing procedures as well as spatial and temporal continuity between and among the along-track remote-sensing database.

This investigation embodies all three trends. It addresses the issue of cause-effect terms as manifested in the global-scale climatology of the ionospheric-thermospheric system during the solar-maximum equinoctial period of the ATLAS 1 period. It also addresses the issue of model testing and verification, with an emphasis on quantifying agreement and disagreement of observations as well as between and among various models. The approach also focuses on remote sensing techniques and necessary elements of ground-truth with a global network of 53 SUNDIAL ionosondes operating around-the-clock for the entire period of the ATLAS 1 mission (March 23 through April 7, 1992).

1.2. Investigative Components

ATLAS 1 was launched on March 23, 1992, at 1301 UT into a 57° inclination 297-km circular orbit with a complement of photometric devices covering atomic, molecular and ionic emissions over the wavelength regime from the EUV through the IR [Torr et al., 1993]. The intent was to measure the density and temperature distributions of the neutral and ion constituents from mesospheric altitudes up through the F-region. This was to be the first of a series of flights which would develop one of the first data sets dealing with the mesosphere and make available ionospheric and thermospheric results for the test and verification of developing large-scale models.

The ground-based SUNDIAL component had a separate but synergistic set of continuing objectives that supported the ATLAS 1 interests while developing a synoptic view of the global-scale ionospheric-thermospheric (I-T) system. Among those objectives has been the development of a worldwide database [Szaszczewicz et al., 1988, 1990, 1993] for an empirical specification of the state and condition of the I-T system and its weather and climatology as influenced by seasonal, solar cycle, diurnal and storm time controls. Working with that database, the objectives also contributed to the development of a predictive capability through the test, verification and validation of empirical and first principle models. Past SUNDIAL efforts have included the empirically derived International Reference Ionosphere (IRI) [Rawer, 1981; Rawer and Ramanamury, 1985; Schunk and Szaszczewicz, 1988] and the auroral oval models of Feldstein and the Television and Infrared Observation Satellite (TIROS) and Defense Meteorological Satellite Programs (DMSP) [Szaszczewicz et al., 1993]. Test and validation efforts have also included a number of first-principle models: the time-dependent ionospheric model of Utah State [Wilkinson et al., 1988; Sica et al., 1990], the Thermospheric-Ionospheric General Circulation Model TIGCM of the National Center for Atmospheric Research (NCAR) [Wilkinson et al., 1992, Szaszczewicz et al., 1995] and the Rice Convection Model [Spiro et al., 1988; Fejer et al., 1990].

In this investigation we concentrate on the IRI, the NCAR/TIGCM, and the Field Line Interhemispheric Plasma (FLIP) [Richards et al., 1994] models. They represent a unique set of capabilities that reflect major differences in approaches to the representation of the prevailing physics along with a substantial difference in computer requirements. The IRI, FLIP, and TIGCM are PC, work station, and Cray based, respectively. The IRI is purely empirical in its approach. The FLIP code is a one-dimensional model which inputs thermospheric controls by incorporating composition from the Mass Spectrometer and Incoherent Radar (MSIS) model [Hedin, 1987] and meridional winds from empirical models or from radars, ionosondes or Fabry-Perot interferometers. The NCAR/TIGCM might (arguably) be considered the closest to an all first-principles model since it self-consistently solves the system of equations which defines the coupled I-T domain. Rigorously speaking however, it is not entirely first principle since it invokes the empirically defined electric field model of Richmond et al. [1980] and the assimilative mapping of ionospheric electrodynamics (AMIE) specification [Richmond, 1992; Lu et al., 1995] of particles, fields, and currents at high latitudes. We present here the very first joint comparison of all three approaches against the global database with a perspective that highlights the differences in the approaches and the relative agreement with the data. We focus on climatology, drawing perspectives on diurnal variability, thermospheric wind controls, plasmaspheric fluxes, and phenomenological domains. We also include a perspective on the overall characteristics of the F region ionosphere as manifested during the March 23 to April 2, 1992, SUNDIAL/ATLAS 1 period.

2. The Database and the Models

2.1 The Data

2.1.1 Prevailing conditions. The SUNDIAL/ATLAS 1 campaign was conducted during the first equinoctial period of the descending phase of solar cycle 22. The months of March and April showed a dramatic decrease in energetic solar events compared with previous months. While there were 49 energetic solar X ray events (M class or higher) in February, there were only four in March and eight in April. Four M class
events occurred during the ATLAS 1 period: one on March 31 at 1545 UT and three on April 1 at 0100, 1017, and 1748 UT. The average uncorrected 10.7 cm flux for the ATLAS 1 period was 182, compared with averages of 171 and 158 for the entire months of March and April, respectively.

Solar wind observations for the campaign period were not available through March 26. But observations for the remainder of the campaign showed rather steady behavior with the solar wind velocity in the 325-425 km/s domain, and densities holding relatively steady near 10 cm\(^{-3}\). There was a brief period in which the density was enhanced by a factor of 3 (around 2100-2400 UT on March 31) but that was not accompanied by any increased velocity.

To develop a perspective on geomagnetic activity for the season and phase of the solar cycle in which the investigation took place, we present in Figure 1 the \(Kp\), \(Ap\) and \(-Dst\) indices. The three panels in that figure, top-to-bottom, show \(Kp\), \(Ap\), and \(-Dst\), respectively, for March and April 1992. The \(Kp\) and \(Ap\) values, 3±2 and 20±15, respectively, for the ATLAS 1 period (March 23 to April 2) were generally typical for the entire month of March. Overall, the conditions can be described as moderately active.

In the bottom panel of Figure 1 it can be seen that \(-Dst\) values were generally depressed during the ATLAS 1 period with values less than -50 considered "stormy" [Taylor et al., 1994; Kamide, 1994]. These results show that the entire ATLAS 1 period occurred during the recovery phase of a storm that took place at about 2300 UT on March 21. That recovery can be seen to be populated by a number of substorms [Lester et al., 1995] with \(Dst\) values gradually returning to zero on April 2, the quietest day of the entire campaign.

One final perspective on energy inputs and activity levels was provided by the interhemispheric power and activity indices determined by measurements on the NOAA/TIROS satellite [Foster et al., 1986]. These results corroborate the generally active nature of the ATLAS 1 period. The NOAA/TIROS activity indices were routinely between 4 and 8, with only two brief excursions to the maximum level of activity (i.e., 10) at about 1700 UT on March 24 and 0400 UT on March 29.

In companion papers, substorm influences on high-latitude conductivities are investigated during the periods March 27 and 30-31 [Lester et al., this issue], while the overlapping Geophysical Environmental Modeling (GEM) campaign days of March 28 and 29 are studied from perspectives focused on high-latitude inputs and substorms [Lu et al., this issue] and global-scale ionospheric-thermospheric variability [Emery et al., this issue]. These studies, along with the investigations of Wilkinson et al. [this issue], Abdu et al. [this issue], and Pulinets et al. [this issue], who focus on ionospheric weather and climatology in the 115°-165° longitude zone, at equatorial latitudes, and in the Russian sector, respectively, represent the ground-based measurement and modeling component of the campaign.

2.1.2. The Ionosonde database. The ground-based data were acquired by the global network of 53 ionosonde stations listed in Table 1, complemented by radar observations at EISCAT, Sondrestrom, Millstone Hill, Arecibo, and Jicamarca. While the radar observations operated only on a short schedule, the ionosonde measurements were carried out around-the-clock for the entire ATLAS 1 period with emphasis here on measured \(F\) region characteristics (\(f_0F_2\) and \(M(3000)F_2\)). The parameter \(M(3000)F_2\) is the transmission factor from which the real height \(h_mF_2\) of the \(F_2\) peak can be derived [Dudeney, 1983] and \(f_0F_2\) is the critical frequency of the \(F_2\) peak. Throughout the text \(f_0F_2\) will be used
<table>
<thead>
<tr>
<th>Name</th>
<th>Latitude</th>
<th>Longitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ahmedabad</td>
<td>23.0</td>
<td>72.6</td>
</tr>
<tr>
<td>Akita</td>
<td>39.7</td>
<td>140.1</td>
</tr>
<tr>
<td>Arkhangelsk</td>
<td>64.6</td>
<td>40.3</td>
</tr>
<tr>
<td>Buenos Aires</td>
<td>-34.6</td>
<td>301.7</td>
</tr>
<tr>
<td>Cachoeira Paulista</td>
<td>22.7</td>
<td>315.0</td>
</tr>
<tr>
<td>Canberra</td>
<td>-35.3</td>
<td>149.0</td>
</tr>
<tr>
<td>Canton</td>
<td>23.1</td>
<td>113.3</td>
</tr>
<tr>
<td>Chanchun</td>
<td>43.8</td>
<td>125.3</td>
</tr>
<tr>
<td>Cling-Li</td>
<td>25.0</td>
<td>121.2</td>
</tr>
<tr>
<td>Chungking</td>
<td>29.6</td>
<td>106.5</td>
</tr>
<tr>
<td>Churchill</td>
<td>58.7</td>
<td>265.8</td>
</tr>
<tr>
<td>Dakar</td>
<td>14.8</td>
<td>242.6</td>
</tr>
<tr>
<td>Darwin</td>
<td>-17.3</td>
<td>130.9</td>
</tr>
<tr>
<td>Fortaleza</td>
<td>-3.7</td>
<td>321.0</td>
</tr>
<tr>
<td>Hankou</td>
<td>30.7</td>
<td>114.3</td>
</tr>
<tr>
<td>Ileobart</td>
<td>-42.9</td>
<td>147.3</td>
</tr>
<tr>
<td>Irkutsk</td>
<td>52.5</td>
<td>104.0</td>
</tr>
<tr>
<td>Kaliningrad</td>
<td>54.7</td>
<td>20.6</td>
</tr>
<tr>
<td>Karaganda</td>
<td>49.8</td>
<td>73.1</td>
</tr>
<tr>
<td>Khabarovsk</td>
<td>48.5</td>
<td>135.2</td>
</tr>
<tr>
<td>Kiruna</td>
<td>67.8</td>
<td>20.4</td>
</tr>
<tr>
<td>Kodaikanal</td>
<td>10.2</td>
<td>77.7</td>
</tr>
<tr>
<td>La Reunion</td>
<td>21.2</td>
<td>55.6</td>
</tr>
<tr>
<td>Lanchou</td>
<td>36.1</td>
<td>103.8</td>
</tr>
<tr>
<td>Lantion</td>
<td>48.8</td>
<td>356.6</td>
</tr>
<tr>
<td>Leningrad</td>
<td>60.0</td>
<td>30.7</td>
</tr>
<tr>
<td>Lycksele</td>
<td>64.6</td>
<td>18.8</td>
</tr>
<tr>
<td>Magadan</td>
<td>60.1</td>
<td>151.0</td>
</tr>
<tr>
<td>Manchouli</td>
<td>49.6</td>
<td>117.4</td>
</tr>
<tr>
<td>Moscow</td>
<td>55.5</td>
<td>37.3</td>
</tr>
<tr>
<td>Mundaring</td>
<td>-32.0</td>
<td>116.2</td>
</tr>
<tr>
<td>Norfolk Is.</td>
<td>-29.0</td>
<td>168.0</td>
</tr>
<tr>
<td>Novosibirsk</td>
<td>54.6</td>
<td>83.2</td>
</tr>
<tr>
<td>Okinawa</td>
<td>26.3</td>
<td>127.8</td>
</tr>
<tr>
<td>Ottawa</td>
<td>45.4</td>
<td>284.1</td>
</tr>
<tr>
<td>Ouagadougou</td>
<td>17.4</td>
<td>358.5</td>
</tr>
<tr>
<td>Peiping</td>
<td>40.0</td>
<td>116.3</td>
</tr>
<tr>
<td>Pulieters</td>
<td>46.6</td>
<td>0.3</td>
</tr>
<tr>
<td>Resolute Bay</td>
<td>74.7</td>
<td>265.1</td>
</tr>
<tr>
<td>Rome</td>
<td>41.8</td>
<td>12.5</td>
</tr>
<tr>
<td>Sverdlovsk</td>
<td>56.7</td>
<td>61.1</td>
</tr>
<tr>
<td>Tahiti</td>
<td>-17.7</td>
<td>210.7</td>
</tr>
<tr>
<td>Tashkent</td>
<td>41.3</td>
<td>69.0</td>
</tr>
<tr>
<td>Townsville</td>
<td>-19.6</td>
<td>146.8</td>
</tr>
<tr>
<td>Trivandrum</td>
<td>8.2</td>
<td>77.0</td>
</tr>
<tr>
<td>Tungsuka</td>
<td>61.6</td>
<td>90.0</td>
</tr>
<tr>
<td>Uppsala</td>
<td>59.8</td>
<td>17.6</td>
</tr>
<tr>
<td>Ushuaia</td>
<td>-54.8</td>
<td>291.7</td>
</tr>
<tr>
<td>Vanimo</td>
<td>-2.7</td>
<td>141.3</td>
</tr>
<tr>
<td>Wakkanai</td>
<td>45.4</td>
<td>141.7</td>
</tr>
<tr>
<td>Waltair</td>
<td>18.0</td>
<td>83.0</td>
</tr>
<tr>
<td>Wulumuqui</td>
<td>43.7</td>
<td>87.6</td>
</tr>
<tr>
<td>Yamagawa</td>
<td>31.2</td>
<td>130.6</td>
</tr>
</tbody>
</table>

These references report that accuracy in $f_0F_2$ is typically 0.1 MHz (or equivalently 125 cm$^3$ in the density of the $F_2$ peak, $N_mF_2$) and inaccuracies in $M(3000)F_2$ can amount to random errors in $h_mF_2$ not in excess of ±10 km.

Figures 2 and 3 provide an introductory perspective on observed ground-based responses of the ionosphere to the prevailing solar-terrestrial conditions. Figure 2 presents a subset of "events" primarily observed in the northern hemisphere. These include a series of substorms, as identified in the European sector by the successors to the European incoherent scatter (EISCAT) magnetometer cross [Luhm et al., 1984], the midlatitude Subauroral Magnetometer Network (SAMNET) [Yeoman et al., 1990], and by positive and negative storm phases in the Russian and French ionosonde results.

In general, "stormy" ionospheric conditions were observed at high latitudes, while nominally quiet characteristics were observed at mid- and equatorial latitudes. This can be seen to some extent in Figure 3 which displays diurnal plots of $f_0F_2$ (and $N_mF_2$) for pairs of stations in the Russian/Fast-African, the Asian/Australian, and the American sectors. (The $f_0F_2$ values are expressed in megahertz throughout the paper with $f_0F_2 = 5$, 10 and 15 MHz equal to $N_mF_2 = 0.3$, 1.3 and 2.9(10$^6$) cm$^{-3}$, respectively. For convenience, the $N_mF_2$ scaling is displayed along the right-hand ordinates in Figure 3.) These station pairs (each pair formed by a high-northern and low-southern latitude site in approximately the same longitude sector) are Moscow/La Reunion, Magadan/Townsville, and Ottawa/Buenos Aires, respectively. The plots present the following: 1) the hourly observations as solid dots, 2) the average diurnal characteristic determined for the entire campaign period as a solid line, and 3) the IRI specification as a dashed line. It can be seen in the figure that the high-latitude sites show irregular day-to-day behavior. Note, for example, the negative and positive storm phases at Moscow on March 25 and 27, respectively. By contrast, the mid-latitude stations show very regular behavior, with the day-to-day characteristics virtually identical with the campaign-period average.

All ionosonde data were subjected to three levels of editing and quality control: the first by regional coordinators (e.g. S. Pulinets for Russian data, M.A. Abdu for South American data, etc.), the second at the initial staging center for data fusion (the Ionospheric Prediction Service, at Sydney, Australia), and the third at the SAIC Data Archive and Analysis Center where errors and inconsistencies were screened on regional and global bases using special purpose data handling and visualization tools [Szuszczewicz et al., 1995b]. This process mitigated possible contamination of subsequent model-measurement comparisons by spurious or otherwise erroneous data.

### 2.2. The Models

#### 2.2.1 The IRI

Currently the empirically derived IRI model is the most used and widely-tested climatological specification of the global ionosphere [Rawer, 1981; Rawer and Ramanamurty, 1985; Schunk and Szuszczewicz, 1988]. It describes the F and G regions of the ionosphere in terms of location (geographic or geomagnetic), time, solar activity (sunspot number), and season (month). By definition it is based on accumulated data, including critical frequencies ($f_0F_2$, $f_0E_S$, $f_0F_1$, and $f_0F_2$), altitudes of peak concentrations ($h_mF_2$, $h_mF_1$, and $h_mF_2$), half thicknesses of the individual layers, and
ion composition and plasma temperatures. The database forming the foundation of the IRI was compiled from ionosondes, topside sounders, incoherent scatter radars, and in situ satellite and rocket profiles (wherever available). Averaged profiles were synthesized and analytical fits developed, with the end product representing an intelligent mix of empiricism, physics, extrapolation, and intuition. In general, the quality of the IRI representation of actual climatological conditions is directly related to available data, its global distribution, parameter space coverage, time base, and accuracy. Obvious gaps in the database include ocean regions and ionogram definition of the E-F region valley.

Notwithstanding the limitations of the IRI, it can (and does) provide a valuable resource that includes a broad experience base in ionospheric profiling and data archiving. Its output provides a global specification not only on the macroscale features of electron density but also on the ion and electron temperatures and the ion composition. These additional features make the IRI an attractive model for comparison with satellite and ground-based data as well as first-principle endues.

The IRI is an evolutionary model, with its primary emphasis being an attempt at a global-scale representation of observations. In this regard the IRI makes its limitations clear, particularly with regard to the relative paucity of ion composition results. Work has continued in improving its specification of the topside domain [Bilitza, 1994], its specification of plasma temperatures and ion composition [Bilitza et al., 1992a], and its specification of global and mesoscale electron density distributions [Bilitza et al., 1992b; Zusczewicz et al., 1990, 1993]; while its inaccuracies at high latitudes can be partially mitigated by inclusion of the NOAA/TIROS, DMSP, or Feldstein auroral oval models [Zusczewicz et al., 1993, and references therein].

For comparisons in this investigation the IRI was run with the observed smoothed sunspot number of 108 and a seasonal behavior equivalent to the end of March and beginning of April (i.e., month = 3.5), a period conforming closely to the SUNDIAL/ATLAS-I campaign.

2.2.2. The FLIP model. The field line interhemispheric plasma model has been described previously by Richards and Torr [1988] and more recently by Torr et al. [1990] and Richards et al. [1994a]. The main component of this one-dimensional model calculates the plasma densities and temperatures along entire magnetic flux tubes from 80 km in the northern hemisphere through the plasmasphere to 80 km in the southern hemisphere using a tilted dipole approximation to the Earth's magnetic field. The FLIP model solves the continuity and momentum equations for $\phi^+$, $H^+$, $N^+$, and $He^+$. The electron and ion temperatures are obtained by solving the energy equations [Schunk and Nagy, 1978] with electron heating due to photoelectrons provided by a solution of the two-stream photoelectron flux equations using the method of Nagy and Banks [1970]. The solutions have been extended to encompass the entire field line on the same spatial grid as the ion continuity and momentum equations.

Once all the interhemispheric solutions have been determined, densities of $NO^+$, $O_2^+$, $N_2^+$ ($n \to 6$), $O^+(2D)$, and $O(2P)$ are obtained from local chemical equilibrium. Then diffusion equations are solved on separate vertical grids in both hemispheres to obtain the densities of the minor neutral species $N(2D)$, $N(4S)$, NO, and vibrationally excited $N_2$ on vertical grids extending from 80 km to about 500 km in each hemisphere.

The FLIP model includes provisions to solve the continuity and momentum equations for the first six vibrational levels of $N_2$ in order to take into account the strong dependence of the $O^+ + N_2 \to NO^+ + N$ reaction rate
Figure 3. Diurnal characteristics of measured values of $f_0F_2$ (in megahertz) at six of the ionosonde stations covering the period March 24 through April 4, 1992. Dots, solid curve, and dashed curve represent the data, the average diurnal behavior for the period, and the IRI "predictions", respectively. Note that $f_0F_2$ (hertz) = $8.9 \times 10^3 \left( N_mF_2 (\text{cm}^{-3}) \right)^{1/2}$ with the right-hand abscissa showing the conversion of $f_0F_2$ = 5, 10, and 15 MHz to $N_mF_2$ = 0.3, 1.3, and $2.8 \times 10^6$ cm$^{-3}$, respectively.

on vibrational excitation of $N_2$ [Richards et al., 1986; Richards and Torr, 1986a]. However, vibrationally excited $N_2$ was not included in the present investigation, noting that in a recent study of ionospheric behavior at Millstone Hill and Hobart during 1990, the FLIP model was able to reproduce the observed daytime densities extremely well without the need for an accounting of vibrationally-excited $N_2$ [Richards et al., 1994a].

Most important reaction rates and cross sections that are used in the FLIP model are now reasonably well established [Torr et al., 1990], and some innovative algorithms have been developed in recent years to incorporate actual measurements (e.g., F region heights or meridional winds) into the FLIP model when they are available [Richards et al., 1995].

The FLIP model requires the following three key inputs: (1) the neutral atmosphere, (2) either the meridional component of the neutral wind or $V_mF_2$ (which is converted into a wind), (3) and the solar EUV flux. The neutral densities and temperatures are provided by the mass spectrometer and incoherent scatter (MSIS-86) model [Hedin, 1987]. Once the time and location of the required thermospheric parameters are specified, the only adjustable parameters in the MSIS model are the solar index (F10.7) and magnetic index (Ap). Both the solar activity and magnetic activity indices were stable in late March 1992. Under these circumstances, there was little day-to-day variation in the model calculation and it was possible to simulate just one day as being representative of the whole period, with no adjustments made to the input parameters to obtain better agreement with the data.

For this investigation the FLIP model winds were obtained
from the International Reference Ionosphere (IRI) model specifications for \( h_m F_2 \) during March 1992 using the servo-analysis method of Richards [1991]. For climatological studies, this method is preferred over using winds from the HWM90 model of Hedin et al. [1991] because the IRI \( h_m F_2 \) is based on a much larger database with greater global coverage.

For the EUV fluxes in FLIP we have used a new solar flux model called EUVAC [Richards et al., 1994b] which is based on the F74113 solar EUV reference spectrum and the relative variations of the EUV flux observed by the Atmosphere Explorer E satellite. The EUVAC model produces 50-575 Å integrated fluxes that agree well with the most reliable rocket measurements.

Apart from the uncertainties in the neutral densities, winds, and EUV flux, there are uncertainties in the degree of particle and energy coupling between the ionosphere and plasmasphere. Although the FLIP model self-consistently includes coupling to the plasmasphere and the conjugate ionosphere, there can be large uncertainties in the plasma and heat content of the plasmasphere. There is currently no way of measuring or predicting the state of the plasmasphere as it recovers after magnetic storm depletions. These uncertainties in the plasmasphere have little effect on the daytime ionosphere, but there may be up to a factor of two difference in the calculated \( N_m F_2 \) (or equivalently a factor of 1.4 in the calculated \( f_0 F_2 \)) at night for extremes of assumptions about the plasmaspheric conditions. In this paper, the flux tubes were moderately depleted at the beginning of the run but filled up during the simulation periods. Thus flux tubes below \( L = 2 \), which fill up quickly, were close to full, while longer flux tubes were about half full during the simulation period.

The ionospheres in the auroral and equatorial regions are extremely difficult to model quantitatively with FLIP (and most other models). In the auroral zone there are electric fields, particle precipitation, and changes in neutral composition which are difficult to predict, while in the equatorial region the variation of the electric fields is the main cause of uncertainty. The FLIP model does not model the electron densities well in the equatorial and auroral zones primarily because it does not currently include convection electric fields. The model is now being modified to account for this.

### 2.2.3. The TIGCM

The NCAR TIGCM self-consistently calculates thermospheric and ionospheric structure, taking into account the dynamic coupling between the thermospheric neutral wind and the ionospheric plasma [Roble et al., 1988]. It is an outgrowth of the NCAR Thermospheric General Circulation Model (TGCM) which included a non-reactive empirical ionosphere and solved the primitive equations of dynamic meteorology adapted to the physics appropriate to thermospheric heights. The basic TIGCM and associated developments have been described in detail by Dickinson et al. [1981, 1984], Roble et al. [1987], Fesen et al. [1986], and Roble and Ridley [1987].

The TIGCM employs an Eulerian approach to modeling the ionosphere as an alternative to the Lagrangian approach used by Sofka and Schunk [1985] and Fuller-Rowell et al. [1987]. Equations describing the ionosphere and thermosphere are both solved on the TGCM geographic grid and ion drift for the ionospheric calculation is obtained from the empirical model of Richmond et al. [1980] for low latitudes and mid latitudes while the empirical model of Heelet al. [1982] or the Assimilative Mapping of Ionospheric Electrodynamics (AMIE) procedure [Richmond, 1992] is used for high latitudes. Therefore the model does not self-consistently solve for the naturally occurring electrodynamic coupling in the magnetospheric-ionospheric-thermospheric system. The TIGCM also employs a simple dipole magnetic field model with displaced geomagnetic and geographic poles. For inputs, the model requires specifications of several external sources. These include solar EUV and UV fluxes, auroral particle precipitation, the ionospheric convection pattern, and the amplitudes and phases of semi diurnal tides at the mesopause boundary.

The odd nitrogen chemistry is the same as that discussed by Roble et al. [1987]. \( \text{N}^2 \text{D} \) is assumed to be in photochemical equilibrium throughout the thermosphere and transport equations are solved for \( \text{N}^4 \text{S} \) and \( \text{NO} \). The model considers \( \text{N}^4 \text{S} \) in photochemical equilibrium and specifies a constant number density of \( 8 \times 10^{6} \text{cm}^{-3} \) for \( \text{NO} \) at the lower boundary.

The ion-chemistry scheme, rate coefficients, and calculation procedure for the ionosphere can be reviewed in the works by Roble et al. [1987] and Roble and Ridley [1987], while the bulk motion includes both magnetic field-aligned diffusion and \( \mathbf{E} \times \mathbf{B} \) forces in solving transport equations.

The upper boundary condition is diffusive equilibrium. The upper altitude limit of the model is about 550 km where an analytical formulation is used as a boundary condition for positive and negative flux transfer. This flux, which contributes to the overall density-height profile and equilibrium values of \( f_0 F_2 (N_m F_2) \) and \( h_m F_2 \), results from normal topside dynamics (e.g., wind and electric field induced transport) as well as inputs from the conjugate ionosphere and fluxtube-coupled regions of the outer plasmasphere. Under equinox conditions with full fluxtubes, \( \text{O}^+ \) flows up from the ionosphere during the day, charge exchanges with neutral \( \text{H} \) to produce \( \text{H}^+ \), which in turn is stored in the outer plasmasphere. The process reverses at night, with the end result being an oversimplified diurnal characteristic which consists of an ebb and flow of plasma between the ionosphere and outer plasmasphere. This process is referred alternatively as plasmaspheric flux, magnetospheric flux, protonospheric flux, or ionospheric-plasmaspheric flux.

The lower boundary condition assumes photochemical equilibrium. Once the \( \text{O}^+ \) density distribution is defined, the \( \text{NO}^+, \text{O}_2^+, \text{N}_2^+ \) and \( \text{N}^+ \) distributions are determined by assuming photochemical equilibrium [Roble and Ridley, 1987]. The temperature, dynamics, composition, and the global ionosphere are calculated at each time step and the heating and cooling rates from chemical reactions, IR emissions, etc. arc coupled into the thermodynamic equation at the next time step. Likewise, neutral winds, temperature, composition, and ionization rates are updated at each time step for the ionospheric calculation.

The TIGCM output provides global distributions of the neutral gas temperature and winds, ion drift velocities, the heights of constant pressure surfaces, and the number densities of \( \text{O}_2, \text{N}_2, \text{O}, \text{N}^4 \text{D}, \text{N}^4 \text{S}, \text{NO}, \text{He}, \text{and} \text{Ar} \). It also solves for the ionospheric structure, giving global distributions of \( \text{O}^+, \text{O}_2^+, \text{NO}^+, \text{N}_2^+, \text{N}^+, \) electron density, and ion and electron temperatures. The TIGCM has an effective 5° latitude-longitude geographic grid with 25 constant pressure levels in the altitude range from approximately 97 km to 500 km. The time step can be flexible, varying from 3 min to 6 min, depending on the inputs. The model was exercised in a “weather” mode with an emphasis on March 28 and 29 in
Plate 1. Global representation (1200 UT) of station locations and model-measurement comparisons of observed vs modeled values of \( f_0F_2 \). The color legend on the bottom left presents the percent differences between the data at each station (colored circles) and the IRI, the TIGCM and the FLIP model predictions in the top, middle, and bottom panels, respectively. See text for other details. Extrema on the \( f_0F_2 \) color scale (legend on bottom right side) for the isocontours are 3 and 16 MHz corresponding to extrema in \( N_mF_2 \) of \( 1.1(10^5) \) and \( 3.2(10^9) \) cm\(^{-3} \), respectively. The red-to-blue color scale for the percent differences of model vs measurements covers the +40% to -40% range. (Note that differences of 20% and 40% in model-measurement comparisons of \( f_0F_2 \) are approximately equal to differences of 45% and 100%, respectively, in corresponding \( N_mF_2 \) comparisons.)
Plate 2. Global representation (1200 UT) of station locations and model-measurement comparisons of observed vs modeled values of $h_mF_2$. Color legend presents the percent differences between the data and the IRI, the TIGCM, and the FLIP model specifications in the top, middle, and bottom panels, respectively. See text for other details.
Plate 3. Global representation (1200 UT) of measured and modeled meridional winds at 300 km (for the models) and at $h_mF_2$ for the observations. Magnitude and directions are shown by vectors and color code at each of the station locations. Each panel is identified, and the $f_0F_2$ contours in the individual panels correspond to the model in question. In the data and HWM panels, the contours of $f_0F_2$ are from the IRI model.
studies of high-latitude electrodynamics [Li et al., 1994; Emery et al., 1995]. It was also run in a “climatological” mode in order to specify the baseline conditions upon which to initiate the study of electrodynamics on the 28 and 29. The "climatological" run, described more fully by Emery et al. [this issue], is the basis for the TIGCM model-measurement comparisons carried out in this investigation.

For March 28 and 29 the average solar radiation index 5a (i.e., the 10.7 cm flux) was 188.5 x 10^{-22} W/m^2 Hz. This was the fundamental input driving the specification of the solar EUV and UV fluxes. For energetic inputs at high latitudes, ionospheric convection and particle precipitation were derived using the Assimilative Mapping of Ionospheric ElectroDynamics (AMIE) procedure with approximately 50-min resolution, while low latitude and mid latitude ionospheric convection were derived from the empirical model of Richmond et al. [1980].

The semi-diurnal tide model, described by Fesen et al. [1991], was included at the lower boundary of the TIGCM, and the model was first run to obtain a diurnally reproducible solution. For a “climatological” run, the patterns of the ionospheric convection and particle precipitation used in this investigation were an averaged AMIE prescription, derived from detailed data acquired during the 48 hour period March 28-29, 1992, and described in the companion paper by Emery et al. [this issue]. The time-dependent AMIE pattern were then incorporated into the TIGCM for a simulation of high-latitude thermospheric dynamics.

3. Model-Measurement Comparisons

3.1. Global Perspectives

Plates 1 and 2 provide global comparisons of the observed values of \( f_0 F_2 \) and \( h_m F_2 \), respectively, at 1200 UT with the "predictions" of the IRI, the TIGCM, and the FLIP models (top, middle, and bottom panels, respectively). The observations at each station represent the value averaged over the campaign period for the UT (or L.T) in question. The colored circles in the figures locate the stations, with the percent difference between the model predictions and the individual station observations defined by the color legend on the bottom left side of Plates 1 and 2. The bottom right-hand color bar defines the range of values for the color-coded isocontours in the plates (for \( f_0 F_2 \) and \( h_m F_2 \) in Plates 1 and 2, respectively) which correspond to the outputs of each of the models in their respective panels. The abscissa is longitude with 0° at the center. The ordinate is latitude, extending north and south to ±90° from the geographic equator. The \( f_0 F_2 \) contours are color-coded from a low value of 3 MHz (dark blue, \( N_m F_2 = 1 \times 10^5 \) cm⁻³) to a high value of 16 MHz (red, \( N_m F_2 = 3.2 \times 10^6 \) cm⁻³). The \( h_m F_2 \) contours are also color-coded with a red-to-blue range extending from 300 to 230 km. The black curves at high northern/southern latitudes specify the boundaries of the auroral oval according to the Feldstein model for a moderate activity level \( Q = 4 \) [Feldstein, 1963; Holtworth and Meng, 1975].

The \( f_0 F_2 \) features (e.g., in the IRI panel) which dominate Plate 1 include (1) the sunrise terminator in the American sector rising in a span of approximately one hour from 5 to 10 MHz (i.e., \( N_m F_2 \) from 0.3 to 1.3 \( 10^6 \) cm⁻³), and (2) the Appleton Anomaly with its peak densities somewhat symmetrically centered about the dip equator and extending in local time from 0900 to midnight. (The center of the plate corresponds to LT=1200 hours, while the left and right-hand edges correspond to LT=0 hours.) These large-scale features, along with the auroral oval boundaries, help identify important phenomenological domains while providing visual aides to the local time.

The \( f_0 F_2 \) contours in Plate 1 show that the largest densities are at the daytime peaks of the Appleton Anomaly, with the highest value near 16 MHz (\( N_m F_2 = 3.2 \times 10^6 \) cm⁻³) at LT = 1400 utpating off from the individual peaks to lower values equatorward and poleward. The lowest values are in the pre-dawn sector, with magnitudes near 5 MHz (\( N_m F_2 = 3 \times 10^5 \) cm⁻³).

A review of the IRI panel in Plate 2 reveals maximum values of \( h_m F_2 \) at the daytime equator with magnitudes in the 425 km range. F region heights fall off poleward of the equatorial region and at all latitudes in evening and night time hours (LT ≥ 8 PM). Lowest F region heights are in the dawn sector at middle-to-high latitudes where values are near 275 km.

The FLIP and TIGCM panels in Plates 1 and 2 show characteristics similar to those described in the IRI panel with the exception that the FLIP model does not reproduce the well-known double peak of the Appleton Anomaly. This is primarily due to the fact that FLIP is designed for midlatitudes were it accounts for winds, chemistry, and plasmaspheric fluxes. It has no convection electric fields and therefore its results are compromised at low-to-equatorial latitudes where electric fields are important.

Each ionosonde site in Plates 1 and 2 is identified by a colored circle, with the color representing the percent difference between the observed value (\( f_0 F_2 \) or \( h_m F_2 \)) and the model prediction (i.e., percent difference = 100 (data - model)/data). The differences are presented on a scale of -40% to +40% in \( f_0 F_2 \) and -25% to +25% in \( h_m F_2 \) (in both cases over the range blue to red as scaled in the legend).

We offer the following observations relative to Plate 1:

1. The IRI underestimates the values of \( f_0 F_2 \), with the primary differences ranging from 0% (green) to 20% (yellow). There is no obvious LT dependence in the differences. Within a few percent this has been found to be true for all UT intervals.

2. Both the FLIP and TIGCM predictions demonstrate better agreement with the \( f_0 F_2 \) data in daytime (more greens), and both depart in major ways from the data in the post-sunset to pre-sunrise period. In the nighttime hours the FLIP model and TIGCM tend to underestimate the critical frequencies of the F region peak by 20-40%. This is a consistent feature that prevailed for all UT intervals.

Reference to Plate 2 leads to the following observations regarding \( h_m F_2 \):

1. All three models either agree with the data (green) or underestimate the heights of the F region peaks by as much as 20 - 25% (orange to red).

2. The level of agreement between data and model are comparable for both the FLIP and TIGCM, whereas the overall results suggest that the IRI has the poorest representation of \( h_m F_2 \) among the three model comparisons displayed in Plate 2. This will be shown to be generally true but not by a large margin.

Before leaving the discussion of Plate 2 we note that the ionosonde values of \( h_m F_2 \) were derived from the \( M(3000)F_2 \) algorithm of Dudeney [1983]. Had we used the \( h_m F_2 \)
formulism of Bilitza et al. [1979] instead, there would have been little-to-no difference at night and smaller ionosonde-derived values of $h_m F_2$ during the day. But the daytime difference between the Dudeney and Bilitza et al. approaches would generally be less than 2.5%.

3.2. An Integrated View of LT and Latitudinal Effects

Figure 4 provides an integrated view of the model-measurement comparisons with a perspective that establishes quantitative figures of merit for each of the models. The results in Figure 4 also help develop an understanding of inaccuracies in the model predictions that might be identified with geophysical domains or LT dependence.

The top, middle, and bottom panels present comparisons of the campaign-averaged data with the IRI, FLIP and TIGCM climatological "predictions" in regions for which abs(MLAT) < 25°, abs(MLAT) > 25°, and 0° ≤ abs(MLAT) ≤ 90°, respectively. The latter case involves all stations, while abs(MLAT) < 25° and >25° refer to model comparisons with the equatorial-to-low latitude and mid-to-high latitude
stations, respectively. The left column deals with $f_0F_2$ (with appropriate scaling for $N_mF_2$ along the right-side ordinate) and the right panel deals with $h_mF_2$. The abscissa is local time (LT), while the ordinate is the percent difference between the data and the model. (Positive values occur when the magnitudes of the observations are greater than the predictions of the model, that is, when the models underestimate the observable.) Each LT plot has been integrated over all universal times with a sliding 3 hour LT window for each UT from 0 through 23 hours.

Focusing first on the bottom panels of Figure 4 (comparisons with all stations) we offer the following conclusions:

1. The FLIP and TIGCM do very well in $f_0F_2$ ($N_mF_2$) daytime comparisons ($6 < \delta$[hours] < 18), underestimating and overestimating the data by +3.2% (+6.1%) and -3.1% (-6.2%), respectively. (We note that positive percentages mean that the data values are in excess of the model values.) In the same LT period the IRI underestimates the data by +12.2% (+25.3%) with an overall diurnal characteristic that is virtually LT-independent.

2. The FLIP and TIGCM predictions for $f_0F_2$ ($N_mF_2$) show a marked departure from agreement with the observations in the post-sunset to pre-sunrise period (i.e., 18 < LT < 6), rising sharply from excellent agreement at sunset to disagreements at levels near 45% (110%) and 28% (64%) for the FLIP and TIGCM models, respectively, near LT = 4. An initial suggestion for the cause of this discrepancy is the "old unsolved problem" involving the maintenance of the nighttime ionosphere via plasmaspheric fluxes [e.g., Evans, 1975; Sica et al., 1990]. We will discuss this issue more fully in subsequent sections, but first move on to other observations in Figure 4 including discrepancies in $h_mF_2$ and their associated impact on $f_0F_2$ values.

3. The bottom right panel in Figure 4 compares the observed values of $h_mF_2$ (all stations) with the model predictions. The results show that all models underestimate the observations by values less than 13%, with the IRI having the poorest and the TIGCM the best overall agreement with the data. In the postsunset to presunrise period the FLIP and TIGCM are nearly identical, generally underestimating $h_mF_2$ by 5±2%. At 300 km this is a 15 km underestimate, or equivalently a half scale height for $N_2$. If $O^+ + N_2 \rightarrow NO^+ + N$ and $NO^+ + e \rightarrow N + O$ were the primary loss channels for F peak densities, a theoretical underestimate in $h_mF_2$ by a half scale height in $N_2$ (i.e., 15 km) would account for a 50% increase in the loss rate of $O^+$ ions. This is indeed a potential contribution to the observed discrepancies in $f_0F_2$.

4. Continuing to review the results in the bottom-right panel of Figure 4, it can be seen that the FLIP and TIGCM models also underestimate the values of $h_mF_2$ during the daytime, with FLIP generally worse than the TIGCM. If all other physical processes in the two models were identical, one would expect (assuming the same principle loss mechanism discussed in item 3) that the FLIP model would predict lower values than the TIGCM for $f_0F_2$. On a relative scale this is consistent with the daytime findings in the bottom left panel of Figure 4 where FLIP underestimates and the TIGCM overestimates the values of $f_0F_2$.

5. Turning attention to the middle and upper panels of Figure 4, it can be seen that the post-sunset to pre-sunrise discrepancy in $f_0F_2$ exists at middle-to-high and low-to-equatorial latitudes, respectively. This would suggest that downward transport of $H^+$ at night and subsequent charge exchange with $O$ to form increased levels of $O^+$ is not the cause for the $f_0F_2$ model-measurement discrepancy. This is because such downward plasmaspheric transport of $H^+$ is effectively zero at low-to-equatorial latitudes. This interpretation of results is not without counter and supporting arguments. Note for example that the pre-sunrise TIGCM results for $f_0F_2$ are generally better at low latitudes than at mid-to-high latitudes, suggesting that flux is at least part of the problem. And while the FLIP results at night are worse at low-to-equatorial latitudes, the model includes no electric fields which have a dominant control of heights and densities at equatorial latitudes. At night these electric field controls tend to move $h_mF_2$ to lower altitudes, thus reducing the plasma densities below those currently predicted by the FLIP model. This suggests that the problem is neither electric fields nor plasmaspheric fluxes.

6. Final observations in Figure 4 focus on $h_mF_2$, where the equatorial-to-low latitude results (top right panel) are uniquely different from those at mid-to-high latitudes. The primary difference is at post-sunset and to a lesser degree at presunrise. The postsunset discrepancy is most likely the result of the well-known postsunset prereversal enhancement in F region heights at equatorial latitudes. This is a dynamo effect, not included in the simple electric field model of Richmond et al. [1980] that is employed in the TIGCM. It is interesting to note that it is also not accounted for in the empirically based IRI, suggesting too little data during that period of the day or too much smoothing of the results. The postsunset enhancement, if we allow ourselves to call it that, is a new observation, with causality most likely associated with ionospheric expansion due to heating at F region sunrise. This issue will be taken up again in section 1.4.

3.3. Effects of Thermospheric Winds

The results in Figure 4 suggest a consistency between the $f_0F_2$ and $h_mF_2$ observations, in that model underestimates of $f_0F_2$ were correlated with underestimates of $h_mF_2$. And this correlation was supported by simple arguments of recombination chemistry. In this section, further analysis of the data allows causality to be traced to model representations.
of meridional winds, known to have strong influences on $h_mF_2$ at mid latitudes.

Plate 3 presents a global perspective of magnetic meridional winds as predicted by the TIGCM, the empirically-based Horizontal Wind Model (HWM) [Hedin, 1991], and the FLIP model. Plate 3 also includes the magnetic meridional winds measured by the SUNDIAL network of ionosondes using the $h_mF_2$ servo analysis procedures of Miller et al. [1993]; and “A fast method for the derivation of meridional winds from the height of the ionospheric $F_2$ region,” submitted to J. Geophys. Res., 1995) with the Dudeney [1983] formulation for $h_mF_2$. (With regard to the accuracy of this $h_mF_2$ servo-analysis procedure we note that several comparisons have been made of winds derived from $h_mF_2$ with meridional winds derived from other methods. These have been primarily with incoherent scatter radars [see, e.g., Iligi et al., 1995], although Miller et al. [1986] and Miller [1989] also compared results with winds measured by Fabry-Perot interferometry. In each of these comparisons the winds derived from $h_mF_2$ measurements were found to be consistent with the results of the other techniques, both in magnitude and in the pattern of variability.) This $h_mF_2$ servo-analysis technique was applied to all ionosonde stations regardless of latitude, but with the caveat that the results are not expected to be realistic representations of meridional winds in regions for which abs(MLAT) < 20°.

The presentation of meridional winds in Plate 3 is at UT = 12 with $f_0F_2$ contours and auralor oval boundaries as previously described. Wind magnitudes are depicted by scaled vector lengths and by color codes at each of the stations. Plate 3 is useful in so far as it develops an appreciation for station distributions and general perspectives on variations with latitude and local time. A review of Plate 3 shows marked differences between and among the findings of the TIGCM, the HWM, and the FLIP predictions, with the TIGCM appearing to be in best agreement with the ionosonde-derived winds. This is in fact consistent with a general finding when a plot showing LT-dependence (integrated over all UTs) is developed for detailed comparisons between measured and modeled winds. Such a comparison is presented in Figure 5 where the abscissa is LT and the ordinate is the magnitude of meridional winds predicted by each of the models or measured by the ground-based network of ionosondes. Only stations in the region 20° < abs(MLAT) < 60° were considered in this comparison since this is the general bound of applicability of the servo-analysis technique. Positive winds are poleward, while negative winds are equatorward.

Review of Figure 5 leads to the following observations:
1. There is a large spread in values in both daytime and nighttime sectors. The daytime spread is 0 to 40 m/s, while the nighttime spread is 30 to 90 m/s.
2. The model in best agreement with the data is the TIGCM.
3. At night the FLIP and TIGCM winds are both less than the observations, suggesting that their values of $h_mF_2$ should be less than those observed. This is consistent with the findings in Figure 4 as well as with the argument that weaker equatorward winds result in lower values of $h_mF_2$. Hence we have established (at least in part) model underestimates of nighttime meridional winds as the cause for lower nighttime model predictions of $h_mF_2$ and $f_0F_2$ (i.e., $N_mF_2$).
4. During the day the observations show very light winds, generally ±10 m/s, whereas FLIP and HWM are typically in the 25 m/s range. Both FLIP and the TIGCM predict stronger daytime winds than those observed. FLIP winds are stronger than those of the TIGCM suggesting lower values of $h_mF_2$ than those predicted by TIGCM. This is consistent with the comparisons discussed in Figure 4.

In closing the discussion of Figure 5 we note that the low ionosonde values for daytime winds are corroborated by the mid-latitude observations of Iligi et al. [1995] during summer and equinox periods using both ionosonde and MU radar techniques. This corroboration, taken in concert with the global observations presented here and associated model comparisons, suggests that current wind models must be more carefully evaluated not only under nighttime conditions (as discussed above in connection with maintenance of the nighttime ionosphere) but in the daytime as well.

### 3.4 Modeling the Plasmaspheric Flux.

The previous sections have established a cause-effect path which links model underestimates in nighttime values of $f_0F_2$ to underestimates in $h_mF_2$ and finally to underestimates in the meridional component of the prevailing thermospheric winds. Ancillary arguments have also suggested that the nighttime discrepancy in $f_0F_2$ might be due in part to inaccurate model specification of plasmaspheric fluxes.

The issue of plasmaspheric fluxes and their contribution to the maintenance of the nighttime ionosphere has itself been somewhat of a controversy without definitive resolution. Part of the problem is the difficulty in the measurement itself. The early work of Evans [1975] suggested that the incoherent scatter radar technique was perhaps the best approach, but even it, according to Evans, has difficulty in uniquely and quantitatively determining the flux. Among the difficulties is the signal-to-noise ratio at night and the need to eliminate layer dynamics (i.e., induced by winds and electric fields) in order to determine the true plasmaspheric contributions. In his study of five nights (near local midnight at Millstone Hill) spanning four seasons, he found average fluxes at (3(10^-7) cm^-2 s^-1. He also found that H+ was a minor ion (never greater than 25%) over the altitude range over which useful measurements were made (i.e., < 800 km), and that the majority of the results exhibited a decrease in the O+ flux with altitude. This altitude dependence indicated that the layers were descending and corrections had to be made. The maximum flux values were at or near $h_mF_2$ and approached 2(10^-9) cm^-2 s^-1, a value clearly dominated by topside layer dynamics. In the end he concluded that large increases in plasmaspheric fluxes, that could give rise to nighttime increases in total electron content, did not appear to be seen. This conclusion was in keeping with that of Jain and Williams [1974]. We pursue this issue further with a discussion of these fluxes as specified within the TIGCM and the FLIP models.

The TIGCM invokes a sinusoidal latitudinal variation of this flux in which it is defined as zero at the magnetic equator and constant above 60° mlt. The flux model is discontinuous at 60°, with day/night values just equatorward set at +2(10^-6/ -2(10^-6) cm^-2 s^-1, and values of 3(10^-6/ -1(10^-6) cm^-2 s just poleward of that. The region above 60° is constant in latitude, while below 60° it goes to zero at the magnetic equator in a cosine fashion. There is no LT dependence in the flux model with the exception of an interpolation routine from the day/night levels between zenith angles of 80° and 110°. We note that the TIGCM includes no H+ transport or
associated charge-exchange chemistry. This flux model is applied at the upper altitude limit of the TIGCM (approximately 550 km) and is intended to include all effects on O\(^+\) (i.e., topside layer dynamics and protonospheric fluxes).

By contrast, the FLIP model approaches the plasmaspheric flux issue self-consistently along fluxtubes from hemisphere to hemisphere. It has already been pointed out that uncertainties in the fluxes between the ionosphere and the plasmasphere have little effect on the daytime ionosphere, but there may be up to a factor of two difference in the calculated \(N_{\text{p}}F_2\) densities (or equivalently a factor of 1.4 in the calculated values of \(f_0F_2\)) at night for extremes in assumptions about the plasmaspheric conditions. In this work, the fluxtubes in the FLIP simulations were moderately depleted at the beginning of the run but filled up after several cycles of the simulation. The expectation therefore is that maximum plasmaspheric flux is already making its contribution to the nighttime \(f_0F_2\) values.

Figure 6 plots the flux levels derived in the FLIP model in the 20\(^\circ\) meridian for 30\(^\circ\), 50\(^\circ\), and 70\(^\circ\) northern geographic latitudes at the 550-km level. These fluxes represent values that integrate the effects on O\(^+\) of layer movement due to winds and electric fields, photoproduction (in the daytime), heating and cooling of the plasma, and ionospheric-plasmaspheric exchange. The 550 km height is the approximate topside boundary for the TIGCM and has been selected for comparison with its analytical flux model.

The interpretation of the fluxes in Figure 6 is complicated by the fact that 550 km is in an altitude regime where measurements by Evans [1975] have indicated that the daytime flux changes sign from downwards at lower altitudes to upwards at higher altitudes. Below this "transition" level, excess ions flow down under the influence of gravity to regions of highest loss in reactions with N\(_2\) and O\(_2\), while above this altitude O\(^+\) ions flow upwards, and through charge exchange with H\(_2\), produce H\(^+\) ions. Some of the flux variation in Figure 6 is due to movement of the altitude of the transition between upward and downward flux. Figure 6 shows that there are three prominent features of the diurnal flux variation: the sunrise and midday upward flux maxima, and the post-sunset downward flux maximum. (The southern hemisphere O\(^+\) flux variation shows similar features to the northern hemisphere variation discussed here.) The sunrise peak on the O\(^+\) flux is principally a consequence of the thermal expansion of the ionospheric plasma. The low density in the morning allows the electron temperature to rise steeply, causing a large increase in scale height that is seen as a large flux in the topside ionosphere [Richards and Tarr, 1986]. Because of the tilt of the magnetic field, the sun rises more than half an hour earlier in the southern hemisphere than in the northern hemisphere and this causes pre-dawn heating which leads to the initial upward surge in flux. From the point of view of photoelectron heating, sunrise occurs at about 100\(^\circ\) solar zenith angle compared to about 90\(^\circ\) for photoionization. The upward flux receives an additional boost at local sunrise. The separation of the two sunrises produces an inflection on the flux curves for 30\(^\circ\)N and 50\(^\circ\)N geographic latitudes. The sunrise effect is not as dramatic at 70\(^\circ\)N because the changes in electron temperature are likewise not as dramatic. Heat from the plasmasphere keeps the ionospheric electron temperature high at night.

After the sunrise increase in flux there is an abatement until mid morning due to a decrease in the electron temperature as the plasma density builds up and also because of a poleward wind which pushes ions down the field lines.

The flux increases after the morning low as a result of an abatement of the poleward wind and the substantial photo-production of ions which peaks near noon. The decrease in the magnitude of the midday O\(^+\) upward flux with increasing latitude is partially due to the lowering of the layer as a result of a latitudinal increase in the poleward wind. The daytime upward O\(^+\) fluxes at 50 degrees latitude are comparable to the fluxes observed by Evans [1975].

The large downward flux after sunset is caused by the thermal collapse of the ionosphere. This collapse is exaggerated by the fact that the electron cooling is large because plasma density remains high while the heating is shut off. The effect decreases with latitude because the plasmaspheric heat reservoir reduces the severity of the thermal collapse. In addition, larger equatorward winds help to keep the ionization higher at high latitudes.

The flux characteristics in Figure 6 are clearly not like those in the analytical formulation of the TIGCM. For immediate purposes, however, we focus on the postmidnight to presunrise period where we have documented the discrepancy between observed and modeled values of \(f_0F_2\) and where we seek to confirm or deny plasmaspheric flux as a causality. Remembering the cosine characterization of nighttime flux in the TIGCM with a maximum downward value of \(2 \times 10^{10} \text{ cm}^2 \text{s}^{-1}\) at 60\(^\circ\) MLAT, we focus on the FLIP values at 50\(^\circ\)N (\(\approx 50\) MLAT at 20\(^\circ\)E longitude). Those values vary from \(-1.5 \times 10^{10}\) cm\(^2\) s\(^{-1}\) at L.T. = 0 to zero at L.T. = 4. If the FLIP results were to be the guide, and consistency with the observations of Evans [1975] an empirical boundary, we conclude the following: (1) that the filled-fluxtubes fluxes of FLIP represent a maximized "saturated" plasmaspheric input to the postmidnight ionosphere, (2) the TIGCM fluxes, while requiring a more realistic L.T. and latitudinal dependence, need not have fluxes increased in the post-midnight period, and (3) that thermospheric winds, not plasmaspheric flux, is the primary cause for FLIP and TIGCM model underestimates of the densities and heights of the nighttime F region.

**Figure 6.** O\(^+\) fluxes (units/cm\(^2\) s) at 550 km from the FLIP model in the 20\(^\circ\)E meridian at 30\(^\circ\)N (solid curve), 50\(^\circ\)N (dotted curve), and 70\(^\circ\)N (dashed curve) geographic latitude.
4. Comments and Conclusions

The overall comparison between the data and the models suggests that the physics of the daytime F region climatology is reasonably well understood. This is based on the March-April 1992 period of the SUNDIAL/ATLAS 1 campaign and the accuracies between the $f_0F_2$ (i.e., $N_mF_2$) predictions of the IRI, the FLIP and the TIGCM models, and the database of 53 ionosondes. Reference to Table 2 shows that the IRI, FLIP and TIGCM model specifications of daytime $N_mF_2$ values (6≤LT≤18) agree with the full complement of stations within accuracies of 57±19%, $N_mF_2$ and 236±4%, respectively, with variations in these percentages when the comparisons are restricted to equatorial-to-low latitudes (i.e., $\text{abs(MLAT)} < 25^\circ$) and middle-to-high latitudes (i.e., $\text{abs(MLAT)} > 25^\circ$).

There is a marked difference in the accuracies in the afternoon to presunrise period, with the maximum disagreements occurring in the postmidnight period at about 0400 LT. Defining this postmidnight period as LT = 0 to 5 hours, Table 2 shows the inaccuracy in FLIP and the TIGCM specifications for $N_mF_2$ can be 82±21% and 39±17%, respectively, when the results are integrated over all stations. These results, in effect, rediscovered the old but unsettled issue of our understanding of the maintenance of the nighttime ionosphere and quantified it in terms of relative and absolute accuracies among the three models and the observations.

Recently, Sica et al. [1990] have drawn some attention to this problem in their inability to uniquely differentiate the relative influences of meridional winds and protonospheric fluxes on modeled values of $N_mF_2$ and $h_mF_2$. They pointed out that any number of reasonable but nonunique admixtures of meridional wind speeds and protonospheric fluxes could give rise to identical values of $N_mF_2$. Their database was only that of a network of $f_0F_2$ measurements from midlatitude ionosondes.

Pursuing the issue further in this investigation, correlations were carried out that involved measurements of $f_0F_2$ ($N_mF_2$), $h_mF_2$, and meridional winds. And these correlations were integrated over all stations as well as over stations restricted to equatorial-to-low and mid-to-high latitudes. The correlations also involved considerations of the physical assumptions within the FLIP and TIGCM models, wherein FLIP self-consistently solves the plasmaspheric flux inputs while the TIGCM invokes a relatively simple analytical formalism. The correlations showed a cause-effect chain in which the low model values of $f_0F_2$ were traceable to low model values for $h_mF_2$ (and associated increases in recombination losses) and ultimately to low model values of

nighttime meridional winds. The search for causality identified with nighttime protonospheric fluxes tended to suggest that they were of secondary concern. This was shown to be particularly true since the self-consistent solution of FLIP flux values was in general agreement with those in the post-midnight TIGCM flux formalism.

No consideration was given to electric field effects, where at least at mid-latitudes their influences are expected to be minimal. Nor was there any attention given to thermospheric densities and composition. It is possible that thermospheric densities and composition played a role in the discrepancies, but the relatively good agreement between the models and the data in the daytime suggested some confidence in the individual model treatments of thermospheric densities and composition.

Current efforts are focusing on the addition of convection electric field formalisms to the FLIP and an improved representation of upside fluxes into the TIGCM. The intent is to include local time and latitudinal controls. This will represent a convergence of predictive methodologies and a closer scrutiny of cause-effect relationships.

Acknowledgments. The authors wish to thank the institutions and agencies in each of the countries which participated in this investigation. The overall effort was supported by the Space Physics Division of the National Aeronautics and Space Administration under contract NASW-4755.

The Editor thanks U. Bilitza and another referee for their assistance in evaluating this paper.

References


Dickinson, R.E., E.C. Ridley, and R.G. Robie, A three-dimensional, time-
Pulzet, S., D. Evans, and M. Lester, UT and LT variations in the development of ionospheric disturbances over Russian longitudes, J. Geophys. Res., this issue.


M. Abdu, Institutional Nacional de Pesquisas Espaciais, 201 Ari. Dos Astronautas, 1758, Sao Jose Dos Campos, Caipa Postal 515, Sao Paulo, Brazil, e-mail: abdu@dae.inpe.br

P. Blanchard and E.P. Szuszczyewicz, Science Applications International Corporation, MS 2-3-1, 1710 Goodridge Drive, McLean, VA 22102, e-mail: blanchard@apso.sain.com; szusz@apso.sain.com

B. Emery, G. Lu, and R. Roble, National Center for Atmospheric Administration, 1850 Table Mesa Drive, Boulder, CO 80307, e-mail: emery@ncar.ucar.edu; roble@ncar.ucar.edu

D. Evans and J. Joselyn, National Oceanic and Atmospheric Administration, 325 Broadway, Boulder, CO 80303

R. Hanbaba, France Telecom/CNEN, SPI F. 22301 Lannion, France, e-mail: rudi.hanbaba@lannion.cnet.fr

K. Igarashi, Communications Research Laboratory, Ministry of Posts and Telecommunications, 2-1 Nukui-Kita-machi 4-Chome, Tokyo, 184 Japan, e-mail: igarashi@cri.jp

P. Jiao, China Research Institute of Radio Propagation, P.O. Box 138/97, Xinxiang, Henan 453003, China, e-mail: cri@crp.ac.cn

M. Lester, Department of Physics, University of Leicester, Leicester, LE1 7RH England, e-mail: mle@uk.ac.le.ac

K. Miller, Center for Atmospheric and Space Science, University of Utah, Logan, UT 84322-4405

S. Pulinets, IMIIRAN, 142092 Troitsk, Moscow Region, Russia, e-mail: 19.323.24.52@charley.izmiran.rssi.ru

B.M. Reddy, National Physical Laboratory, Dr. K.S. Krishnan Road, New Delhi, India 110012

P. Richards and D. Torr, University of Alabama, Huntsville, Al. 35899

P. Wilkinson, IPS Radio and Space Services, Level 4 15 Help Street, West Chatswood, * NSW, 2057 Australia, e-mail: phil@ips.gov.au

(Received September 6, 1995; revised April 26, 1996; accepted June 4, 1996.)