The ionosphere’s total electron content (TEC) is a parameter widely used in studies of the near-Earth plasma environment. The scientific use of TEC appeared early in the artificial satellite era, and among its many contributions were fundamental insights into how the ionosphere responds to geomagnetic storms. While many excellent reviews of solar-terrestrial disturbances exist in the literature, none have concentrated on the TEC parameter per se. With new TEC data sets increasingly available from the Global Positioning System (GPS), a comprehensive summary of pre-GPS storm studies is needed to set the base for progress in the GPS era. This review summarizes past case studies, describes statistical occurrence pattern, and identifies responsible mechanisms validated via modeling. It presents a new set of results of TEC disturbance patterns during 180 geomagnetic storms to describe seasonal and solar cycle effects. It concludes with a set of open questions that require additional study.

1. INTRODUCTION

Among the founding problems of solar-terrestrial physics was the origin of the terrestrial ionosphere. In his retrospective of ionospheric physics, Rishbeth [2001] credits Oliver Lodge with the photochemical explanation in 1902. Three decades later, Sydney Chapman [Chapman, 1931a, 1931b] described the theory of ionization for an atmospheric layer, including a solution for the grazing incidence problem. At nearly the same time he and Vincent Ferraro [Chapman and Ferraro, 1931] put forth a proposal for the source of geomagnetic storms, and then he along with Julian Bartels [Chapman and Bartels, 1940] proposed a comprehensive treatment of geomagnetism. Thus both the aurora and the causes of ionospheric perturbations during storms became linked to physical drivers in the solar wind. Decades later, when space physics became sufficiently prominent to warrant textbooks, the early ones dealt with the full scope of the solar-terrestrial chain [e.g., Hess and Mead, 1968; Rishbeth and Garriott, 1969; Papagiannis, 1972]. Today, as we approach the fiftieth anniversary of artificial satellites, students and researchers have multiple choices of specialized treatments to read on each of the geospace components (e.g., Kivelson and Russell [1995] on heliosphere-magnetosphere and Schunk and Nagy [2000] on the ionosphere). In this paper, specialization will be even more extreme in that a comprehensive discussion will be given of the storm time behavior of a single ionospheric parameter, total electron content (TEC), the integral with height of the ionospheric electron density profile. The reasons for this are threefold:

1. Scientifically, the ionosphere is a highly coupled neutral plasma system in which strong altitude effects upon photochemistry and plasma dynamics occur. Integrating over the entire ionospheric electron density profile to get TEC offers a convenient way to portray, to assess, and to understand the overall behavior of near-Earth thermal plasma. Such an integral has its maximum contribution from the $F_2$ layer, with approximately 2/3 of the TEC coming from regions above the altitude ($h_{\text{max}}$) of peak density ($N_{\text{max}}$). If substantial effects occur in TEC, simple vertical redistributions of $F$ layer plasma cannot be the main cause.

2. Observationally, TEC is a relatively simple parameter to measure that involves the use of radio diagnostics that are a fortiori transionospheric in nature. Thus the radio frequencies employed (VHF and UHF) generally do not suffer severe degradation during storms (as can happen with HF ionosondes), and thus the times of peak interest for solar-terrestrial physics are not lost to the very effects under investigation. This means that both regional and global networks can be used for TEC studies on an essentially continuous basis.

3. Finally, the use of TEC derived from the Global Positioning System (GPS) has revolutionized ionospheric physics in ways similar to those caused by the advent of ionosondes in the 1930s and incoherent scatter radars in the 1960s. Global and regional networks of GPS receivers are capable of providing TEC over large geographic areas (with the exception of large ocean areas where ground-based GPS receivers cannot easily be deployed). Nevertheless, it is now possible to get TEC from essentially everywhere on the globe and often in near-real time. The impact of such GPS-based research is still in its first decade, and thus use of the TEC parameter will surely advance the three foci of current-day ionospheric science (understanding ionospheric structuring, creating forecasting capabilities, and data-assimilative modeling). Moreover, from an applications perspective, TEC and its variabilities in space and time can have significant effects upon radio communications and navigation systems that have crucial “now-casting” as well as forecasting requirements. For all of the new GPS studies to contribute in substantive ways it is important that they build upon, and not repeat, past work so that valuable time and talent can be used efficiently.

The goal of this review is to provide the first comprehensive summary of the TEC storm phenomenon. This is possible because of the significant accomplishments achieved prior to the GPS era, both in defining TEC morphology patterns as well as in identifying all of the physical drivers of $F$ layer storm effects. That era ended in the early 1980s, and a new generation of GPS TEC experts has emerged in recent years. Thus, in utilizing the remarkable GPS capabilities now at hand, attention should be given to the areas most needed (e.g., modeling) and not to using wonderfully new diagnostics to rediscover known morphologies and the physics that drives them.

Section 2 provides historical context for TEC storm studies and discusses briefly the content of past review articles. The lack of adequate coverage of past TEC storm studies in those reviews is addressed in section 3 via summaries of published results at individual sites, using both low-Earth orbit and geostationary orbit satellites, and the results from networks of stations are presented in section 4. In section 5, results from a new study that draws upon a very large number of geomagnetic storms in the 1967–1976 period are used to provide a state-of-the-art summary of average TEC storm effects. Section 6 deals with several specialized topics related to TEC changes associated with low-level geomagnetic activity and their relationship to overall TEC variability. Section 7 offers a brief status of modeling ionospheric storms. Points of discussion and personal interpretation are given in section 8, and suggestions for future work are offered in section 9.

2. HISTORICAL CONTEXT

The ionosphere is not the same every day. All components of this highly coupled system (production, loss, and transport) change over multiple timescales, ranging from an impulsive solar flare or auroral intensification (approximately minutes) to solar cycle durations (~11 years). The identification of periods when known input functions change markedly can be used to study the system response to a
<table>
<thead>
<tr>
<th>Author</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matsushita [1959]</td>
<td>First comprehensive analysis of a large number of storms (100) at midlatitudes, giving positive and negative phases in storm time and local time, derived from a &quot;million or so&quot; ionosonde values, from ionosondes taken over a solar cycle (1939–1951), &quot;scaled by a team under the supervision of Miss B. Hardwick, B.A., to whom I am specially indebted.&quot; The conclusion was that all ionospheric disturbance variations and also the worldwide disturbance magnetic variations, are attributed to a single cause, the electrostatic fields. This of course meant enhanced chemical loss. The result is expected for the negative phase as introduced earlier by Searle (1959).</td>
</tr>
<tr>
<td>Somayajulu [1964]</td>
<td>It was the first to stress that TEC effects indicated large-scale changes to the overall ionosphere, i.e., that the previous ionosonde-derived results did not result from redistributions. The conclusion is that two modes of latitude progression occurred during storms: high to low latitudes simultaneous with equatorial to midlatitudes.</td>
</tr>
<tr>
<td>Seaton [1971]</td>
<td>Consistently, the positive phase dominated. A comprehensive set of storm patterns obtained for different longitudes. It is important for the suggestion that two modes of latitude progression occur during storms: high to low latitudes simultaneous with equatorial to midlatitudes.</td>
</tr>
</tbody>
</table>

**TABLE 1. Milestone Papers and Review Articles on Ionospheric Storms**

<table>
<thead>
<tr>
<th>Author</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matsushita [1959]</td>
<td>First comprehensive analysis of a large number of storms (100) at midlatitudes, giving positive and negative phases in storm time and local time, derived from a &quot;million or so&quot; ionosonde values, from ionosondes taken over a solar cycle (1939–1951), &quot;scaled by a team under the supervision of Miss B. Hardwick, B.A., to whom I am specially indebted.&quot; The conclusion was that all ionospheric disturbance variations and also the worldwide disturbance magnetic variations, are attributed to a single cause, the electrostatic fields. This of course meant enhanced chemical loss. The result is expected for the negative phase as introduced earlier by Searle (1959).</td>
</tr>
<tr>
<td>Somayajulu [1964]</td>
<td>It was the first to stress that TEC effects indicated large-scale changes to the overall ionosphere, i.e., that the previous ionosonde-derived results did not result from redistributions. The conclusion is that two modes of latitude progression occurred during storms: high to low latitudes simultaneous with equatorial to midlatitudes.</td>
</tr>
<tr>
<td>Seaton [1971]</td>
<td>Consistently, the positive phase dominated. A comprehensive set of storm patterns obtained for different longitudes. It is important for the suggestion that two modes of latitude progression occur during storms: high to low latitudes simultaneous with equatorial to midlatitudes.</td>
</tr>
</tbody>
</table>

The response of the ionosphere to geomagnetic activity, commonly called ionospheric storms, is an old topic in space physics. The field has been enriched by a remarkable set of comprehensive studies and review articles; some of the most important ones are listed in Table 1. It is a tribute to the earliest of such studies that the basic patterns were so well identified. Thus, while early researchers had defined global geomagnetic disturbances to have a storm commencement (SC) time, initial phase, main phase, and recovery phase, Sato [1957] and Matsushita [1959] found it necessary only to describe the ionospheric response as a positive phase (or increase type) followed by a negative phase (or decrease type). This remains true today, with current researchers occasionally suggesting new names for these phases but, as we shall see, not new mechanisms. All of the early studies were based upon ionosonde measurements, and thus while individual storms at individual stations may have suffered diagnostic degradations, the overall message from the statistical analyses of ionosonde data identified clear and unambiguous patterns. While not cast in the terminology of current-day space physics, nor in the graphical representation styles of contemporary space science, a serious reading of Matsushita [1959] reveals virtually all of the known characteristics of global F region storms: (1) The positive phase lasts longer and becomes more prominent with decreasing latitude, followed by a prolonged negative phase that is stronger with increasing latitude; (2) when the positive phase has a noticeable local time component, its maximum amplitude occurs near 1800 LT at subauroral latitudes, earlier at higher latitudes and later at lower latitudes; and (3) the influence of season is that positive storms are more pronounced in winter and negative storms are more pronounced in summer. Similarly, careful readings of Obayashi [1964] and Matuura [1972] offer comprehensive descriptions of mechanisms and their roles within the overall context of solar-terrestrial disturbances.

The earliest studies of ionospheric storms gave emphasis to the negative phase because it often lasted for several days and was thus the statistically dominant component. The choices of mechanisms proposed to account for such effects centered on dynamical effects due to electric fields, as well as changes in the composition (and therefore chemistry) of the thermosphere. For the positive phase, electrodynamics was also the suggested cause, and simulations of both types were conducted as far back as Sato [1957]. Obayashi [1964] proposed overall scenarios that involved enhanced chemistry and plasma transport and...
specifically how the new TEC observations then becoming available might be related to whistler observations of ionization above the ionosphere (i.e., what today we would call plasmasphere-ionosphere coupling). The only mechanism not really stressed in the early studies was the role of storm time thermospheric winds, though that would happen soon [Jones and Rishbeth, 1971].

[11] The next step was to test such ideas. This introduced the well-known argument in geophysics about the science yield that can flow from studies of individual events versus the insights gleaned from statistical patterns derived from a large number of events. Clearly, both have contributed greatly to our understanding. Naturally, an individual storm first had to be seen in order to identify perturbations from nonstorm conditions. Equally clear is the mantra that no two storms are precisely alike, thus fostering the seemingly endless pursuit of “case studies” that now dominates the field. Fortunately, as Matsushita [1959] showed, there are common elements to all ionospheric storms, and the statistical treatment of such perturbations has probably helped us learn more about the basic mechanisms of storms than have case studies. Yet it is the case studies that show how mechanisms that operate “on the average” can have extreme magnitudes, and thus theory and simulation work must be able to handle the largest spatial and temporal effects proposed for a given mechanism. In this review both methods will be used, case studies and statistics, to characterize how total electron content responds to geomagnetic storms.

[12] In conducting a reappraisal of the review papers presented in Table 1, the “in progress” reports by Obayashi [1964] and Matuura [1972] are not, today, as useful as the more recent review by Prölls [1995]. Curiously, this is not due only to the elapsed decades between them. Matuura’s [1972] approach, for example, is to summarize all possible mechanisms, with only a limited attempt at a critical evaluation of their correctness, importance, or impact upon the field. On the other hand Prölls [1995] did just that, and his book chapter (now 10 years old) remains the definitive summary of F layer storm effects available today. While the author’s individual preferences were given prime visibility, a noteworthy effort was made to make that clear to the reader, and thus his review was highly relevant at the time and remains so today. For example, Prölls does not dwell excessively on the negative phase of ionospheric storms since the physics and chemistry of enhanced loss rates were well understood at the time, owing mostly to Prölls’ own work with the ESRO-4 satellite and ground-based ionosondes [Prölls and von Zahn, 1974; Prölls and Najita, 1975; Prölls et al., 1975]. The positive phase, however, was left as a challenge in that several types of enhancements, each with its own mechanism, were introduced and a comprehensive synthesis was not achieved. This was due, in part, to Prölls’s view that the electrodynamics proposed by earlier workers was not as important a mechanism as thermospheric dynamics.

[13] Figure 1 shows how Prölls portrayed the five characteristic elements of an ionospheric storm, and they will be used as context in all that follows. The main conclusion will be that features 1 and 5 are not independent patterns but rather ones linked fundamentally by electrodynamics, as had been proposed earlier by several authors. Finally, Prölls’ [1995] review dealt overwhelmingly with the ionosphere’s maximum electron density ($N_{max}$) derived from ionosonde data; total electron content was mentioned only once, and TEC data sets were not included in any of the paper’s 25 figures. While the physical causes for $N_{max}$ and TEC storm time variations are the same, there are differences in each parameter’s ability to characterize the overall ionosphere during storms. Moreover, in the current era of GPS-dominated application studies, TEC is the parameter most often used in storm studies. Thus sections 3 and 4 will be devoted to the first comprehensive summary of observations and model results of TEC during ionospheric storms.
3. SINGLE-SITE TEC MEASUREMENTS AND THEIR USE IN STORM STUDIES

3.1. Observational Methods

[14] Early attempts to measure the ionosphere’s TEC involved transionospheric radar studies, popularly known as the “Moon reflection” technique [e.g., Evans, 1956, 1957; Bauer and Daniels, 1959]. With a double passage of the RF signal through the ionosphere, reliable TEC values could be obtained (for as long as the Moon was available as a target from a given radar site). With the advent of artificial satellites, above-the-ionosphere beacon transmissions received by a ground station could be used to obtain suborbit track TEC versus latitude (or longitude, depending on the satellite’s inclination) for the approximately 15-min tracking opportunity at a given site. These data provided excellent spatial coverage, but local time patterns were dependent on the slow precession of the LT orbital plane, thus mixing seasonal and diurnal effects. The use of geostationary satellite radio beacons solved the local time issue, and TEC could be monitored continuously at fixed sites within any longitude band capable of viewing the same geostationary satellite. Latitude chains of receivers could then provide both spatial and temporal patterns of TEC over vast regions. Moreover, when a new type of radio beacon system was installed on the ATS-6 satellite, one that allowed for simultaneous Faraday rotation measurements of the ionospheric TEC and group delay measurements of the ionospheric-plus-plasmaspheric electron content, studies of coupled plasma-sphere-ionosphere interchange patterns could be conducted.

[15] We are now in the era of GPS satellites, and the advantages (and disadvantages) of both the low-orbit and geostationary TEC methods are embodied in this new technique. Thus, while a GPS satellite is capable of being viewed for only a few hours per pass, several satellites (in different directions) can be observed at any given time (providing a new diagnostic of medium-scale horizontal structure). This review is concerned with large-scale structure, and thus while early radio beacon studies had shown TEC to exhibit mesoscale patterns from low-orbit satellites [e.g., da Rosa and Waldman, 1970], GPS offers a far better method. For diurnal coverage, GPS relies upon averaging over localized gradients in the multipoint observations from each site. This provides a good estimate of diurnal TEC patterns at each GPS receiver site, provided all the calibration “dark secrets” of satellite and receiver biases are handled satisfactorily. Far more reliable diurnal coverage is also becoming available from the new set of GPS transmitters on satellites in geostationary orbits (GEO). Expanded use of the GPS/GEO data from all GPS receiver sites offers the best possible set of TEC observations for ionospheric research, but this is still years away. The clear and major advance offered by current GPS data sets is, unquestionably, their large-scale spatial coverage: Vast networks of GPS receivers can provide TEC over specific countries and continents and, when merged (using appropriate interpolation techniques over oceanic regions), to full global maps with excellent time resolution.

[16] Each of the experimental methods mentioned above has a significant heritage in the published literature in which all of the observational techniques, including calibration, are discussed. Summaries also are given by Hargreaves [1992] and Hunsucker [1991]. For the purposes here, it will simply be assumed that all TEC data in the published literature have passed their validation criteria, and thus all reported “storm-induced perturbations” are reliable measures of ionospheric effects determined by careful experimenters.

[17] As a concluding point to this introduction one can speculate that if artificial satellites had not been placed in orbit about the Earth, the lunar reflection technique probably would have grown in use, and storm studies would have been conducted. As it turned out, early storm effects were indeed reported by Taylor [1961] using Moon echoes, with somewhat similar effects reported a few years later by Klobucharet al. [1964]. However, the launch of Sputnik I initiated an era of satellite radio beacon studies of the ionosphere that quickly became the dominant mode for TEC research. These can be divided into two relatively distinct categories, and so observations of TEC storm effects obtained from low-orbit (~1000 km) satellites are treated in section 3.2, and section 3.3 gives results from geostationary satellites.

3.2. TEC Behavior During Storms Observed Using Low-Orbiting Satellite Beacon Experiments

[18] Within a year of the placement of the first satellites into orbit, papers describing TEC and its variations during storms started to appear. In many of these papers the primary focus was simply on the new parameter itself, and thus variations reported during disturbed periods were often fragmentary. Yet, as we will see, a consistent picture emerged from these pioneering studies. Table 2 presents an attempt to summarize the central findings about TEC during storms, as taken from papers in the 1960–1967 time frame. As a perusal of Table 2 will show, there is an embedded message of storm phase and storm time. Table 2 ends in 1967 because in that year Hibberd and Ross [1967] published the first statistical treatment of TEC during storms using 30 TEC values taken from 18 disturbed periods. It is worthwhile to quote directly from their paper [Hibberd and Ross, 1967, p. 5333]:

The disturbance variation in TEC is almost always positive in the first 24 hours of the storm. The increase in content may be very large: The content may be as much as twice that of the undisturbed ionosphere. In the second 24 hours or so the content is decreased below that of the undisturbed ionosphere. The disturbance variation in content has disappeared by early in the third storm day, if not earlier. For most of the weaker or shorter duration magnetic disturbances the ionospheric variations do not persist for more than 24 hours and appear to be positive only.

[19] The Hibberd and Ross results encompass all of the findings listed in the diverse set of reports given in Table 2. It is a remarkably concise and important work, and one that points clearly to different mechanisms for the initial positive phase versus the subsequent negative phase. That weak storms only had a positive phase also offered a substantial clue about the character of the mechanisms driving the positive and negative phases. The answers to come, namely, that short-timescale dynamical mechanisms (electrodynamical and thermospheric) dominate the positive phase, while
longer-timescale composition changes the negative phase, required more complete observations of local time patterns versus latitudes and storm time.

3.3. TEC Behavior During Storms Observed Using Geostationary Satellite Radio Beacon Techniques

[20] Hibberd and Ross [1967, p. 5334] concluded their study of TEC storm time behavior by commenting on the limitations imposed by low-Earth-orbit satellites as follows: “The disturbance variations described above almost all refer to daytime and evening hours. They were not widely enough distributed to warrant examination for possible diurnal or seasonal disturbance effects.” This limitation to continuous TEC coverage was solved by the use of radio beacons onboard the new series of communications satellites being placed into synchronous orbit.


<table>
<thead>
<tr>
<th>Reference</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ross [1960]</td>
<td>inverse correlation between TEC and $\Sigma Kp$ in the summer but not in the winter</td>
</tr>
<tr>
<td>Garriott [1960]</td>
<td>TEC slightly reduced during two winter storms, nearly unchanged during three equinox storms; large horizontal gradients to the south</td>
</tr>
<tr>
<td>Hame and Stuart [1960]</td>
<td>TEC increased shortly after storm commencement (SC), decreases 12–24 hours later</td>
</tr>
<tr>
<td>Yeh and Swenson [1961]</td>
<td>TEC considerably decreased for measurements taken a few days after seven SC storms</td>
</tr>
<tr>
<td>Taylor [1961]</td>
<td>TEC depleted on the day following a SC</td>
</tr>
<tr>
<td>Evans and Taylor [1961]</td>
<td>small increases in winter nighttime TEC with $Kp$</td>
</tr>
<tr>
<td>Munro [1962]</td>
<td>TEC depressions on the day following a severe magnetic storm</td>
</tr>
<tr>
<td>De Mendonca [1962]</td>
<td>increase in north to south gradients during daytime positive phase and decrease in gradients for negative phase</td>
</tr>
<tr>
<td>Lawrence et al. [1963]</td>
<td>decreases in TEC occurring much more often than increases for $Kp &gt; 5$</td>
</tr>
<tr>
<td>Klobuchar et al. [1964]</td>
<td>large decreases in TEC on the day following a SC</td>
</tr>
<tr>
<td>Potts [1965]</td>
<td>no clear correlation between TEC and $\Sigma Kp$</td>
</tr>
<tr>
<td>Lyon [1965]</td>
<td>inverse correlation between TEC and $\Sigma Kp$ in summer but not in winter</td>
</tr>
<tr>
<td>Taylor [1965]</td>
<td>TEC $\sim 1/Kp$; decrease in sunrise slope; long-period oscillations during disturbed winter nights</td>
</tr>
<tr>
<td>Bertin et al. [1966]</td>
<td>TEC enhancements over an extended geographical area; latitude gradients greatly increased</td>
</tr>
<tr>
<td>Liszka [1966]</td>
<td>first studies of auroral and subauroral TEC patterns</td>
</tr>
<tr>
<td>Rai and Hook [1967]</td>
<td>auroral latitudes with increased gradients at all local times in winter; large nighttime enhancements attributed to energetic particle precipitation</td>
</tr>
<tr>
<td>Hibberd and Ross [1967]</td>
<td>first use of multiple TEC storm periods to define average storm time behavior showing a short positive phase followed by a longer negative phase</td>
</tr>
</tbody>
</table>

*Low-Earth orbit is $h \sim 1000$ km.


<table>
<thead>
<tr>
<th>Reference</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Klobuchar and Whitney [1966]</td>
<td>For 16 days of summer data the TEC dependence on $Kp$ was unclear; several large evening increases were observed as well as some daytime and evening depressions.</td>
</tr>
<tr>
<td>Nakata [1966]</td>
<td>First complete set of diurnal TEC curves during a magnetic storm was presented; TEC increased by 100% on the daytime period following SC and was reduced by 50% on the following day.</td>
</tr>
<tr>
<td>Titheridge and Andrews [1967]</td>
<td>Following the SC of the great June 1965 storm, TEC was enhanced by 20%; on the next day, TEC was reduced 35% with a gradual recovery over 5 days.</td>
</tr>
<tr>
<td>Goodman [1968]</td>
<td>Positive phases of two large storms were discussed; for the great storm of May 1967, the TEC enhancements of 100% nearly coincided with enhancements in the geomagnetic field.</td>
</tr>
<tr>
<td>Webb [1969]</td>
<td>The May 1967 storm was again studied; the 100% enhancements were followed by a 50% depletion on the following day.</td>
</tr>
<tr>
<td>Smith et al. [1968]</td>
<td>A TEC positive phase followed by a negative phase for two periods was discussed, only one of which was magnetically active.</td>
</tr>
<tr>
<td>Klobuchar et al. [1968]</td>
<td>Winter nighttime TEC enhancements during magnetically disturbed periods are often spatially and temporally confined phenomena.</td>
</tr>
<tr>
<td>Mendillo et al. [1969]</td>
<td>The June 1965 and February 1968 storm periods were examined; the TEC pattern of an initial positive phase followed by a longer negative phase occurs irrespective of season.</td>
</tr>
<tr>
<td>Taylor and Earnshaw [1969]</td>
<td>Analysis of June 1965 storm showed positive and negative phases; enhancement in topside TEC at dusk and transport effect at dawn were suggested.</td>
</tr>
<tr>
<td>Taylor and Earnshaw [1970]</td>
<td>During periods of weak geomagnetic activity ($Kp \leq 4$), summer midday TEC showed increases with $Kp$; transport by disturbance winds and/or electric fields was suggested.</td>
</tr>
<tr>
<td>Mendillo et al. [1970]</td>
<td>Examples of afternoon sector enhancements in TEC and total magnetic field (soon termed the dusk effect), and suggested linkage to changes in the plasmasphere were presented.</td>
</tr>
</tbody>
</table>
Within a year of doing so, observations and reports of storm effects and their diurnal signatures upon TEC appeared in the literature. Table 3 gives a summary for the period 1966–1970, pointing to the important role of case studies for TEC during the great storms of June 1965 and May 1967. Emerging from these studies was a sense that the diurnal pattern of the positive phase had a pronounced longitude pattern. At longitudes away from the North American east coast the positive phase was generally an all-day phenomenon [Nakata, 1966; Taylor and Earnshaw, 1969], while at the subauroral regions near 70°W longitude the afternoon/dusk sector was singled out for particularly dramatic positive phases in concert with increases in the local magnetic field [Goodman, 1968; Mendillo et al., 1969, 1970; Papagiannis et al., 1971]. Both the large amounts of plasma involved and the dusk meridian occurrence zone led to suggestions that the ionosphere-plasmasphere system was highly coupled during storms.

[21] The mid-1960s and early-1970s was the period when virtually all of the plasmasphere’s spatial and temporal patterns now being shown in exquisite data sets by the IMAGE satellite were being discovered and successfully explained, and their influence upon the ionosphere was being explored. While some false starts occurred that suggested “compression effects” that proved to be of minor importance, as well as overly optimistic “plasmaspheric dumping” scenarios [e.g., Papagiannis et al., 1971; Evans, 1973; Taylor, 1973; Park, 1970, 1973], the fact that the disturbed $F$ layer was being connected to plasmaspheric processes offered fundamental insights into a highly coupled system controlled by the degree of magnetospheric convection [Carpenter and Park, 1973; Rishbeth and Hanson, 1974; Moffett et al., 1975; Park, 1976].

[22] The first statistical study for the local time component of TEC storm morphology came from an analysis of 4-day periods following the SC of 28 geomagnetic storms (with $Ap > 30$) that occurred in the 26-month period November 1967 to December 1969 [Mendillo, 1971]. The TEC data came from Faraday rotation measurements from the Air Force Cambridge Research Laboratory (AFCRL) station at Sagamore Hill (Hamilton, Massachusetts) at 42.6°N, 70.8°W. The raypath from the ATS-3 satellite had a 420-km subionospheric point at 39°N, 71°W and a geomagnetic invariant latitude of ~53° corresponding to an $L$ value of ~3. The results obtained are reproduced in Figure 2 and point to several LT features that are common to most storms at a subauroral site. For example, the occurrence of a “dusk effect” in ionospheric storms is shown, meaning that the timing of the maximum of the TEC positive phase did not depend critically on the time of the storm’s SC but rather was a fundamental characteristic of the dusk sector at latitudes where plasmasphere-ionosphere coupling was most pronounced. Second, the contraction of the plasmasphere resulted in the $F$ layer trough moving to lower $L$ values after sunset, resulting in a dramatic termination of the dusk effect, followed by recurrent trough excursions to midlatitudes in the 0300–0600 LT sector for at least three nights of a storm.
period. Such dynamical causes of TEC negative phase effects are quite distinct from those due to enhanced loss via chemistry. Composition changes resulting from storm-induced modifications to thermospheric circulation are the dominant drivers of the daytime negative phase at midlatitudes. Surprisingly, the negative phase of the overall TEC perturbations was long lasting, i.e., persisting for more than 3 days after the positive phase (a topic discussed in more detail in section 5.4). These first statistical patterns demonstrated the benefit of continuous TEC data sets for storm studies (and their implications for full profile changes ($\Delta N_e(h)$) in contrast to peak density ($N_{\text{max}}$) variations alone) and provoked considerable modeling of the results displayed.

[23] It had long been appreciated that the negative phase of ionospheric storms results from enhanced ionospheric chemical loss driven by changes in the thermosphere. Thus the new results in Figure 2 were that such effects persist in statistical descriptions for at least 3 days after the geomagnetic activity subsided. In thermospheric studies at the time the Jacchia model widely in use linked upper atmospheric response to concurrent magnetic activity via the use of indices, and thus (statistically) the thermospheric storm was over when the magnetic storm was over. The advent of the Mass Spectrometer Incoherent Scatter (MSIS) model did not change this. It was the power of general circulation model (GCM) simulations that showed how the relaxation time of the thermosphere is not quick, and thus today we are not surprised by the longevity of some negative phase storm effects [e.g., see Burns et al., 2004; Fuller-Rowell et al., 1991, 1994].

[24] The cause of the “dusk effect” in Figure 2 was debated rather vigorously throughout the 1970s, and an excellent summary of the competing theories and interpretations appears in the reviews by Prölls [1995] and Buonsanto [1999]. As they point out, it was long appreciated that the competing theories of neutral winds and electrodynamics could each provide the required plasma uplifting to regions of reduced loss (thereby allowing solar production
Thus the crucial experimental need was to measure either a sudden onset of strong equatorward winds from auroral heating in the daytime sector (not possible in sunlight using the ground-based Fabry-Perot interferometer techniques then available) or the prompt appearance of electric fields of magnetospheric origin. The latter was possible using incoherent scatter radar methods, and the pivotal observation came decades ago at Millstone Hill on 14 May 1969. Figure 3 shows the TEC for this storm, one of the 28 storm periods used to obtain the statistical pattern at nearby Sagamore Hill in Figure 2. The storm began with a SC at 1929 UT, and $Ap$ reached 131. The TEC from Sagamore Hill showed the classic patterns of a dusk effect, its abrupt termination, followed by a prolonged negative phase. At nearby Millstone Hill a local ionosonde recorded effects at the ionospheric peak and the incoherent scatter radar (ISR) observed at low-elevation angles toward the west, thereby maximizing sensitivity to horizontal plasma motions [Evans, 1973]. The ionosonde’s $f_{o}F_2$ values and the ISR’s determinations of $h_{max}$ obtained on that day are shown in Figure 4a. Abrupt increases in the $F$ layer’s density and height followed the SC. In Figure 4b the ISR recorded a simultaneous onset of strong westward drifts within 30 min of the SC. Taken together, and shown in the context of auroral activity (Figure 4c), these offered convincing proof of a rapid penetration of magnetospheric convection electric fields to midlatitudes, causing the observed cross-L (upward and westward) drifts. In their review of ISR contributions to storm studies, Testud et al. [1975] pointed to these observations (together with results from other radars) as among the most important to come from the ISR measurement technique.

[25] A consequence of strong westward convection is the introduction of postsunset ionospheric plasma into the dusk area. Evans [1973] pointed out that this is equivalent to saying that the plasmapause contraction noted during storms...
has an ionospheric signature, i.e., the sharp termination of the dusk effect as seen in Figures 3 and 4. Figure 5a shows four examples of this pattern observed from Sagamore Hill, all showing TEC decay rates far in excess of those possible from recombination chemistry because of storm-enhanced or vibrationally excited N\textsubscript{2} concentrations. Two of these storms (February 1969 and March 1970) were also studied by Arendt [1971] and Kane [1975] using TEC from other sites. Figure 5b illustrates that this effect is a consistent occurrence feature at subauroral latitudes by showing same day results of a fifth example from two sites widely separated in longitude: Sagamore Hill (Massachusetts, eastern United States) versus Edmonton (Alberta, western Canada), both with \(L \sim 3\). Reprinted from Mendillo et al. [1974] with permission of Elsevier.

Figure 5. (a) Four examples of the abrupt termination of the “dusk effect” in TEC storm effects at a subauroral site (Sagamore Hill, \(L = 2.8\)) that are attributed to the appearance of the plasmapause. (b) Examples of the abrupt termination of the “dusk effect” in TEC storm effects on 15 April 1971, as observed at two widely separated subauroral sites in North America: Sagamore Hill (Massachusetts, eastern United States) versus Edmonton (Alberta, western Canada), both with \(L \sim 3\). Reprinted from Mendillo et al. [1974] with permission of Elsevier.
3.4. Early Simulation Studies of the TEC Positive Phase

[26] Modeling of the positive phase became a hot topic in the 1970s. Both neutral dynamics [e.g., Jones and Rishbeth, 1971; Davies, 1974a, 1974b] and electrodynamics [e.g., Tanaka and Hirao, 1973; Anderson, 1976] were used as drivers, each with some success (again see Prölls [1995] for a full discussion). For the TEC “dusk effect” of prime interest here, perhaps the most convincing results came from Tanaka and Hirao [1973] and Anderson [1976]. In Figure 6a the development of midlatitude electron density altitude profiles under the influence of a suddenly imposed eastward electric field is shown. As pointed out by Tanaka and Hirao [1973], the topside enhancements result in the TEC being “remarkably increased,” and the development of such effects agrees with storm patterns at Millstone Hill [Evans, 1970, 1973] (see Figure 4). The confinement of such patterns to the dusk meridian was then investigated by Anderson [1976], specifically examining the consequences of the zonal drifts of magnetospheric origin observed by Evans [1973] in conjunction with the role played by meridional winds. In Figure 6b the results are summarized for TEC, with the solid curve indicating diurnal pattern for quiet time winds and no external $E$ field, the dashed curve indicating disturbed (i.e., equatorward) meridional winds and no external $E$ field, and the dash-dot curve indicating both disturbed winds and a strong $E$ field of magnetospheric origin. The larger magnitude of the dusk effect and its rapid termination clearly follow from the strong westward plasma drifts; the causes are both stagnation of afternoon plasma during sunlit hours (making the wind-induced upward transport even more effective) and the transport of depleted post-sunset plasma into the afternoon sector.

[27] While the modeling results of Anderson [1976] were prompted by a need to demonstrate that the dusk effect was linked fundamentally to magnetospheric convection, there is a secondary aspect in Figure 6b, one not discussed by Anderson [1976], that also points to the penetration of the magnetospheric $E$ field along the dawn-dusk meridian. There is a corresponding dawn effect due to sunward magnetospheric convection: The positive slope of TEC versus LT during the postsunrise hours is reduced by the motion of low-density plasma from presunrise hours into the daytime ionosphere, a pattern captured schematically in Figure 1 (Prölls’ case 4), statistically in Figure 2, and in the case study shown in Figure 3.

4. MULTISITE TEC STUDIES DURING GEOMAGNETIC STORMS

4.1. AFCRL Chain of TEC Stations at 70°W

[28] Encouraged by the discovery of the dusk effect at their Sagamore Hill TEC-observing station, the Air Force Cambridge Research Laboratory mounted an effort to observe TEC from polar cap stations to the tropics. Figure 7a shows this set of stations with their subionospheric points, i.e., where the radio raypath from the ATS-3 geostationary satellite intersects the ionosphere at 420 km en route to each station. This network provided fundamental insights into the behavior of TEC during storms, an example of which is given in Figure 7b [Mendillo and Klobuchar, 1975].

[29] The storm of 17 December 1971 depicted in Figure 7b is of particular importance in the history of ionospheric disturbances. What used to be called “great storms” are now called “superstorms,” and this was surely one. The TEC “dusk effect” values recorded at midlatitudes were the highest ever seen, and the coherent pattern of the positive phase versus latitude and local time defined a fundamental characteristic to be seen in virtually all of the great storms that followed. In Figure 8 the TEC hourly values from the stations available for this storm are presented on a grid of latitude versus local time. The important result to flow from this format is noticing that the dusk...
effect is seen to be a subauroral feature linked to a much larger-scale TEC enhancement that occurs earlier at higher latitudes and later at lower latitudes. Moreover, the values of the TEC maxima are themselves latitude-dependent; that is, the positive phase is less pronounced as one moves from lower midlatitudes to auroral latitudes. Thus the pattern shown in Figure 8 is of a positive phase spanning all of North America, with a southeast to northwest alignment,

Figure 7. (a) Air Force Cambridge Research Laboratory (AFCRL) latitude chain of TEC-observing stations along the ~70°W meridian. (b) TEC data taken during the great storm of 17 December 1971. From Mendillo and Klobuchar [1975].

Figure 8. Spatial-temporal morphology of TEC storm effects using contours of TEC hourly values upon a grid of magnetic latitude and local time from the four high-latitude stations in the AFCRL chain shown in Figure 7, augmented by data from the Stanford University TEC station in Rosman, North Carolina. From Mendillo and Klobuchar [1975].
from the Caribbean to polar latitudes. Because of its electro-
dynamical origin this pattern is established early in a storm,
and as the north-south latitude chain of stations rotates past
the noon meridian, the onset of a TEC positive phase is
experienced first at high latitudes and then at lower ones.

4.2. Sagamore Hill (Massachusetts)—Arecibo
Observatory (Puerto Rico) Joint Storm Study

The growing acceptance that magnetospheric con-
vection patterns can be suddenly imposed upon the iono-
sphere prompted several studies of the latitude dependence
of TEC during storms. The crucial message from Figure
8 was that the TEC storm effects grew in magnitude and
occurred later in LT at lower midlatitudes, and thus empha-
sis should be placed on sites away from subauroral latitudes.
The study by Lanzerotti et al. [1975] did that and played a
pivotal role in linking effects observed at Arecibo to those
observed simultaneously at Sagamore Hill. Their results
from 12 storms from 1968 to 1970 appear in Figure 9. The
subauroral site showed the “classic” pattern (as in Figure 2)
of a dusk effect positive phase, followed by a sharp
transition to trough effects at 0300 LT and then transition
to the negative phase on the following days. At Arecibo the
daytime enhancement was followed by a persistent “ledge”
extending into the nighttime hours and no prominent
daytime negative phase.

Lanzerotti et al. [1975] pointed out that magneto-
spheric convection (described in the equatorial plane in \( L \)
shell and LT coordinates) maps to the ionosphere as three-
dimensional motions: vertical, zonal, and meridional. A
schematic very popular at the time from Chappell [1972]
was used by Lanzerotti et al. to relate the “peeling away” or
“detachment” of the plasmasphere in the afternoon quad-
rant to the northward, westward, and vertical motions
experienced in the ionosphere (Figure 10). The prompt
effect of enhanced convection is to move plasma from the
dusk sector to higher \( L \) values and earlier LTs. This causes
the dramatic westward and upward motions responsible for
the dusk effect and its abrupt termination, as seen in the ISR
observation by Evans [1973] at Millstone (Figure 4) and
modeled by Anderson [1976] in Figure 6. At higher
latitudes the pattern continues because of the poleward
component of convection, as shown in Figures 7 and 8.

The Arecibo pattern described by Lanzerotti et al.
[1975] is of particular importance because it showed for the
first time the presence of convection effects upon the TEC
at lower latitudes. At \( L < 2 \) the plasmasphere is rarely, if
ever, “peeled away,” and thus convection is entirely within

Figure 9. Results from the Arecibo–Sagamore Hill joint TEC storm study showing median storm patterns
(using the ratio of disturbed/quiet day) for the two stations over a 3-day storm period. The distribution of SC
times is arranged to ensure that the positive phase occurs on day 1, and results from less than half of the
storms are shown by the dashed curve. Note the termination of the “dusk effect” at Sagamore Hill occurs
when a lingering enhancement (“ledge”) occurs at Arecibo. From Lanzerotti et al. [1975].

13 of 47
a plasmasphere reduced in size. At these latitudes, there are no dramatic boundaries to regions of depletions (i.e., plasmapause and trough) to move about but certainly plenty of plasma. As will be discussed in section 4.5, the low-latitude ionosphere experiences dramatic storm time changes in TEC because of expansions and contractions of the equatorial ionization anomaly (EIA), and thus poleward horizontal motions of this always positive equatorward gradient can bring high TEC values into the lower midlatitudes. Thus, as suggested by Rishbeth and Hanson [1974, p. 706] in their criticism of convergence effects for the positive phase, “the major effect probably comes from the ‘advection’ term (in the continuity equation), resulting from the transport of relatively dense plasma into the volume monitored.” The true legacy, then, of the Lanzerotti et al. [1975] study was to portray that horizontal transport is of equal importance to vertical motions and that they are unified via the magnetospheric convection-plasmasphere-ionosphere connection. While a host of new names have been introduced by new researchers to the field (e.g., storm-enhanced densities for the positive phase/dusk effect in F layer storms and drainage plume for peeling/detached plasma tails in the disturbed plasmasphere), the fundamental mechanisms identified decades ago remain valid.

4.3. Case Studies of Severe Spatial Gradients at Midlatitudes

The undisturbed ionosphere depends on processes linked to temporal and spatial origins, such as photoionization being tied to the elevation angle of the Sun and low-latitude electric fields tied to the unique geometry of the dipole magnetic field. As a result the ionosphere exhibits strong spatial gradients across the sunrise terminator (LT effects) and at latitudes spanning the geomagnetic equator (the EIA). These affect conductivities, which, in turn, govern electrodynamics. Storms do not create new processes but rather intensify known processes to extents that are not fully known. The “case study” approach plays a crucial role in challenging proposed mechanisms to account for extreme events. There were several important studies of great storms (superstorms) that brought to light the remarkable spatial gradients that can occur in TEC during periods of extreme geomagnetic activity. Two of these dealt with midlatitudes gradients and will be treated here, while a third focused on longitude gradients and will be discussed in section 4.4. Storms causing severe modifications to the low-latitude EIA are treated in section 4.5.

The first major “global” study of storm effects in TEC used 20 geostationary satellite observing sites [Schödel et al., 1974]. This was done for the same great storm of 17 December 1971 shown in Figures 7 and 8, and the results had relevance to seasonal behavior since stations from both hemispheres participated in the study. Of the several strong morphology patterns shown, the clearest were that (1) a severe negative phase occurred only in the summer (southern) hemisphere and (2) the strong daytime latitude gradient for the positive phase observed at east coast North American longitudes (Figure 8) were matched at western North American longitudes. This demonstrated the persistence of characteristic effects for the positive phase versus latitude across several time zones (longitudes), suggesting that the electrodynamical mechanism that had been proposed for TEC storm effects near 70°W was not unique to that sector.

A follow-up study by Essex et al. [1981] using nearly the same network of stations, but now for the storm of 17 June 1972, advanced our understanding of seasonal, local time, and coupled longitude-latitude effects upon TEC during storms. The seasonal effects observed during the two storms were consistent; that is, positive phases dominated winter storms, while summer storms are dominated by negative phases, though often preceded by a brief positive phase.

One of the important uses of case study research is to seek temporal relationships between characteristics of the geomagnetic storm and the development of the ionospheric storm [Rishbeth, 1963]. A key accomplishment in the study...
of ionospheric storms came from the finding by Thomas and Venables [1966] that the F layer perturbations are not necessarily “turned on” at the time of the SC but rather at the time of main phase onset (MPO). The sharp decrease in the horizontal component of the global geomagnetic field (or equivalently in the \( Dst \) index) marks the intensification time of a geomagnetic storm. The penetration of electric fields and the onset of auroral input occur most strongly after MPO. Following MPO, \( \frac{dDst}{dt} \) decreases at its maximum rate to a minimum (most negative) value. While the temporal separations of these three “markers” of a geomagnetic storm’s main phase (MPO, \( \frac{dDst}{dt} \) max, \( Dst \) min) vary from storm to storm, statistically they all occur within a few hours [Mendillo, 1973], thus providing a UT interval of “intensification.” \[37\] Essex et al. [1981] examined the concept that the time of maximum geomagnetic activity plays an important role is how the ionospheric perturbations evolved globally (i.e., versus longitude). If TEC maxima occur at the same instant of storm time (e.g., during the most magnetically active UT period), they would be recorded progressively in local time at different longitudes. Such a case is shown in Figure 11 (top left) for the midlatitude stations in the North American subset of their network. The linear fit specifies the 2100 UT period when the peak geomagnetic activity occurred, as described by the \( Dst \) index. Noting that all of the stations shown are in the Northern (winter) Hemisphere, Essex et al. [1981] looked for, and found, the same effect during the June 1971 storm as sampled in the Southern (winter) Hemisphere (not shown in Figure 11). Thus they concluded that a winter storm seems to have the timing of its positive phase peak determined by the UT time of storm intensification (at least for the longitude sectors sampled). Using a similar network of six TEC stations spanning the United States, Hearn [1974, p. 133] had found that the average response during 21 storms in 1969 showed that the “westernmost stations tend to have earlier starting and longer lasting afternoon and evening sector enhancements.” \[38\] For summer storms (i.e., the June 1972 event for stations in the Northern Hemisphere and the December 1971 event for stations in the Southern Hemisphere), Essex et al. [1981] examined the longitude dependence of TEC maxima versus local time and did not find the same trend (there is no single UT time that accounts for the distribution shown in the bottom left plot of Figure 11). The best ordering parameter they found was latitude, as shown in Figure 11 (right), with TEC peaks occurring earlier in LT at high latitudes and progressively at later Lts equatorward. As will be discussed in section 5, seasonal effects upon ionospheric storms depend, in part, on day length and ambient O/N\(_2\) ratio conditions, both of which have strong seasonal dependencies.
4.4. Interplay Between Positive and Negative Phases

In the first use of TEC-observing stations distributed in longitude across North America, Klobuchar et al. [1971] showed how the negative phase can overwhelm the positive phase. Figure 12 shows a dramatic TEC longitude gradient for the great storm of 8 March 1970. That is, on the same day in the same hemisphere both strongly positive and strongly negative TEC daytime perturbations were observed at mid-latitudes separated by only a few time zones. In the eastern United States the positive phase was dominant (Prölls' type 1), while in the western United States it was the negative phase (Prölls' type 4). The implication was that the mechanism responsible for the negative phase (neutral atmosphere changes that enhance loss) can act quickly; if the storm starts early in the LT day and intensifies rapidly, then by the time that site rotates to the afternoon sector, the positive phase mechanism is ineffective, and hence only a negative phase storm is recorded.

Figure 12. The great storm of 8 March 1970 led to the first “joint study” of TEC storm effects using a network of TEC-observing stations at northern midlatitudes. In each plot the dashed curve is the control curve, and the solid curve shows the TEC pattern observed on the day of the storm in local zone time. The SC time is noted by an arrow in each plot. Note the severe daytime gradient spanning longitudes from Hawaii (157°W) to Arecibo in the Caribbean (67°W). From Klobuchar et al. [1971].
This scenario of auroral-heating-induced effects (i.e., storm time circulation changing compositions) causing a negative phase that dominates over the electrodynamical/wind effects that cause a positive phase is not limited to great storms only; it can be seen several times per year at a given site. This is illustrated in Figure 13 for the subauroral TEC at Sagamore Hill. The TEC patterns fall into three categories: For geomagnetic storms developing during the

Figure 13. Definitions of (a) regular positive phase (RPP), (b) delayed positive phase (DPP), and (c) no positive phase (NPP) storms in TEC at the subauroral site Sagamore Hill [after Mendillo, 1973]. The SC denotes storm commencement time, and main phase onset (MPO) is the time (when the sharp transition to Dst < 0 occurs) for the three geomagnetic storms [from Mendillo, 1973]. Figures 13a and 13b are prompt and delayed occurrences, respectively, of Prölls’ type 1, while Figure 13c is Prölls type 4 (see Figure 1). Reprinted with permission from Elsevier.
daytime hours (SC and MPO between sunrise and mid- afternoon), a regular positive phase (RPP) ionospheric storm occurs that dusk period; a magnetic storm commencing after dusk can have either a delayed positive phase (DPP) ionospheric storm, i.e., one occurring the following day, or a no positive phase (NPP) ionospheric development, i.e., one that goes directly to the negative phase.

[41] An early accomplishment of ionospheric storm studies came from Rishbeth [1963, p. 42], who noted that the local time variations of the disturbed geomagnetic field were different for the different types of ionospheric storms ("the auroral zone currents are more intense for storms which are ‘negative’ in midlatitudes than for those which are ‘positive’"). Goodman [1968], Webb [1969], Mendillo et al. [1969, 1970], and Papagiannis et al. [1971] noted that a strong TEC positive phase at dusk coincided with a strong enhancement in the geomagnetic field. However, the behavior of the magnetic field during “non-dusk-effect storms” had not been studied, and this offered a way to separate the two phenomena, consistent with Rishbeth’s concept of seeing if the geomagnetic field behavior could itself discriminate between types of ionospheric storms. This is not meant to imply that the same mechanism accounts for the geomagnetic and ionospheric perturbations but that it is rather a correlation of processes. This proved to be the case, as shown in Figure 14. The discriminator between the NPP and DPP cases (essentially between Pröll’s types 1 and 4 in Figure 1) is whether the geomagnetic field during the dusk to dawn hours exhibits the types of strong perturbations associated with an intense westward auroral electrojet. When this occurs, auroral heating (Joule and particle) is strong, and the resultant equatorward propagation of neutral atmosphere composition changes seeds the negative phase response for the following day. For cases when signatures of the auroral currents are minimal, then the opportunity for a positive phase mechanism can survive until the subsequent dusk period. This approach was validated by Mendillo [1973], with additional verification by Lanzerotti et al. [1975], showing it to be a reliable forecasting tool. Application of this approach to TEC perturbations during a much larger number of storms is described by Balan and Rao [1990].

Figure 14. Behavior of the disturbances in the total geomagnetic field strength during ionospheric TEC storms (a) with NPP and with a DPP in comparison to (b) field perturbations during RPP TEC storms. The total field at the Weston Observatory (Massachusetts) is dominated by the vertical (Z) component and hence is sensitive to auroral electrojet patterns to the north of its location near Sagamore Hill. Thus \((\Delta B)_{DPP} < (\Delta B)_{RPP} < (\Delta B)_{NPP}\) implying a gradation in Joule heating that determines the onset and severity of the \(\Delta TEC\) negative phase patterns in Figure 13. Reprinted from Mendillo [1973] with permission from Elsevier.
4.5. TEC Storm Effects at High and Low Latitudes

4.5.1. Auroral Zone Patterns

The TEC storm studies summarized in sections 4.1–4.4 dealt overwhelmingly with midlatitudes. Yet important coupling processes from auroral latitudes were described in few cases (e.g., Figures 8–12). Prölls [1995] gave an excellent summary of high-latitude $F$ layer storm effects, and no attempt is made here to improve on his discussion.

Figure 15. Examples of TEC storm effects at three high-latitude stations during the great storm of 16–17 December 1971 that had SCs at 1906 UT on the 16th and 1418 UT on the 17th. (a) Kiruna (Sweden) results showing that the midnight periods (2300 UT) on 16 and 17 December had TEC values approaching half the daytime values and considerably above the poststorm values for nighttime. In Greenland from (b) Thule and (c) Narssarsuaq, significant enhancements in TEC during comparable LT periods. For the latter stations the dotted curves give monthly median behavior. In each case the latitudes and longitudes given are for the geographic locations of the station (subscript 0) and of its 420-km subionospheric measurement point (subscript i). Using magnetic coordinates, the subionospheric latitude for Kiruna’s observations is at $58^\circ$, with Thule’s at $72^\circ$ and Narssarsuaq’s at $63^\circ$. Note that the precipitation-induced enhancements in the Greenland stations are larger at Narssarsuaq than at Thule because of the equatorward shift in the auroral oval (with an enlarged polar cap) during large storms. Reprinted from Schödel et al. [1974] with permission from Elsevier.
However, as pointed out in section 2, TEC data were not shown in his work, and so, for completeness, the relatively sparse results for TEC at auroral and polar cap latitudes can be summarized here.

Figure 16. TEC average storm effects at Goose Bay (Labrador), an $L = 4$ station. Results come from 67 storms during 1971–1975 with $Ap > 30$ or $Kp > 5$ during the storm period: (a) all storms, (b) 21 summer storms, and (c) 22 winter storms. Asterisks denote average values coming from less than half of the 67 storms. From Buonsanto et al. [1979].
storms, as shown in Figures 16b and 16c. In all cases the auroral enhancements reappear (reduced in magnitude) on subsequent nights, showing the persistence of precipitation-induced F layer enhancements throughout the storm period. Similar patterns were described for the L = 4 region in the Alaskan sector by Hargreaves and Hunsecker [1981].

[46] The daytime negative phase in TEC and the postsunset motions of the trough, described at Sagamore Hill (L = 3) in Figure 2, are clearly linked to similar effects at L = 4 in Figure 16. Buonsanto et al. [1979, p. 20] describe these as due to “the penetration of an enhanced magnetospheric convection electric field” in accord with the explanations offered by Lanzerotti et al. [1975] for the L = 1.3–3 domain.

### 4.5.2. Patterns at Equatorial and Low Latitudes

[47] The storm time behavior of TEC in the tropical ionosphere has been studied extensively, perhaps even more so than at midlatitudes. The ionospheric communities in India, Japan, and Taiwan did extraordinary amounts of work, with publications appearing in both less widely circulated national journals [e.g., Somayajulu et al., 1971; Chandra et al., 1973; Jain et al., 1978] as well as in those of western Europe and the United States [Rastogi, 1962; Rastogi and Rajaram, 1965; Rajaram et al., 1971; Raghavarao and Sivaraman, 1973; Huang et al., 1974, 1987; Deshpande et al., 1977; Rastogi et al., 1979]. Prior to the widespread availability of TEC data, earlier works dealing with the low-latitude response to geomagnetic activity were conducted using individual ionosonde stations and chains of stations during storms [e.g., Matsushita, 1963; Rajaram and Rastogi, 1969, 1970]. In the Matuura [1972] review of F region morphologies, two figures with TEC data were used to document how both positive and negative perturbations can be related to changes in the “fountain effect” associated with the ambient equatorial ionization anomaly. While no TEC observations are presented or discussed, Prölls [1995] provided an excellent summary of the mechanisms involved. Here both statistical patterns and case studies are used to describe our unified understanding of low-latitude ionospheric storms.

[48] The TEC structure at equatorial and low latitudes is ordered by the geomagnetic field. Throughout this region the TEC pattern responds to three fundamental storm time processes, each linked ultimately to magnetospheric input: (1) prompt effects, (2) delayed effects, and (3) composition changes.

#### 4.5.2.1. Prompt Effects

[49] Following the onset of a storm, the TEC initial positive phase develops at midlatitudes, in part, because of the appearance of magnetospheric electric fields that move low-latitude plasma to higher altitudes/latitudes. At the geomagnetic equator this enhanced “fountain effect” means a reduction in TEC, with corresponding TEC enhancements developing in each hemisphere poleward of their normal EIA crest locations. From an equatorial aeronomy perspective this effect is called penetration electrodynamics, and in the postsunset sector it also leads to the onset of equatorial spread F (ESF), as recently reviewed by Martinis et al. [2005]. The TEC “ledge effect” (Figure 9) found by Lanzerotti et al. [1975] is the consequence of this poleward transport of an enhanced EIA pattern early in a storm. It is the low-latitude signature of the same convection/advection that leads to the patterns shown in Figure 8 at middle and higher latitudes.

[50] Using ionosonde data, Kotadia [1965] showed that the transequatorial F layer responds very differently over small latitude spans, depending on the times when the storm begins and intensifies. An example of this is shown in Figure 17a. The solid curves in the top two panels of Figure 17a give \( f_o F_2 \) storm time changes at stations near the magnetic equator (Kodiakanal) and the EIA crest (Ahmedabad); the dashed curve shows changes in the local magnetic field (representative of the \( Dst \) index). There is a clear opposite phase relationship between changes in \( f_o F_2 \) at the two sites early in the storm, as well as during the main and recovery phases of the geomagnetic storm. The transition begins at main phase onset in \( Dst \) and is established as the \( dDst/dt \) reaches its largest negative value.

[51] A convincing demonstration that the EIA enhancements occur early in a storm was offered by Raghavarao and Sivaraman [1973] using Alouette topside sounder data (recall that two thirds of TEC is above \( h_{max} \)); these are shown in Figure 17b. On two consecutive days, each with a SC (solid arrow) and observation times shortly after MPO (dashed arrow), there was a dramatic growth of the EIA in magnitude and poleward extent (dashed versus solid curves).

#### 4.5.2.2. Delayed Effects

[52] Electric field penetration scenarios are short-lived, and thus past statistical studies [e.g., Kotadia, 1965; Somayajulu, 1963; Kaner, 1975] tended to emphasize the fact that the dominant effect at low latitudes is that the EIA is reduced in scale during storms. For TEC this was shown clearly by Basu and Das Gupta [1968] using low-orbit TEC data (see Figure 17c). The suppression of the fountain effect is then the longer-lived effect. Such delayed electrodynamics arises from the winds generated by auroral heating that reach low latitudes, i.e., the “disturbance dynamo” mechanism [Blanc and Richmond, 1980] that reduces the eastward electric fields at equatorial latitudes.

#### 4.5.2.3. Composition Changes

[53] The final (and second longer-lived) mechanism of magnetospheric origin is the composition changes that also result from auroral heating. As described nicely by Prölls [1995], the high-latitude heating that instigates neutral circulation from high to low latitudes has the secondary result of enriching the N2 at midlatitudes and the O densities at low latitudes. This adds to the prolonged nature of the TEC positive phase effect in storms at sites near the equator, accounting for the nonelectrodynamical component when the fountain effect mechanism is not dominant. The interplay between the electrodynamical “fountain” processes 1 and 2 and the auroral-heating-induced process 3 is most pronounced at the interface between low and middle latitudes (\( L \sim 1.3 \)), a topic discussed in greater detail in section 5.5.

#### 4.5.3. Superstorms at Low Latitudes

[54] Perhaps the most dramatic example of equatorial electrodynamics of magnetospheric origin occurred during the highly disturbed period of 13–14 March 1989 when,
after a daily index of $Ap = 246$ on 13 March, early on the 14th the activity levels included an hourly index of $ap = 400$, $Dst = -598$ (with the hourly index $SymH = -707$), and $Kp = 9$. Huang and Cheng [1991] used TEC data recorded in Lumping (Taiwan), obtained by both geostationary satellite and low-orbiting satellite signals, to show very dramatic changes in the magnitudes and positions of the EIA (crests unusually poleward early in the storm and then nonexistent the following day). In Brazil, approximately 12 hours away in local time, the poleward excursion of the EIA crest, driven by ~200 m/s vertical plasma drifts, was equally dramatic [Batista et al., 1991]. Using DMSP passes in the same longitude sector, Greenspan et al. [1991] showed that the base of the equatorial ionosphere, driven by this extraordinary fountain effect, was above the altitude of the satellite (840 km)! Modeling by Batista et al. [1991] and

Figure 17. Examples of different ionospheric storm effects at low latitudes: equatorial ionization anomaly (EIA) enhancement and suppression that depend upon storm time and latitude. (a) (first plot) The $f_{o}F_2$ variations versus storm time at Kodaikanal (EIA trough site, solid curve), averaged from 59 SC storms in 1956–1957, shown in comparison to changes in the horizontal component of the geomagnetic field (dashed curve, right axis); (second plot) results from Ahmedabad (EIA crest site, solid curve), averaged from 65 SC storms in 1953–1957). Early in a storm (but after MPO), there is a “mirror image” transition from an enhanced to a suppressed EIA in the two top plots, in contrast to the standard midlatitude results during the same years shown in the bottom two plots. Reprinted from Kotadia [1965] with permission from Elsevier. (b) A case study using topside sounder data during the storm of 21–22 September 1963 ($Ap = 44–126, Kp = 7^o–7^o$). SC times are indicated by solid line arrows, with dashed line arrows indicating the times of the observations shown below. These data show dramatic enhancements (dashed lines) in electron densities at 450 km and 500 km with EIA crests shifted toward higher latitudes in both hemispheres in contrast to quiet time patterns (solid curves) for midday. Reprinted from Raghavarao and Sivaraman [1973] with permission from Elsevier. (c) Case studies of the suppression of the EIA for disturbed days (solid curves) versus quiet days (dashed curves) during conditions of winter, fall, and summer using latitude profiles of TEC taken at the University of Calcutta [from Basu and Das Gupta, 1968]. These “late time” EIA contraction effects are in marked contrast to the enhanced EIA that occurs early in a storm as in Figure 17b.
Rasmussen and Greenspan [1993] succeeded in showing that the electric field penetration process could account for the TEC and \( F \) layer effects observed in this extreme case study event. Such studies offer convincing examples of how storm effects can truly be exaggerated cases of ambient processes, as shown for the low-latitude EIA in TEC via modeling during quiet times by Anderson and Klobuchar [1983].

[55] The TEC modulations associated with such expansion/contraction effects of the fountain effect are now well understood because of the patterns of disturbance vertical drift versus storm time and local time presented by Fejer [2002]. In Figure 18, annotations to this model are made to unify how the EIA changes can be related to storm time vertical drifts. The top plot shows a "parameterized storm" using the geomagnetic index \( AE \), with seven elapsed times noted in a disturbance that lasts for 13 hours. At each phase of the "storm" (t0 to t6) the vertical drifts from the Jicamarca radar were used to create average departures.
versus local time. These form the seven plots below. The absolute values of the averaged departures are small, so it is the sense of their direction that is the important message. Thus, at times t0 and t1 (early in the disturbance) the perturbation in vertical drift is negative in the postmidnight (PM) hours (0200–0600 LT) and positive in the postsunrise (PS) period (1800–2200 LT). The color coding and annotations added on the right indicate their consequences: a suppression of the EIA/fountain effect for PM times and an enhancement of the EIA/fountain effect for PS times. Following this penetration phase, the disturbance dynamo causes opposite trends. The key point to note is that PS growth in the fountain effect is limited to early storm times and that the longer-lived effect of a storm is fountain effect suppression, thereby accounting for all of the observed morphology patterns shown in Figure 17.

5. A NEW STATISTICAL STUDY OF TEC DURING STORMS

[56] In sections 2, 3, and 4 a broad summary was given of the history of TEC storm studies from the years when Moon echo radars were first used to the era when low-orbit and geostationary satellite radio beacon experiments were widespread. Considerable insights into the dominant occurrence patterns and the responsible physical processes were achieved. Both case studies during “great storms” and statistical studies for groups of storms resulted in consistent effects and interpretations. Prior to the Buonsanto et al. [1979] study that used 67 storms during 5 years for the L = 4 ionosphere, the largest sample sizes used for the statistical treatment of TEC during storms had only 28 storms during 2 years [Mendillo, 1971; Mendillo et al., 1972] at L = 3 and only 12 storms (within 2 years) at two stations (L = 3 and L = 1.3) [Lanzerotti et al., 1975]. To put the statistical study approach on a more firm basis over the full L = 2–5 range, a retrospective analysis was recently conducted [Mendillo and Klobuchar, 2006] from the largest set of TEC storm observations ever assembled in a systematic way. This refers to the AFCRL chain of stations near 70°W that operated in the 1960–1970s, shown schematically in Figure 7. A comprehensive summary of TEC data from Sagamore Hill was presented in an “atlas” of 75 storm scenarios during the years 1967–1972 [Mendillo and Klobuchar, 1974]. A second compendium (with an additional 109 storms from 1973 to 1976) appeared via two technical reports [Mendillo, 1978] that were widely circulated among the satellite radio beacon community. While partial sets of results from these data sets were presented for Sagamore Hill by Hargreaves and Bagenal [1977] and for Goose Bay by Buonsanto et al. [1979] and in a discussion of prediction capabilities of latitude patterns by Davies [1981], the full scope of the work was never published in the literature. Moreover, the AFCRL TEC data sets from stations at other longitudes were left unused. Mendillo and Klobuchar [2006] provided only a brief description of how these data could be cast into average storm patterns (with disturbances as percentage changes from monthly median conditions) and how they might be portrayed using more modern visualization methods than shown, for example, in Figure 8 where absolute TEC values were used. A more comprehensive description of this vast TEC archive is now presented.

5.1. Selection Criteria and Analysis Method

[57] Table 4 gives a summary of the stations used and the number of storms available for each of those stations. The events selected had a specified sudden or gradual storm commencement (SC) time in UT and the daily geomagnetic index Ap > 30 units (or Kp > 5) on the day of the SC or the following day. The numbers of storms per year meeting these requirements are shown in Figure 19a.

[58] The analysis protocol relied on the well-understood patterns of regular positive phase, no positive phase, or delayed positive phase storm effects at Sagamore Hill (as defined in Figure 13), and thus all of the other station results were analyzed in accordance with such phasing at Sagamore Hill. The scheme is illustrated in Figure 19b using the 75 events in the AFCRL atlas. The approach is to arrange the ΔTEC (percent) time series into daily/LT averaging bins that will preserve features (such as the “dusk effect”) in the

<table>
<thead>
<tr>
<th>Station (Subion Point)</th>
<th>Time Period</th>
<th>Number of Storms</th>
</tr>
</thead>
<tbody>
<tr>
<td>North America</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Narssarsuaq, Greenland (53°N, 52°W)</td>
<td>Apr 1971 to Dec 1975</td>
<td>70</td>
</tr>
<tr>
<td>Sagamore Hill, Massachusetts (39°N, 71°W)</td>
<td>Dec 1967 to Dec 1975</td>
<td>109</td>
</tr>
<tr>
<td>Kennedy Space Flight Center, Florida (26°N, 80°W)</td>
<td>Nov 1973 to Sep 1976</td>
<td>70</td>
</tr>
<tr>
<td>Other Sites</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Athens, Greece (34°N, 28°E)</td>
<td>Oct 1972 to Dec 1976</td>
<td>63</td>
</tr>
<tr>
<td>Osan, Korea (33°N, 137°E)</td>
<td>Jan 1974 to Jun 1976</td>
<td>31</td>
</tr>
</tbody>
</table>
storm day averages. That is, if the 13 storms with delayed positive phases had been placed to start in the late afternoon of day 1, their dusk effects would have occurred on day 2 and hence would be averaged with the negative phase perturbations from the regular positive phase and no positive phase storms. Similarly, if the 7 NPP storms that commenced early in the local time day had been placed in day 1, their negative phase perturbations would have been averaged with the positive phase effects. To avoid this, they are placed to start directly on day 2. By identifying these 20 storm patterns in advance, and adjusting their placement within the averaging bins, the resultant perturbation values have magnitudes more representative of those found during individual case studies. The actual UT onset time of an SC thus has no absolute priority in determining the resultant LT disturbance patterns. The choice of a monthly median pattern for the control curve was made because it best addressed the question of how geomagnetic activity contributes to the even distribution of days above and below the median for a given month (a topic treated in more detail in section 6).
5.2. Statistical Pattern: Results for All Storms

Figure 20 contains the final summary of TEC storm effects from auroral to lower midlatitudes near 70°W. This pattern comes from the largest number of TEC storm events analyzed in a consistent manner, and yet it does not suffer from the types of excessive averaging that yield small magnitudes that question the utility of applying statistical methods to highly variable geophysical phenomena. By the superposition of LT-dependent features (via the approach shown in Figure 19) and by following the pattern daily in local time, consistent characteristic patterns are obtained with magnitudes representative of those found during individual storms. There are several important responses in TEC storm effects captured in Figure 20:

1. Because of the placement of the delayed positive phase storms in averaging bins prior to the start of the statistical day 1, the TEC storm-induced effects can be seen during the early morning hours of the first day. Such TEC enhancements are produced by the ionization of the thermosphere by energetic particle precipitation from the magnetosphere (the Prölls’ type 3 in Figure 1). Note that the latitude pattern of enhancement is peaked after midnight at the highest latitudes ($L = 4–5$) but is peaked at midnight at $L = 3$. An auroral oval with its lowest latitude at midnight has clearly made its equatorward excursion prominent in TEC effects.

2. Following sunrise on day 1, a positive phase in TEC occurs with enhancements larger and later in local time from auroral to midlatitudes. Near $L = 2$, however, there is a clear separation of the positive phase into two components. The first (postnoon) enhancement appears as a broad daytime effect (Prölls’ type 2 in Figure 1), consistent with a wind-induced uplifting from high to lower latitudes. The second enhancement (postset) is the consequence of magnetospheric convection, as described by Lanzerotti et al. [1975]. This advection and uplifting effect occurs in concert with the dramatic termination of the “dusk effect” at subauroral latitudes where the trough/plasmapause boundary suddenly appears at midlatitudes (Prölls’ types 1 and 5 in Figure 1). At equatorial and low latitudes the simultaneous effect is the enhancement of the equatorial ionization anomaly early in a storm (Figures 17a and 17b).

3. The final day 1 feature is the precipitation-induced poleward wall of the trough seen as TEC enhancements to either side of midnight at $L = 4–5$ and at midnight at $L = 3$ (Prölls type 3 in Figure 1). The intrusion of the trough to lower latitudes is a prominent nighttime effect.

4. On day 2 of the storm period a daytime negative phase in TEC occurs from middle- to high latitudes because of changes in thermospheric composition (Prölls’ type 4 in Figure 1). The trough’s excursion to midlatitudes is most dramatic. At lower midlatitudes, there is again evidence of two distinct components of the positive phase, one at midday and one postsunset (to be discussed in more detail in section 5.5).

5. Days 3 and 4 of the storm period show a daytime recovery of the negative phase to prestorm conditions at
most latitudes, while signatures of the nighttime trough and the double enhancements near $L = 2$ persist to the very end of the storm period.

### 5.3. Statistical Patterns: Seasonal Effects

The effect of season has long been noted in ionospheric storm studies. It was pointed out in early works using ionosonde data in the 1950s and confirmed in the TEC case studies of Schödel et al. [1974] and Essex et al. [1981] described in section 4.3. The data sets used to obtain Figure 20 are sufficient in number that they can be divided into three 4-month groups (winter, equinox, and summer storms) to yield a latitude versus LT daily analyses for seasonal effects. These are shown in Figure 21 over the 4-day storm period. The strong message to come from Figure 21 is that all characteristic signatures of TEC storm effects

---

![Seasonal effects upon TEC storm patterns using the same format as in Figure 20, with (top) winter, (center) equinox, and (bottom) summer. From Mendillo and Klobuchar [2006]. See text.](image-url)
occur in all seasons. Their modulation in magnitude depends on ambient conditions. That is, virtually all past studies of seasonal effects show that the effect of a storm is to enhance the photochemical mechanism that is dominant for that season [Duncan, 1969]. Thus, for winter seasons, when the so-called seasonal anomaly results in higher electron densities than found in summer, a geomagnetic storm tends to accentuate the ambient production-dominant situation: The positive phase of an ionospheric storm is far greater for storms that occur during winter months. In local summer, when the ambient ionospheric loss rates are at their annual maxima, a geomagnetic storm further increases the loss rates, and the negative phases of ionospheric storms are extreme. Thus vertical motions in winter are more effective in making “dusk effect” signatures when the O/N$_2$ ratio is high, while storm-induced reductions of the already low O/N$_2$ ratios in summer cause extra strong negative phases.

[66] The consistency of the day 1 positive phase showing two distinct drivers is also evident in the seasonal plots: The midday enhancement due to winds is always there (and on multiple days near $L = 2$). In contrast, the afternoon enhancement due to convection of plasma from lower altitudes/latitudes follows day length and thus the longevity of solar production (TEC peak enhancements occur later in summer, earlier in winter). Finally, for the auroral-particle-induced enhancements the effect of season is to accentuate the percentage enhancements in winter when the control curve (monthly median nighttime TEC) values are lowest (i.e., there is no seasonal anomaly for the nighttime ionosphere).

5.4. Characteristic Patterns: Influence of the Solar Cycle

[67] The TEC database for Sagamore Hill is far greater than for any of the other stations in the AFCRL network, and so the influence of solar cycle can only be addressed at that site. This may be done by noting that three independent studies of TEC storms at Sagamore Hill relate to different parts of solar cycle 20. The 28 storm periods analyzed by Mendillo [1971] occurred under conditions with (1) $\langle F_{10.7} \rangle = 151$, while the 75 events analyzed by Hargreaves and Bagenal [1977] had (2) $\langle F_{10.7} \rangle = 139$, and finally, the 109 storms described by Mendillo and Klobuchar [2006] had (3) $\langle F_{10.7} \rangle = 99$. The effects of solar flux and of season upon the “dusk effect” from all of those studies are summarized in Table 5. The equinox period has been divided into separate spring and fall seasons and shows no appreciable differences. The point to note is that the afternoon enhancements, when portrayed as a percentage change, are inversely related to solar flux (simply because of the monthly median control curves having smaller magnitudes). As shown in Figure 21, the timing of the dusk effect peak is latest during storms in summer months.

[68] Additional solar cycle effects occur with the negative phase. Figure 22 gives the average storm patterns for Sagamore Hill for the three solar flux conditions. While the average magnitude on the first local time day (SD1) of the negative phase remains near $-10\%$ in all plots, there are small changes in its local time structure; that is, its amplitude decreases with solar flux. Such patterns were termed the “DC/AC” components of the negative phase by Rodger et al. [1989], results derived from a solar cycle of ionosonde data in Antarctica. Rodger et al. gave a comprehensive discussion of neutral wind and composition changes as the possible drivers of such patterns.

[69] The dominant message about the negative phase offered by Figure 22 concerns its duration: (1) At solar maximum the negative phase persists for more than 4 days; (2) for solar moderate conditions it is essentially over by the fourth day; (3) under solar minimum conditions the negative phase is only a 2-day statistical event. It is unclear how such effects would arise. Storm time thermospheric circulation and composition perturbation effects upon the $F$ layer are usually considered to be appreciable for only a day or two, but their time constants may well depend on the ambient levels of solar flux. The refilling times for the depleted flux tubes of the plasmasphere from the ionosphere below (to be treated in section 6.2) may also depend on the state of the prestorm $F$ layer (and therefore upon solar cycle phase). Such considerations add a long-term timescale to the concept that ionospheric storms are exaggerated cases of the same processes that drive normal daily and seasonal variability patterns. They are the Sun’s “active experiments” in solar-terrestrial physics.

5.5. Longitude Effects at Locations with $L < 2$

[70] The important result to come from the joint Arecibo-Sagamore Hill storm study [Lanzerotti et al., 1975] was that the positive phase at Arecibo persisted in local time well beyond the dusk effect seen at Sagamore Hill (see Figure 9).
As shown in Figures 20 and 21, this postsunset “ledge” is the second of two persistent enhancements that occur on day 1 of a storm near \( L = 2 \) during all seasons. Its appearance in the statistical patterns shown in Figures 20 and 21 arises from the TEC observations at the Kennedy Space Flight Center (KSFC) site in the AFCRL latitude chain. Prior studies of ionospheric storms had long considered two types of positive phases but usually as alternates, i.e., not as simultaneous events in the same latitude-longitude sector. For example, Evans [1973] had described two types of positive phases that can occur at Millstone Hill (“all day” versus “dusk effect”); he speculated that they might operate “independently or in concert” but had no observational evidence to offer for simultaneity. Tanaka [1979] also discussed the dual pattern of the daytime positive phase (called simply “positive” or “evening peak” types of storms). He pointed out that the dusk effect could be superimposed upon a broader diurnal enhancement, as shown in his data from Fort Monmouth (New Jersey). The two effects were attributed to storm time equatorward winds and to magnetospheric electric fields that affect midlatitude regions, respectively, particularly at American longitudes. The results in Figures 20 and 21, portrayed nearly 3 decades later via statistically derived average storm patterns, serve as an excellent confirmation of Tanaka’s early insights from case study events.

To determine if this dual-positive-phase feature is unique to the \( 70^\circ \text{W} \) longitude sector or if it is a global pattern, TEC data from the AFCRL stations in Athens (Greece) and Osan (Korea) were examined. The Athens data included 63 geomagnetic storm periods during the years 1972–1976, and the Osan data included 31 storms during 1974–1976. Their subionospheric points have \( L \) values of 1.4 and 1.3, respectively, only slightly different from that for the KSFC observations (\( L = 1.8 \)). Figure 23 gives the average storm patterns for all three of these sites. Figure 23a gives the KSFC pattern that was the input for that latitude in Figure 20, clearly showing the origin of the double maxima on days 1, 2, and 3. At European longitudes the TEC pattern from Athens (Figure 23b) shows the separation of the two peaks less distinctly, but it is clearly there. Far to the east, the pattern from Osan (Figure 23c) shows the effect prominently, with the nighttime maxima as persistent signatures throughout the storm period. Similar patterns derived using TEC from Hawaii during 27 storms in 1967–1968 can be seen in the results of Basu et al. [1975] and particularly so for storms in winter months.

The cause(s) of the double enhancements at \( L < 2 \) has (have) yet to be explained in full. The pattern on day 1 of a storm, as described above, can be attributed to the effects of equatorward winds (first peak) and horizontal convection (second peak). It is tempting to suggest that these two mechanisms simply act at reduced magnitudes on subsequent days. However, the electrodynamics of the disturbed equatorial and low-latitude ionosphere do not support this view. For example, as shown in the Fejer and Scherliess [1997] model in Figure 18, the period of enhanced EIA is confined to the dusk sector early in the storm. This penetration of

![Figure 22. TEC average storm patterns at Sagamore Hill during different phases of the solar cycle. Note the decrease in duration of the negative phase from conditions characterized by \( F_{10.7} \) values of (a) 151, (b) 139, and (c) 99 from top to bottom.](image-url)
magnetospheric electric fields is replaced by disturbance dynamo electric fields in later postsunset periods and thus there is a decrease in the EIA and, by implication, a reduction of poleward horizontal convection effects. The clue may come from the auroral enhancements in TEC that do persist on subsequent days. With reduced, but persistent, auroral activity on days 2, 3, and 4, small evening sector surges in equatorward winds would lift the $F$ layer to regions of reduced loss, creating small positive phases in comparison to the control curve (monthly median) pattern. The effect would maximize at $L < 2$ because of the inclination effects of the geomagnetic field (i.e., the $\sin I \cos I$ term causes meridional wind uplifting effects to peak at $I = 45^\circ$). More work on understanding ionospheric storms in this latitude region is clearly required.

6. SPECIAL TOPICS IN GEOMAGNETIC ACTIVITY EFFECTS UPON TOTAL ELECTRON CONTENT

[73] Part of the reason for the systematic study of ionospheric storms is to understand the role played by all levels of geomagnetic activity in determining the spread about a monthly mean or median pattern of TEC. The TEC database described in section 5 spanned nearly 11 years that included 180 geomagnetic storms of sufficient strength to warrant study. Thus, on average, there are one to two good storms per month, and as has been shown, they cause 3–5 days of perturbed conditions per event, thus affecting about one third of the days in a month. The truly pronounced variations are only on days 1 and 2 of a storm, and so most of the days in a month are not exceptionally disturbed because of storms. Yet a monthly mean pattern of $F$ layer behavior always shows a standard deviation ($\sigma$) of about 20–30%, and thus these variations are not due entirely to isolated geomagnetic storms (see Johanson et al. [1978] for TEC and Forbes et al. [2000] and Rishbeth and Mendillo [2001] for $N_{\text{max}}$).

[74] Figure 24 illustrates this point by showing a full year of TEC data (1971) from Sagamore Hill (Massachusetts), separated by month, with each day’s diurnal curve overplotted on a TEC versus LT format. Several of the highest curves are due to storms described earlier (e.g., the extremely large “dusk effect” in TEC in December (Figures 7 and 8) and the rapid decrease in TEC in April (Figure 5b) attributed to a plasmapause crossing). Smaller dusk effects are seen in several other months. The lowest curves in most months are due to the negative phase in TEC storms (i.e., the day after a dusk effect, e.g., in March, April, May, June, July, September, and October). Removing these obvious storm days still leaves a significant monthly variability ($\sim 15–20\%$) that needs explanation. In this section several types of TEC perturbations associated with geomagnetic disturbances, but not storms per se, are described. They contribute to the variability but do not account for all of it. That is, during periods of geomagnetic quiet, the ionosphere is still variable because of coupling from the lower atmo-
Figure 24. A year of total electron content data separated by month to portray overall ionospheric variability. The observations were made using Faraday rotation measurements of the 137-MHz plane polarized beacon on the geostationary satellite ATS-3 from the AFCRL station at Sagamore Hill, Hamilton (Massachusetts), 42.6°N, 70.8°W. The 420-km subionospheric point is at 38.7°N, 70.7°W, having an invariant magnetic latitude of 53° corresponding to L = 3. The traditional MKS units of TEC (10^16 el/m^2) are given on the left axes, and their correspondingly imposed time delay (in nanoseconds) for a GPS-type (L1 = 1.575 GHz) frequency is given on the right. A value of 20 TEC units results in ~10 ns of delay, causing a range error of ~3 m.
If auroral physics is the longest-studied aspect of magnetosphere-ionosphere coupling, the second most studied area must be plasmasphere-ionosphere coupling. Nighttime increases in TEC at latitudes well equatorward of the auroral zone [Titheridge, 1968; Klobuchar et al., 1968] were always associated with some form of plasma transport since photoionization or auroral precipitation were, by definition, not available mechanisms at such latitudes. Called protonosphere dumping or draining of magnetospheric flux tubes, these TEC changes were characterized by smooth nighttime increases spanning a few hours (in marked contrast to the jagged, irregular enhancements sketched as Pröls’ type 3 in Figure 1). Examples of this effect are shown in Figure 25, with TEC data from Sagamore Hill taken in two directions. Note that the quiet time behavior is a nearly constant TEC versus LT (right center plot) and that small geomagnetic disturbances cause large increases in TEC (left axes) and \( N_{\text{max}} \) (right axes). Identifying these effects with magnetospheric substorms became a major focus of activity and particularly so in the comprehensive studies by Park [1970, 1971, 1973], Park and Meng [1973] and Park and Banks [1974] and in the modeling study of Moffett et al. [1975].

Magnetospheric convection effects became the accepted cause of nighttime increases in TEC, providing the station was at a latitude within the plasmasphere and hence had full flux tubes above it. For a site poleward of the plasmapause the magnetospheric electric field would simply redistribute the existing ionospheric plasma, leaving TEC unchanged but \( N_{\text{max}} \) enhanced. This behavior for the trough region was examined in detail by Buonsanto [1976], and examples are given in Figure 26. Figures 26a–26e show a smoothly decaying TEC but with a nighttime increase at time \( P_2 \) in \( N_{\text{max}} \) at \( L = 4 \) during a substorm. In Figures 26f–
26h, simultaneous data at $L = 3$ show a nighttime increase in both parameters. The bottomside true-height $N_e(h)$ profiles from the St. Johns ionograms reveal the nature of the downward motions that result in increased peak density but, with compensating topside distortions, no appreciable change in TEC.

### 6.2. Plasmasphere Electron Content Variations During Storms

Ionospheric TEC obtained from Faraday rotation observations of VHF geostationary satellite radio signals are weighted strongly by the geomagnetic field strength, and thus most of the rotation occurs below $2000\ km$ [Titheridge, 1972]. This makes the Faraday electron content ($N_F$) a bona fide ionospheric measurement. If a geostationary satellite has multiple frequencies, these can be used to determine both the Faraday content (ionospheric) and total raypath content (ionosphere plus plasmasphere) total content ($N_T$), as described early in the satellite era by Garriott et al. [1970] and Almeida et al. [1970]. In 1974 a new and sophisticated radio beacon experiment was launched onboard the ATS-6 satellite making possible both Faraday and differential group delay observations of electron contents out to the geostationary distance of $6.65\ R_E$ [Davies et al., 1975, 1976; Hartmann et al., 1977]. Since the group delay method is not sensitive to magnetic field weighting, it gives the total integral of electron densities ($N_T$) from the satellite to the ground station. The difference between this total content and the Faraday content is the electron content of the plasmasphere ($N_P = N_T - N_F$). The ATS-6 mission had a global impact on such studies because the satellite was repositioned in longitude during its lifetime from $90^\circ W$ to $30^\circ E$, enabling observations and results from groups first in North and South America and then in the United Kingdom, continental Europe, and India.

The plasmaspheric contribution ($N_P$) to the total raypath content ($N_T$) is typically a few TEC units ($2 \times 10^{16}\ e^-/m^2$). Changes of this size are not highly significant perturbations to the ionospheric content except at night (as discussed in section 6.1, above). The plasmasphere is depleted during a storm, however, and the ATS-6 differential content methods offered a wonderfully simple way to monitor its refilling from the ionosphere below. Studies of coupling between the ionosphere and plasmasphere using ATS-6 thus provided results that were far more extensive than the earlier whistler investigations that were confined mostly to Amer-
can longitudes [e.g., \textit{Park}, 1970]. This ATS-6 legacy remains as the definitive database, and the one that generated the most comprehensive set of observations and modeling for plasmasphere-ionosphere coupling studies to date.

\cite{79} Examples of ATS-6 results during storms are given by \textit{Soicher} [1976], \textit{Degenhardt et al.} [1977], \textit{Kersley et al.} [1978], and \textit{Sethia et al.} [1978]. Scenarios showing the refilling process are given by \textit{Poletti-Liuzzi et al.} [1977] and \textit{Webb and Lanzerotti} [1977], with model results given by \textit{Poulter et al.} [1981]. A key result to come from stations in different longitude sectors that were able to observe ATS-6 during its extended mission was the study of how the offset between the geographic and geomagnetic equator determined the family of flux tubes sampled in their respective regions. This had consequences for the time it took to refill those flux tubes and therefore for the duration of the plasmaspheric storm effect. Thus, as shown by \textit{Kersley and Klobuchar} [1980], midlatitude stations in the United States (Sagamore Hill and Boulder) had longer refill scenarios than a station in Wales (see Figure 27a). As shown in Figure 27b, this is due to the minimum \(L\) values sampled along the raypath being smaller in Aberystwyth \(L = 2.3\) than at Sagamore Hill \(L = 2.3\). Thus smaller volumes can be refilled more quickly by ionospheres at lower midlatitudes where solar production is greater. When the underlying ionosphere exhibits seasonal effects in its TEC storm, namely, that summer storms have more long-lived negative phases, then the replenishment time is affected as well; that is, it takes several days longer to replenish the local plasmasphere flux tube after summer storms than after winter storms [\textit{Kersley and Klobuchar}, 1980].

\cite{80} Finally, the ATS-6 radio beacon experiments were also able to address the TEC nighttime winter enhancement effect.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure27.png}
\caption{(a) Average changes in plasmaspheric content (\(\Delta N_p\)) as observed by three midlatitude stations (Aberystwyth, Wales; Boulder, Colorado; and Sagamore Hill/Hamilton, Massachusetts) over 16-day periods following the onset of geomagnetic storms. Within each day shown, local time runs from 0000 to 2400 hours. The progression in replenishment times (fastest at Aberystwyth, longest at Sagamore Hill) depends on the minimum flux tube volumes sampled by each station, a feature that depends on longitude (see text). (b) Comparison of geomagnetic field line configurations for the ATS-6 raypaths from 6.65 \(R_E\) on the geographic equator to stations at Sagamore Hill (Hamilton, Massachusetts) and Aberystwyth (Wales, U. K.). Because of the tilted dipole effect versus longitude, the minimum \(L\) value sampled from Sagamore Hill is \(L = 2.3\) and is \(L = 1.7\) from Aberystwyth. Reprinted from \textit{Kersley and Klobuchar} [1980] with permission from Elsevier.}
\end{figure}
issue and its relationship to plasmaspheric dynamics (discussed in section 6.1). Davies et al. [1979] used ATS-6 observations from Bozeman (Montana), Boulder (Colorado), and Dallas (Texas), sites well within the plasmasphere at $L = 3.1$, $2.3$, and $1.9$, respectively, to show that the plasmaspheric flux tubes above those sites did not experience measurable depletions during TEC nighttime increases. They showed that the westward $E$ fields of magnetospheric origin described by Park and Meng [1973] resulted in compensating lateral (cross $L$) motions of ionization from higher-to-lower $L$ shells, thus allowing for a replenishment of flux tubes above the ionosphere that provided the plasma flow to regions below [Moffett et al., 1975]. This was an important result showing that both vertical (downward) and horizontal (equatorward) plasma drifts of magnetospheric origin had an ionospheric signature at night, complimenting the earlier findings of afternoon sector increases in TEC, also due to inner-magnetosphere-ionosphere coupling (upward and poleward motions), as treated at length in sections 3 and 4.

6.3. Solar Wind Sector Boundary Crossings and Day-to-Day Variability of TEC

[81] Geomagnetic storms and substorms are initiated by episodic changes in solar wind parameters (principally $B_z$). Two broad characterizations of such episodes are the as-yet unpredictable coronal mass ejection versus the repeatable and thus somewhat more predictable pattern of solar wind sector boundary (SB) crossings. It was thus natural to begin with the latter, and the search for SB-crossing effects upon the ionosphere was started by Low et al. [1975] and Lyon and Bhatnagar [1979] using TEC data and by D’Angelo [1980] and Zevakina and Lavrova [1980] with $f_F2$ observations. Using the AFCRL network of TEC stations shown in Figure 7, Mendillo and Schatten [1983] examined such effects during the period 1973–1975, declining years of a solar cycle when
sector structures are most evident. A superposed epoch analysis centered on the day of a SB crossing yielded results shown in Figure 28. Figure 28 (top) shows that, on average, the final day the Earth spends within a sector (day = \(C_0\)) is generally the one of lowest geomagnetic activity. Wilcox [1968] had shown that such low Ap values were associated with low-density solar wind streams that precede a sector boundary. The transition to disturbed conditions on the SB-crossing day then triggers an ionospheric positive phase response. Thus the SB-crossing effect is simply a restatement of the geomagnetic activity effect but with a lower level ionospheric variation than found after sudden commencement geomagnetic storms.

6.4. Geomagnetic Activity Versus Geomagnetic Quiet

The results in section 6.3 support the concept that all forms of geomagnetic activity should have an effect upon the ionosphere’s TEC. Sections 3, 4, and 5 documented the results for SC storms, section 6.1 gave patterns during substorms, and section 6.3 described SB-crossing patterns. To some extent they all have similar characteristic signatures. If geomagnetic activity were the only source of ionospheric variability, then the opposite of this statement should also be true; that is, periods of geomagnetic quiet should have TEC variations that are in marked contrast to the disturbance patterns. Perhaps the clearest way to demonstrate this is to take advantage of the fact the strong seasonal effects occur in ionospheric storms, as demonstrated in Figure 21. Using the well-understood TEC behavior at Sagamore Hill (\(L = 3\)), Mendillo and Schatten [1983] examined the five most disturbed (DD) and the five most quiet (QQ) days of 37 summer months (May–August) and 40 winter months (November–February). Their average departures from monthly mean conditions were computed, as well as the standard deviations (\(\sigma\)) of the mean QQ and DD patterns. The results are shown in Figure 29. During geomagnetic storms in summer months the dominant (i.e., longer-lived) pattern is the daytime negative phase. Thus the DD pattern is one of depletions, with a variability of 22%. The QQ pattern is its mirror image, with enhancements in the daytime and reduced variability (15%).

Figure 29 (bottom) gives results for winter months. The characteristic signatures of geomagnetic activity dominate the DD patterns: The daytime positive phase (with increased variability), the postsunset enhancements due to auroral precipitation, and the postmidnight trough appearance dominate the DD patterns. The mean results for the QQ days show features that are an unambiguous mirror image of the DD pattern. This is a good demonstration of how geomagnetic activity, in general, appears to dominate a monthly mean pattern: The five most disturbed and the five most quiet days of each month define the extremes (and spread) of day-to-day variability for that month (see Figure 24). Yet it is significant that the quiet time patterns do not have vanishingly small variabilities. The ionosphere is not constant when geomagnetic activity is low. This fact points to the role of coupling from below, perhaps the least well understood solar-terrestrial process in all of space physics. Upward coupling is a highly significant mechanism contributing to the overall day-to-day variability of the ionosphere [Forbes et al., 2000; Rishbeth and Mendillo, 2001; Mendillo et al., 2002]; increased progress in this area is a critical need.

7. MODELING TEC STORM EFFECTS

Models proposed for ionospheric storms span at least 50 years of research. It is considerably beyond the scope of this review to summarize all of the modeling approaches and results obtained over the past decades. In keeping with the focus on the TEC parameter we first point out that most models of F region disturbances concentrate on the peak density (\(N_{max}\)) and thus are fully relevant to TEC. These...
include, for example, many studies that have employed well-known first-principle models (e.g., thermosphere ionosphere electrodynamical general circulation model (TIEGCM), coupled thermosphere ionosphere plasmasphere (CTIP), time-dependent ionospheric model (TDIM), field line interhemispheric plasma (FLIP), and global theoretical ionospheric model (GTIM), all described briefly by Anderson et al. [1998]). They have been used for simulations of individual storms, as well as for in-depth studies of specific disturbance mechanisms. The latter deal with storm effects in general and are therefore more appropriate for comparisons to statistical patterns, as presented above. An early example using the National Center for Atmospheric Research GCM approach was given by Roble et al. [1979], with more comprehensive and recent examples given by Fuller-Rowell et al. [1996, 2000], Field et al. [1998], and Schunk et al. [2003].

Empirical models for storm effects have also been developed, e.g., as correction methods for climatological models of the $N_e(h)$ profile, such as the International Reference Ionosphere; a good example of this approach is described by Araujo-Pradere et al. [2004].

For the TEC parameter per se, storm models of individual features, such as the “dusk effect” in ionospheric storms, were developed by Tanaka and Hirao [1973] and Anderson [1976], as shown in Figure 6. Decades later, once GPS data became widely available to document storm scenarios [e.g., Ho et al., 1996, 1998], the GCM approach began to be used for TEC case study events [e.g., Lu et al., 1998]. The Lu et al. [1998] study pointed out how storm simulation work requires use of many submodels for input, even in self-consistent GCM studies. These include, for example, the assimilated mapping of ionospheric electrodynamics (AMIE) patterns for time-dependent electrodynamics, inputs of upward propagating semidiurnal tides at the lower boundary, and an accurate solar irradiance model for the period in question. Altitude coverage in GCMs is usually limited to a set of pressure levels (with heights $< \sim 700$ km) that do not span the full range of altitudes relevant to TEC. Yet some aspects of TEC storm effects are confined to regions near $h_{max}$ and below, and thus the results from Lu et al. [1998] portray quite well such features as the poleward wall of the trough and the $O/N_2$ changes responsible for the negative phase. For the complete simulation of TEC storm changes a plasmasphere model coupled to a thermosphere-ionosphere global model is required. This began with the coupled Sheffield University plasmasphere-ionosphere model (SUPIM) [Moffett and Murphy, 1973] that was used in pioneering studies of plasmasphere-ionosphere interactions in the 1970s [e.g., Moffett et al., 1975] and during the ATS-6 era by Poulet et al. [1981]. This was subsequently merged with the GCM at University College London [Fuller-Rowell and Rees, 1980, 1983] to form a coupled thermosphere-ionosphere model (CTIM) [Fuller-Rowell et al., 1987].

The most promising developments in TEC storm modeling involve new data assimilation methods (see Schunk et al. [2004] and papers in the special issue “Selected Papers From Ionospheric Effects Symposium” in Radio Science, 39(1), 2004). Examples of global assimilative ionospheric modeling using GPS TEC include studies of specific processes (e.g., electrodynamics by Pi et al. [2003]). Applications to storm time enhancements of the EIA are shown by Hajj et al. [2004], and how O/N$_2$ data can be used to study the negative phase in TEC are shown by Fuller-Rowell et al. [2004].

An important aspect of TEC modeling that has emerged in the last decade is the longevity of storm perturbations [Lu et al., 1998; Fuller-Rowell et al., 2004] and their dependence upon solar cycle conditions in the thermosphere [Burns et al., 2004]. The negative phase of storms was identified early in solar-terrestrial science and, as described in section 2, was considered the dominant effect to study. Yet, even with its composition change mechanism long known and modeled [Fuller-Rowell et al., 1991], the negative phase may well be the most difficult to model successfully over a full multiday storm period. As shown in Figure 22, there is a clear prolongation of the negative phase during solar maximum years, and Figure 21 shows that well-known seasonal effects in the negative phase [Fuller-Rowell et al., 1996] also have different longevities. Modeling thermospheric winds and composition effects, from an abrupt onset that is followed by either a slow or rapid recovery to prestorm levels, involves the correct specification of auroral input (particles and Joule heating) both in space and time, over a 4- to 5-day storm period. No model, to my knowledge, has been applied to such long-lived scenarios.

8. DISCUSSION

In the sections above a comprehensive summary was attempted for a remarkably complicated geophysical system: the response of the ionosphere’s total electron content to all forms of geomagnetic activity. First using case studies and then comprehensive statistical treatments, the spatial-temporal morphology patterns were established for storm effects upon TEC, including seasonal and solar cycle effects. Modeling studies, based on the earliest known patterns for the positive and negative phases of ionospheric storms, were shown to be successful in identifying the three dominant causes of storm effects: thermospheric composition changes, neutral wind perturbations, and the appearance of electric fields of magnetospheric origin.

When conducting a retrospective of a field that spans over 50 years of active research, it is inevitable that key papers are missed and that personal views creep into the narrative. In the material that follows the latter is certainly true. As I see it, space physics has changed during the past 6 decades from a discovery-mode field to one that now focuses mainly on applications-mode activities, i.e., predictions of space weather. New observing systems for the ionosphere (GPS) and the plasmasphere (IMAGE) have introduced data sets far more comprehensive than those used during what might be called the “ionospheric storm study era” of the 1960s, 1970s, and 1980s described in this review. Yet the improvements in “imaging” ionospheric storm patterns via maps of TEC, or in “seeing” plasmaspheric structuring in two dimensions, have not yet added significantly to our understanding of the basic physical processes identified in early studies. I am confident that this will still happen, and, indeed, the origin of this review article
was an attempt to push that agenda ahead faster than we are currently seeing.

The current era of GPS-based studies has introduced a new generation of researchers to TEC storm effects and, in the process, reintroduced the concept of direct magnetospheric-ionospheric coupling via penetration electric fields as a dominant player in storm scenarios. It was an interesting experience for this author to explore how ionospheric storm science evolved during the past century. As we shall see, it was perhaps more of an exercise in the sociology of science than in actual space research. The field started with essentially uniform agreement that electric fields were the source of ionospheric departures from Chapman theory. The legacy of explaining the Appleton anomaly via electrodynamics generated a following that readily applied such ideas to ionospheric storms. The early giants in the field, e.g., D. F. Martyn, T. Sato, and others, simply used $E$ fields to explain everything! Once artificial satellites were launched, thermospheric drag effects were seen to change dramatically during geomagnetic storms. This brought to our attention the importance of changes in total neutral density, implying thermospheric expansion caused by storm time heating. Satellite-borne neutral mass spectrometers then showed composition changes to be the cause of the negative phase of ionospheric storms, and it became a readily acceptable replacement for the electrodynamical explanation, as indeed it should have. This left the positive phase as an electrodynamical effect. Then, in the late 1960s, neutral winds were recognized as a fundamental driver of $F$ layer dynamics, and, to some extent, the pendulum swung to winds as the mechanism of choice for many non-Chapman-like morphologies, including ionospheric storms.

As is often the case in science, the correct solution of “multiple mechanisms” for a phenomenon (in this case winds and $E$ fields for positive ionospheric storms) did, in fact, appear, and, for the most part, that explanation (as put forward by J. V. Evans and by T. Tanaka in the 1970s) summarized very nicely the overall status of the field. Indeed, if a student or new researcher wanted a relevant-to-today summary of ionospheric storm effects and mechanisms from that era, and as put forward by a single author so as to offer a unified approach, the work of Tanaka and Hirao [1973] and Tanaka [1979, 1981] would be the reading assignment of choice. Tanaka [1981] discusses the electrodynamics of the positive phase, its large-scale linkage to equatorial fountain effect changes, and they, in turn, to storm time influences upon equatorial spread $F$. The growing trend of only finding references that are available via electronic publishing databases fosters poor library research, and thus my suggestion to read Tanaka’s work is offered as a way to reacquaint the GPS community with leadership ionospheric storm research right there in the library.

By the mid-1970s a subset of the ionospheric community arrived at what might be called “an acceptable level of understanding” for ionospheric storms, and many of the ionospheric researchers most associated with storms simply moved on to other careers (e.g., John Evans to COMSAT) or to other areas of study (e.g., I to so-called active experiments). In looking back, the interesting development that followed was that the pendulum that had swung from $E$ fields only to winds plus $E$ fields did not stop but swung to actually exclude electric fields as a mechanism available for use in ionospheric storm studies. How or why this should

Figure 30. Histogram of the number of publications dealing with ionospheric storms published per year (using 421 references taken from Prölss [1995]). The reduction in published results from 1975 to 1990 is attributed to two factors: the achievement of a “reasonable level of understanding” and the appearance of studies suggesting that the penetration of magnetospheric $E$ fields would not last long enough to cause significant ionospheric perturbations. The examples shown are Jaggi and Wolf [1973], Southwood and Wolf [1978], Siscoe [1982], and Senior and Blanc [1984]. See text.
happen leads to Figure 30, where the comprehensive review by Prölls [1995] is used to examine publication patterns.

[94] There are many interesting aspects of the statistics in Figure 30 (and some probably involve funding and/or personnel changes at agencies, changes in radio propagation applications needs, or simply the fact that no new diagnostic systems were available to study storms). Nevertheless, the growth pattern to peak activity that led to the “acceptable level of understanding” for ionospheric storms by the mid-1970s is clearly evident, as well as the secondary onset of new studies in the 1990s. Checking the reference list in the review by Buonsanto [1999], one that emphasized post-1996 progress, revealed that the publications in 1997–1998 continued the high activity trend of 25–30 per year.

[95] I have added to the publication pattern in Figure 30 the dates of a few papers that introduced the concept that electric fields of magnetospheric origin are essentially excluded from the ionosphere during storms, that is, that their penetration is limited to such short durations (tens of minutes to an hour or so) that they cannot be of any consequence. Thus I suggest that these important papers on magnetosphere-ionosphere coupling by Jaggi and Wolf [1973], Southwood and Wolf [1978], Siscoe [1982], Senior and Blanc [1984], and others somehow led to many ionospheric physicists minimizing electrodynamical explanations for the effects they observed. This reverence for theoretical explanations of magnetospheric processes and/or for the new MHD computer simulation techniques being developed led to a “magnetosphere-ionosphere-coupling setback” that was neither deliberate nor fatal. Today, “penetration effects” are wildly back in fashion, and thus the community has surely recovered from the ~15 year gap evident in Figure 30. In fact, a careful reading of the literature during those 15 years shows that the concept of magnetospheric E field penetration effects deep within the ionosphere [e.g., Spiro et al., 1988], and for durations long enough to matter, was always there to see (and it even made it into textbooks [e.g., Kelley, 1989]). Yet the rediscovery of penetration today is portrayed as being so completely new that the obvious lesson to learn, that initial theory and idealized simulation scenarios cannot displace observational evidence, was, in fact, overlooked. When simulations and data do not agree, it is most likely the computer models that need fixing. Henry Rishbeth is fond of reminding me that it was perhaps D. R. Bates who said “No observational result should be believed until it’s confirmed by a reliable theory.”

9. NEEDED WORK

9.1. Overview

[96] Is the “acceptable level of understanding” for the TEC response during ionospheric storms arrived at nearly 30 years ago still valid today? If science is to be a cumulative undertaking, then to some extent the answer must be yes. Surely, all can agree that past case studies and past statistical patterns obtained during previous solar cycles must be representative of effects seen today; that is, solar-terrestrial physics has not changed. Yet new directions and modes of understanding occur within scientific disciplines, and thus the answer must also be no to the extent that new observational and modeling capabilities offer a fresh look forward. To formulate such an agenda, it is best to begin with a concise summary for the two latitude domains most studied to date.

9.1.1. TEC Storms at Midlatitudes

[97] The storm time behavior of the ionosphere’s total electron content (TEC) for the L = 2–5 domain (as portrayed statistically in Figures 20–23) comes from the largest number of individual storm periods yet studied, and they offer a set of results that are in concert with all past case studies. The percentage of departures from monthly median conditions shows a consistent set of characteristic patterns versus latitude and local time for all seasons. These include an initial daytime positive phase at all latitudes induced by meridional equatorward winds, supplemented by a distinct subauroral “dusk effect” that affects the full afternoon-to-evening sector at all latitudes. Such a large-scale process is caused by the expansion of the magnetospheric convection pattern at ionospheric heights [Lanzerotti et al., 1975]. The dominance of this electrodynamical process in the North American longitude sector is due to the offset of the dipole, an effect that introduces a competition between solar production, winds, and magnetospheric influence as a function of longitude [Mendillo et al., 1992]. Thus for a given geographic latitude, magnetospheric E fields are stronger and wind effects are milder because of the high dip angles near 70°W. For European longitudes the same geographic latitude (and thus solar production) has a smaller dip angle, and thus winds are more effective than E fields in causing vertical motions. The TEC positive phase over Europe spans the full daytime period except at Scandinavian latitudes, where dusk effect enhancements can occur but are milder because of the less dense solar-produced ionosphere at high geographic latitudes. In the Southern Hemisphere the corresponding longitude sector of electrodynamical dominance is that of Tasmania.

[98] Following the daytime positive phase, persistent nighttime enhancements in TEC are caused by the precipitation of plasma sheet particles; trough excursions into midlatitudes and the severe spatial gradients they cause are associated with the contraction of the plasmasphere [Mendillo et al., 1974]. The most long-lived effects of all storm effects, the multiday and often globally affected negative phase, are due to well-known changes in the O/N2 ratio [Prölls, 1995]. All of these effects vary to some extent with solar cycle conditions, and all of the physical mechanisms proposed have been validated by model calculations. It is worth stressing again that new mechanisms are not required to understand ionospheric storms at midlatitudes. What is needed is an improved knowledge of the storm time mix and magnitudes of the same composition changes, winds, and E fields that govern the ambient ionosphere.

9.1.2. TEC Storms at Equatorial and Low Latitudes

[99] The region spanned by the equatorial ionization anomaly (EIA) is governed during storms by the same mechanism (electric fields) that accounts for its daily occurrence. The difference is that the locally induced E field can be
enhanced or reversed by external $E$ fields of direct magnetospheric origin (penetration) or indirect magnetospheric sources (disturbance dynamo). This growth versus contraction of the EIA in storm time and local time, and the often dramatic effects that follow, has been observed repeatedly (e.g., references in Figure 17 [Huang and Cheng, 1991]) and modeled successfully [Batista et al., 1991; Rasmussen and Greenspan, 1993]. Thus the overall scenario for TEC storms at EIA latitudes is well understood, and it is summarized schematically in Figure 18. Following this electrodynamical phase for TEC at low latitudes, composition changes driven by modified thermospheric circulation lead to enhanced O/N$_2$ conditions that result in a prolonged positive phase [Prollis, 1995]. Seasonal and hemispheric asymmetries can occur, especially during strong storms, and thus decreases in O/N$_2$ can also cause negative phase effects at low latitudes.

9.2. TEC Storm Problems in Need of Additional Work

Limiting the scope of this review to the TEC parameter per se, one might ask: “What can the new GPS data sets, the soon to be operational Advanced Modular Incoherent Scatter Radar (AMISR) capabilities, and the radio beacon and in situ diagnostics of the Communications/Navigation Outage Forecast Satellite (C/NOFS) mission contribute to TEC storm studies? What are the modeling issues for TEC disturbances in most need of study?” Given the firm basis for the TEC storm morphology patterns presented in this review, as well as our understanding of the processes that drive them from the equator to the auroral zone, a call for more midlatitude case studies using the GPS network (i.e., beyond the dozen or so already in the literature) seems of lower priority than observations in specific regions where past documentation has been weak. Modeling of the TEC parameter during storms has not been extensive, and thus an emphasis on model improvements tailored for TEC is a crucial need. The simulation of processes that vary with season and solar cycle and with the severity of the geomagnetic storm are required for detailed comparisons with known TEC morphology patterns. Modeling of case study events helps to identify the range of applicability of individual mechanisms; yet some caution is needed to avoid endless scenarios of “this storm did that” results in favor of more conclusive “process modeling.”

Following are some specific observational and modeling needs:

1. How does the TEC in the polar cap change during storms? Past studies that used geostationary satellite radio beacons were limited by long slant-path geometry issues that resulted in averaging out spatial gradients. Given that the region within the auroral oval can be affected severely by in situ precipitation and strong horizontal dynamics, both varying over small spatial scales, more attention should be given to observations at high latitudes. GPS capabilities are not optimal at the very highest latitudes, and thus the new AMISR facility long anticipated for the polar cap should contribute significantly to TEC storm studies, when it is finally allowed to go there.

2. How does the TEC at lower midlatitudes change during storms? If auroral processes of all types move equatorward during storms and equatorial electrodynamics moves the EIA poleward during storms, what happens at their interface? Can GPS data show a consistent answer? Can optical signatures of the EIA (the so-called “intertropical arcs”) be used to map low-to-midlatitude coupling scenarios throughout the night and yield consistent results? Can space-based UV platforms monitor the intertropical arcs over a full 24-hour period? Can ground-based optical aeronomy under daytime conditions be used to track neutral atmosphere variations reliably during geomagnetic storms [Pallamraju et al., 2004]? Can the new C/NOFS mission, especially configured for equatorial/low-latitude applications, contribute to TEC storm studies in addition to its prime focus on ESF issues?

3. At the interface between low latitudes and midlatitudes ($L \sim 1.3$), are there observations that can help us understand the persistent dual maxima in TEC storm patterns shown in Figures 20, 21, and 23? How can case studies address the delay between TEC negative phase onset and the temporal characteristics of the storm [Basu et al., 1995]?

4. Is it time to solve the thermospheric wind effect during storms? Of all the parameters in the upper atmosphere system the least documented one is winds, and yet they are fundamental to the system. Fabry-Perot interferometers can be used in networks to get nighttime winds; their extension to daytime observations (or at least in the dusk sector) is crucial to progress; and new high-resolution instruments now make this possible [Chakrabarti, 1998]. Space-based methods to measure thermospheric winds as a function of height and local time are the single most needed diagnostic in all of space physics, yet NASA and European Space Agency missions never give serious consideration to that need. The only empirical model for winds, the NASA horizontal wind model, rarely gives adequate results during quiet times and never during disturbed times. How TEC changes might be driven by winds remains a premier observational challenge (as pointed out by Buonsanto [1999] in his suggestion of a vertical shear in storm time meridional winds and in models by Schlesier and Buonsanto [1999a, 1999b]). Will coordinated GPS–C/NOFS data sets be able to address this issue in a systematic way?

5. Can more of the existing ionospheric models be improved to yield better altitude coverage so that TEC can be computed reliably out to GPS altitudes? For decades, only one model existed that coupled the plasmasphere to the ionosphere (“the Sheffield model” or SUPIM). It is time to conduct TEC modeling challenges of the coupled GCM-plasmasphere system under disturbed conditions, as done in the past for the quite time $F$ layer [Anderson et al., 1998].

6. Can the low-latitude TEC flux tube depletions and their drift patterns associated with equatorial spread $F$ be simulated successfully during storms? Can the rate of change of TEC at low latitudes, the most easily available signature of ESF from GPS data sets, be modeled in storm time scenarios?

7. How can globally assimilative ionospheric modeling contribute to TEC storm studies? This is, indeed, the new frontier of ionospheric research and one with no counterpart in any other area of space physics. Given the
absence of crucially needed neutral wind observations, can data assimilation models give thermospheric winds patterns during storms that are computationally rigorous and physically viable?

Can models reproduce the longevity of the TEC negative phase versus latitude, local time, season, and phase of solar cycle? Do computed changes in the $\text{O}/\text{N}_2$ ratio match available observations over a full storm scenario? What are the timescales for thermospheric recovery and plasmaspheric replenishment and is one the dominant processes for ionospheric recovery?

Would a new focus on conjugate point studies at both low latitudes [Rajaram and Rastogi, 1970] and middle latitudes [Yeh, 1972], distributed in longitude, help us understand seasonal patterns, longevity issues, and latitude coupling?

### 9.3. A Final Challenge

As an experimental space physicist I was quick to place atop the above list the observational needs from radio, optical and in situ diagnostics. Yet one is forced to notice that most of the needs expressed were in areas of modeling. To return to the empirical approach and to move beyond “an acceptable level of understanding” from observations to the need for models to reproduce TEC storm patterns, Figures 31 and 32 are offered as the basis of a challenge.

In the earliest days of ionospheric storm studies, geomagnetic storms were classified as weak, moderate, and strong, and ionospheric disturbances were analyzed in the same way [e.g., Matsushita, 1959]. The term “strong storm” was later replaced by “great storm” and today by “superstorm.” Unable to resist the current-day trend of coining new names for known effects, I point out that the ionospheric storm depicted in Figures 31 and 32 has long been consider by TEC workers to be a legendary storm! The solar-terrestrial disturbances of May 1967 produced large flare-induced changes in TEC, a pronounced TEC “dusk effect” prior to its official discovery, and the most dominant negative phase in TEC ever recorded. In Figure 31 (left), TEC measurements made in Illinois by Webb [1969] begin with sudden increases in TEC caused by flares on 23 May. These were followed by a day unaffected by geomagnetic activity (24 May) and then by a storm day with a remarkably LT-confined and high-magnitude TEC positive phase on 25 May. Finally, 26 May showed a typical midlatitude TEC negative phase.

In Figure 31 (right), observations from the Washington (D. C.) area published in an earlier paper by
Goodman [1968] show that the afternoon TEC positive phase coincided with increases in the geomagnetic field, suggestive of an electrodynamical mechanism, later termed the “dusk effect” [Mendillo et al., 1970; Mendillo, 1971]. The peak TEC enhancement occurred at 1800 LT (2300 UT) in Washington, the same UT of the peak observed in Champaign-Urbana. The rapid decrease in TEC was interrupted at both stations by enhancements produced by auroral particle precipitation (the poleward wall of the trough). This effect was so pronounced in the Washington, D.C. area (where magnetic latitudes are higher than in Illinois) that amplitude scintillations of the satellite’s radio beacon signal prevented tracking the disturbance for the rest of the night.

[114] Figure 32 shows TEC patterns on the same day but far to the west in Hawaii, as observed by Low and Roelofs [1973]. Figure 32 (top) shows that two geostationary satellites were monitored from the same station in Honolulu, giving TEC results separated (by virtue of their 420-km subionospheric points) by less than 5° of longitude. In Figure 32 (bottom) the bold curve and dots with hatching give control curves (essentially identical for both directions). The thin curves give the daily TEC patterns to the west and east of Honolulu, and they portray an extraordinary response: The two curves begin to diverge at the onset time of the dusk effects in the eastern United States (2200 UT), peak an hour after those in the eastern United States (2400 UT), and do so while maintaining a pronounced difference in magnitude. This strong longitude gradient persists well into the following day. However, the most remarkable aspect of this storm is the next-day response in Hawaii, namely, a negative phase so severe that the ionosphere is unaffected by sunrise!

[115] A comparison of Figures 31 and 32 with Figure 12 reveals the extreme range of possible ionospheric storm morphologies: On 8 March 1970 the positive phase was confined to eastern United States longitudes, and the onset of the negative phase was also on that day from the Midwest to Hawaii. On 25 May 1967 the positive phase occurred from the eastern United States all the way to Hawaii, indicating that the negative phase was too weak to overwhelm the positive phase on day 1 of the storm. Yet the following day (26 May) had a negative phase far more severe at low latitudes (Hawaii) than at subauroral latitudes (Illinois). Moreover, a strong spatial gradient appeared at the lower midlatitude site (Hawaii) that persisted for nearly a day. Successful modeling of these envelopes of multiday ionospheric storm time behavior is the challenge to consider. Are current-day models capable of showing the blend of temporal and spatial mechanisms acting during these events? If not, what is needed to do so?

[116] ACKNOWLEDGMENTS. The preparation of this review was supported by grants from NSF and ONR, and the assistance of Joei Wroten is gratefully acknowledged. The TEC data used from the AFCRL chain of stations were kindly provided by John A. Klobuchar during our many years of fruitful collaboration. His challenge to the TEC community to explain “The day there was no sunrise in Hawaii” (i.e., Figure 32) is the origin of my “legendary storm” designation, a question that has haunted us for over 3 decades. Is it now happily passed on to the modeling community. Several colleagues read a draft of this review and offered valuable comments and suggestions. I particularly thank Henry Rishbeth and Gerd Prölls and appreciate the hard work of two referees who spotted many inconsistencies and ambiguities.

[117] The Editor responsible for this paper was Peter Riley. He thanks two anonymous reviewers.

REFERENCES


Fuller-Rowell, T. J., D. Rees, S. Quegan, R. J. Moffett, and G. J. Bailey (1987), Interactions between neutral thermospheric com-


Lawrence, R. S., D. J. Posakony, O. K. Garriott, and S. C. Hall (1963), The total electron content of the ionosphere at middle latitudes near the peak of the solar cycle, J. Geophys. Res., 68, 1889–1898.


M. Mendillo, Center for Space Physics, Boston University, 725 Commonwealth Avenue, Room 514, Boston, MA 02215, USA. (mendillo@bu.edu)