ANNALI DI GEOFISICA, VOL. 41, N. 5-6, November-December 1998

My first fifty years in ionospheric research

Pietro Dominici
Dipartimento di Fisica, Università «La Sapienza», Roma, Italy

Abstract – The author, as an active witness, recalls what he has seen of the development of ionospheric physics during the last 50 years.

Key words *ionosphere – geophysics – radio communications*

A very personal introduction: Miss Ionosphere and me

Firstly, I wish to thank heartily Professor Enzo Boschi, the President of the National Institute of Geophysics (ING), where I spent the main part of my scientific career, for this splendid occasion of recalling 70 years of my life. The same thanks are due to the colleagues and friends who participated in this iniziative.

My feeling in writing these lines is like a traveller who during a long journey is urged by some of his dearest friends to rest himself (the bureaucrats call that «retirement») and stop for a moment to cast a retrospective glance on what he has done. As my life has been largely occupied by activities concerning ionospheric physics, looking back is for me just like speaking about my relationship with something I have transfigured into a kind of celestial creature, Miss Ionosphere: a changing and charming creature, sometimes smiling and full of nice promises, at other times fleeting and full of mystery.

I met Miss Ionosphere in 1943. At that time she was about 19 (as I will recall later, she was really born in 1924), and I, just graduated from the high school called in Italy «Liceo classico» and about 17, was admitted to the Academy of the Italian Royal Navy, as a cadet in the «Armi Navali», the corps of engineer officers responsible for weapons and communications. According to the armistice agreed on September 8th between Italy and the Allied Nations, this Academy was moved to the southern town of Brindisi. In this provisional location I had, due to the favour of one of my senior officers, the exceptional opportunity of setting up a small HF radio laboratory and entering the complicated world of ionospheric radio communications. On a certain day of December 1943 a severe disturbance, till some breakdowns, occurred on our long-distance HF radio circuits. I now consider this event and the subsequent discussions I had with some senior officers as my first *rendez-vous* with Miss Ionosphere. She appeared to me like a pretty but complicated girl, and for this reason a fascinating personality for a very young man, as I was. A sudden love was born between us, that has happily lasted until now.

Mailing address: Prof. Pietro Dominici, Viale di Villa Pamphili 73, 00152 Roma, Italy.

This tender feeling grew stronger especially when, due to some personal circumstances, I left the Navy for the technical direction of a division concerning naval HF transmitters of an industrial company in Rome and when, at the same time, I directed my studies at the University of Rome towards the ionospheric physics. I consider the year 1950, when I obtained my doctor's degree in physics discussing a thesis on ionospheric absorption of short radio waves, the year of a definitely established bond between Miss Ionosphere and myself. Indeed, another reason for this ideal tie is that at that time I was already working at the Ionospheric Department of the ING, in Rome. I had a very satisfactory career at ING, reaching the direction of the Ionospheric Department in 1970 till 1988 and then the direction of the Institute in 1976-1979. In 1988 I definitively moved from ING to my current teaching post at the University of Rome. My activity here was from 1960 on general geophysics, from 1962 on electromagnetism, and from 1981 till the present day on geomagnetism and ionospheric physics.

So, I have had plenty of opportunity to follow the progress of ionospheric knowledge over a long period of time, that I call, in order to gain the benevolence of the Gods, the first fifty years of friendship with my dear Miss, *pardon!*, Lady Ionosphere.

But, aside from this joking, what exactly is the terrestrial ionosphere?

On the definition of the ionosphere

One of the effects of the electromagnetic and corpuscular radiations, by which the Sun radiates the Earth, is the ionization of the molecules of the atmospheric gases and the consequent production of free electrons and ions (also by secondary processes arising from ionization). A much less important ionization is produced by non solar agents, such as the cosmic rays.

With reference to the most important and almost exclusive ionizing agent, that is the ultraviolet and X components of the electromagnetic solar radiation, its intensity decreases when the altitude decreases due to atmospheric absorption. On the other hand, as the altitude decreases the atmospheric pressure, and therefore the number of the ionizable gaseous molecules, rapidly increases. A natural consequence of these two circumstances is that the concentration of charged free particles, *i.e.* electrons and ions, reaches its maximum value at an intermediate altitude between the extreme upper atmosphere, where the intensity of the solar radiation is greater, and the terrestrial surface, where the concentration of ionizable molecules is larger. More precisely, at an altitude of 200÷400 km, depending on the geographical position, the hour of the day, the season of the year and the phase in the cycle of the solar activity.

The *ionosphere* (a contracted form for *ionized atmosphere*) is the part of the Earth's atmosphere where the total concentration N_t of charged particles (free electrons and ions) is large compared to that in the lower and upper atmosphere. The conventional interval of the ionospheric heights is from 50 km ($N_t \approx 10^6$ m^{-3}) to 1000 km ($N_t \approx 10^9$ m^{-3}), the maximum value being approximately $N_t \approx 10^{12}$ m^{-3} at the above mentioned height of 200 ÷ 400 km. At heights above 1000 km the *magnetosphere* begins, where the atmosphere is completely ionized and N_t monotonically decreases, and extends up to the height of the *magnetopause* (about 60000 km towards the Sun and over than one million of kilometers in the opposite direction), where the terrestrial atmosphere loses its individuality due to the fact that its material concentration reaches values pertaining to the surrounding circumterrestrial space. A capital difference between the ionosphere and the magnetosphere is that in the first region, especially in its lower part below the maximum of ionization (*i.e.* between 50 and about 300 km), a consistent population of neutral molecules is present whose concentration largely exceeds that of the charged particles (at the maximum of N_t there are about one thousand neutral particles per each charged particle!).

The first definition of the ionosphere, which dominated until a few years ago, was rather different from the above definition, the ionosphere being «the atmospheric region where the concentration of charged particles appreciably modifies the atmospheric index of refraction for radio waves».

This definition reminds me of the early development of ionospheric physics, from the above cited discovery of the ionosphere in the Twenties till about the Seventies, when a strong bond was present between ionospheric physics and the very important industry of long-distance radio communications by short waves.

The first introduction of the concept of the ionosphere

It is well known that the connection between the concept «ionosphere» and long-distance radio communications arose immediatly after the successful experiment of Guglielmo Marconi in connecting Poldhu, in Cornwall, and St. John, in Newfoundland, across the Atlantic Ocean by radio waves at a frequency of about 300 kHz (wavelenght 1000 m), on December 12, 1901; about one year later, the regular radio telegraphic service between Europe and America started, by waves of about 71 kHz (4200 m) in one direction and about 61 kHz (4900 m) in the other.

One of the hypotheses that were advanced in order to reconcile this result with the known laws on the propagation of electromagnetic fields, in particular concerning the straight line propagation in free space, was almost contemporaneously advanced in 1902 by Oliver Heaviside in U.K. and Arthur E. Kennelly in the U.S.A. The problem relating to the curvature of the Earth's surface (like an obstacle about 250 km high on the distance of about 3700 km between Poldhu and St. John) was solved because the radio waves were reflected downwards by a sufficiently ionized, *i.e.* sufficiently conducting, layer in the high atmosphere, the zone which some years later was called the «ionosphere» (fig. 1).

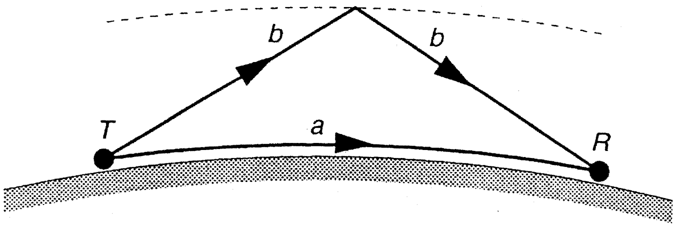

Fig. 1. The two hypotheses advanced in 1901-1902 to explain the successful transatlantic experiment of G. Marconi on December 12, 1901: (a) the radio waves travel from the transmitting station *T* to the receiving station *R* following the terrestrial surface (*ground waves*, or *terrestrial waves*); (b) the radio waves are reflected towards the Earth's surface by a reflecting layer in the upper atmosphere (*sky waves*, or *ionospheric waves*).

This hypothesis was soon laid aside because a theory was developed which considered the penetration in the conducting Earth's surface of the electromagnetic field of the incoming radio signal, whose wave surfaces are modified and inclined forward, thus following the terrestrial curvature.

Nevertheless, the frequent occurrence of unexplainable events kept alive the «ionospheric hypothesis» and contemporaneously gave birth to the interest in radio waves whose length became shorter than those normally used. In particular, this was the case of the radio waves that were often used in naval communications because the available antennas for ships were shorter in geometrical dimensions than those used between terrestrial stations (the wavelengths for naval communications were often of the order of some hundreds of meters instead of some kilometers, *i.e.* at frequencies of some hundreds of kHz instead of tens of kHz). A typical event of this kind was the fact that ships were frequently unable to communicate at nighttime with stations a few hundreds of kilometers away while at the same time they easily communicated with stations several hundreds of kilometers away.

Many investigations were soon carried out by several engineers and physicists on these strange events, including some amazing results obtained by radio amateurs. Undoubtedly, it appeared in the early Twenties that using radio waves of some tens of meters in length (frequency of a few MHz) it was possible to cover enormous distances (even intercontinental ones) with very small radio powers (some tens of watts) compared with those used with long or very long waves (some tens or hundreds of kilowatts). The only mechanism of propagation which appeared to be capable of simply explaining these events was the «ionospheric reflection», so that efforts were then made to support this «ionospheric hypothesis» with direct evidence of the real existence of a radio reflecting layer in the upper atmosphere.

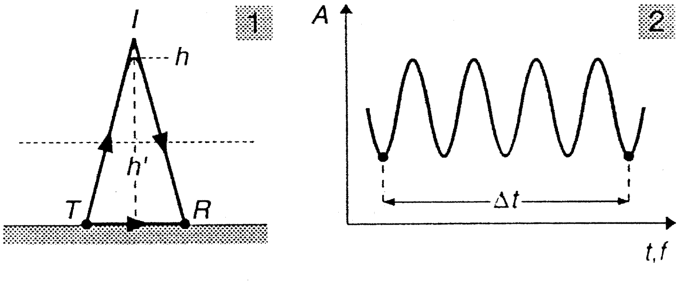

Fig. 2. (1) Scheme of the radiointerferometric experiment of E.V. Appleton and M.A.F. Barnett (December 1924-February 1925). A radio receiver *R* at the Radio laboratory of the Oxford University measured by a galvanometric device the intensity of the signals emitted by the BBC broadcasting station at Bournemouth, whose frequency regularly changed in time (after midnight, when the normal transmissions ceased); there was interference between direct waves *TR* and ionospheric waves *TIR*. (2) Because the frequency f and hence the distance between successive fringes of the space interference pattern changed with time t, the amplitude A of the received signal periodically changed with time. The virtual h' of the reflecting layer was deduced from the number of cycles of A observed during a certain time interval Δt. For instance, in one of the experiments 7 fringes were counted corresponding to a frequency variation of 759 to 779 kHz in 30 s and, being $TR = 125$ km, the result was about 97 km for the virtual height h' and about 92 km for the real height h.

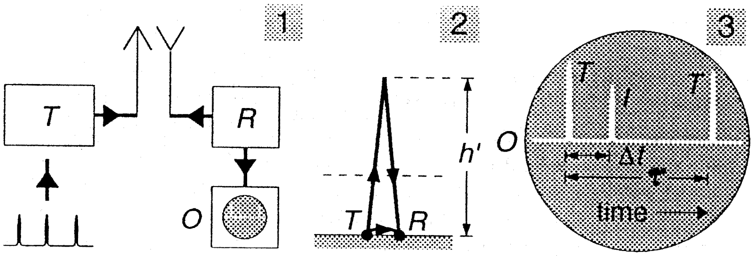

Fig. 3. (1) Scheme of the radioechometric experiment of G. Breit and M.A. Tuve (October 1925). A radio transmitter *T* (4.3 MHz, 10 kW peak power), near Washington, D.C., vertically emitted periodic short pulses (duration 150 μs, repetition period τ = 12,5 ms); a near receiver *R* at the Anacostia district of Washington (13 km apart) presented the possible echoes from the ionosphere on an oscilloscopic device *O*. (2) Scheme of the ray trajectories. (3) Oscilloscopic pattern: if Δt is the time delay between the transmitted pulse *T* and the corresponding ionospheric echo *I*, the virtual height of the reflecting layer is $h = c\,\Delta t/2 = 0.15$ km per μs of time delay, *c* being the speed of light in vacuum.

This evidence was obtained in two famous experiments, the interferometric experiment by Edward V. Appleton and M.A.F. Barnett at Oxford, England (December 1924 - February 1925: fig. 2) and the echometric experiment by Gregory Breit and Merle A. Tuve near Washington, D.C., U.S.A. (October 1925: fig. 3), that showed the existence of a radio reflecting layer at an altitude of about 100 km. In honour of the first proponents of the ionospheric hypothesis this layer was named the *Kennelly-Heaviside layer* and then (with reference to the current symbol, *E*, of the electric component of the electromagnetic field of the radio waves which had been measured in the Appleton-Barnett experiment), the *E layer* (now it is more common to say *E region* because about at the same altitude of this regular layer there are some other irregular layers, pertaining to the so-called *sporadic E*).

At the end of 1925 Appleton discovered another more ionized and higher layer, at an altitude of 250 ÷ 350 km, named the *F layer* (because it was «after» the *E* layer) or *Appleton layer* (now, *F region*). Later on, other minor ionospheric layers were discovered below the *E* layer and alphabetically named according to their altitude from the *E* layer as *D, C,* etc., but now only the *D region* is considered below the *E* region. The former and lower layers *C, B, A* are considered, when really observable, as minor particularities of the *D* region. Similarly, some irregular layers appearing especially at sunrise in the *F* region are now regarded as sporadic particularities of this region.

It should be recalled that the term «ionosphere» was introduced in 1929 by Robert A. Watson-Watt, an English radio engineer to whom the realization in 1935 of the first operating radar from the Breit and Tuve experiment of fig. 3 is due.

The first development of ionospheric physics

As already stated, the first development of the ionospheric physics occurred with a strong interaction with the development of long distance radio communications by short waves. This part of the radio engineering and radio industry supplied physics with a valuable amount of experimental data to be studied and physics, in turn, supplied radio engineers with the necessary indications for the best management of the available radio frequencies.

In fact, while reaching large distances using long radio waves is almost exclusively a question of radiated power, radio communications by short waves exhibit, together with some great advantages (low radiated power, simple and relatively small antenna systems, etc.), great disadvantages. Firstly, the operating frequency between two given terminals must be carefully chosen according to the actual ionization of the upper atmosphere along the radio circuit, *i.e.* according to the hour of the day, the season of the year, the phase of the solar activity. Secondly, it is necessary to have the capability of changing the frequency and sometimes also the geometry of the radio circuit when one of the possible sudden major disturbances in that ionization occurs.

The first important theoretical contribution given by ionospheric physics was the *magnetoionic theory* (Wilhelm Altar 1926, E.V. Appleton 1928-1932) regarding the propagation of radio waves in the ionosphere. This theory showed that the most important parameter for the atmospheric propagation of the radio waves used at that time (ranging from very long waves to short waves, *i.e.* with frequencies from some tens of kHz to a few MHz) is the *concentration of free electrons* (the so-called *electron number density* or, more concisely, *electron density*), from which the index of refraction of such radio waves mainly depends. An already well accepted fact was that free electrons in the upper atmosphere are produced mainly by solar UV and X photons. So, the availability of a theory on atmospheric ionization by solar photons became important and in 1931 just such a theory was developed by Sidney Chapman.

It is important for some of the future considerations to recall the basic hypotheses of this theory.

a) Ionospheric ionization is adequately represented by the *electron density*, N, for which in general one has the equation of continuity

$$dN/dt = p - l \tag{1}$$

t being the time, p the general production rate of electrons, l the general loss rate of electrons; the term «general» means all possible processes of production and loss but in this theory only the *solar photoionization* and *ionic recombination* are considered for electron production and loss respectively, and then (1) becomes:

$$dN/dt = q - \alpha N^2, \tag{2}$$

$$q = q_M \exp\{1 - [(h - h_M)/H] - \exp[-(h - h_M)/H] \sec\chi\}, \tag{3}$$

q (the *Chapman function*) being the electron production rate by solar photoionization, χ the Sun's zenith angle for the site of interest, q_M the maximum value of q (occurring when is $\chi = 0$, *i.e.* when the Sun is at the zenith of the site), h the generic altitude, h_M the altitude where q attains its maximum q_M, H the scale height (the increase in h corresponding to a decrease of exp(− 1) $\approx$ 37% in atmospheric pressure under conditions of hydrostatic equilibrium), α the *recombination coefficient* between electrons and the species of positive ions to be considered.

b) The atmosphere is: 1) chemically homogeneous, *i.e.* composed of only one molecular species; 2) thermally homogeneous, *i.e.* isotherm, and 3) in hydrostatic equilibrium, *i.e.* without any vertical or horizontal movement.

Later on the *capture* of electrons by neutral molecules or atoms was also considered an electronic loss process (it is called the *attachment* process). In this case there is a more general form for (1):

$$dN/dt = q - \alpha N^2 - \beta N, \tag{4}$$

β being a suitable *electron capture coefficient* (or *attachment coefficient*) by the molecular species of interest.

Chapman's theory was a powerful means to coherently organize the large quantity of ionospheric data arising both from the practice of industrial HF radio communication networks and from synoptic world-wide measurements of the vertical distribution of electron density by the special variable-frequency vertical radio transmitting-receiving devices called *ionosondes*.

A necessary parenthesis on ionospheric measurements

The first instruments to directly measure the electron content of the upper atmosphere were *ionosondes* (the radio instruments that are now called *coherent pulse vertical ionosondes*). They were based on the same principle of the previously cited experiment of Breit and Tuve, *i.e.* the vertical emission of short pulses of radio waves in the range approximately 1 ÷ 30 MHz (due to great technical difficulties, it is very hard to lower the frequency range to 50 kHz or so as would be necessary to measure the electron density in the bottomside ionosphere). They were also based on the recording of possible echoes from reflecting ionospheric layers in such a manner to have, when varying the radio frequency, the diagram of the echo time delay *versus* the radio frequency. The echo time delay is easily interpreted as the *virtual height* of the reflecting zone, that is the reflection height if the speed of propagation of the radio pulses was everywhere in the reflecting zone that of light in vacuum. Some results from the above quoted magnetoionic theory allow the electron density to be deduced from the frequency of the reflected signals. In particular, the electron density N_p where a vertically upwards radio wave is reflected is proportional to the square of the frequency itself (*penetration frequency*, f_p, at the height of interest)

$$N_p = 1.24\ 10^{-2}\ f_p^2. \tag{5}$$

This relation also applies to the maximum electron density of a layer, N_M, wich corresponds to the maximum penetration frequency for this layer, called *critical* (*ordinary*) frequency, f_0 (*f.i.* f_0E for the E layer): the waves at higher frequency pass through the layer. The term «ordinary» refers to the ordinary mode in double refraction occurring to the incoming radio wave in the reflecting region. The same results, together with the use of a suitable ionospheric model, allow the corresponding *real height*, h, to be deduced from the virtual height h' and thus the vertical electron density to be plotted in the range from the lower E region (about 100 km) to the absolute maximum in the F region (250 ÷ 350 km). These graphs, both in terms of critical frequency and electron density, are called (*vertical*) *ionograms*.

Ionosondes of this kind were put into regular service in the early Thirties (the first one was that near Washington, D.C., U.S.A., in 1931). It is a pleasure for me to recall that in 1932 a ionosonde was realized in Camerino, Central Italy, by Ivo Ranzi, who at that time was assistant professor of Physics at the local University. In 1936 another ionosonde was put into regular service at ING in Rome by the same I. Ranzi, at that time full professor of Physics at the University of Florence, and Antonio Bolle, who was assistant professor of Physics at the University of Rome: two people who have been very important for me, first as splendid teachers and then as dear colleagues and friends. In order to complete the panorama of the Italian contribution to the development of ionosondes, two other ionosondes, both based on the heterodyne principle were introduced by Peter G. Sulzer in 1946, one built at ING in 1949 by A. Bolle and Carlo Alberto Tiberio (at that time professor of Physics in a Technical High School) and the other built by myself, also at ING, in 1951. The latter sounder remained in operation, care of C.A. Tiberio, until 1958. Around this time, the policy of home-made ionosondes was fruitfully replaced almost all over the world by resorting to excellent instruments built by some specialists with semi-industrial criteria.

The vertical ionosondes are still the basis of ionospheric measurements, by the synoptic operation (at each zero instant of UT hours or more frequently) of the worldwide network of ionospheric observatories.

Many of these observatories use advanced instruments. The first were the *digital ionosondes*, where the generation of sounding signals and the processing of the received echoes are completely performed by digital electronic circuits. The first ionosonde of this type, the famous «Digisonde 128», was designed and built in 1975 by another old and dear friend, Klaus Bibl at the University of Massachusetts at Lowell. Another advanced instrument is the *frequency-variation ionosonde* (also called *chirpsounder*), where a continuous radio wave whose frequency linearly increases in time is vertically transmitted and the height of the received echo is deduced from the difference between the frequency of the transmitted signal corresponding to the echo and the frequency when the echo is received, which obviously measures the echo delay (the first chirpsounder was realized in 1972 by George Barry in the U.S.A.).

Pertaining to these two types are the ionosondes that in recent years have been used at ING ionospheric stations. In 1979 the Rome station was equipped by a Digisonde 128P, to which a simplified digital ionosonde produced by Kel Aerospace Co. in Australia and a vertical/oblique chirpsounder VOS-1/ED 02 Barry were added in progress of time for special uses. In 1993 the Kel ionosonde was moved to the ING geophysical station in Antarctica. In 1997 the Digisonde 128P was moved to the Gibilmanna station, in Sicily, where it is still working, and at the Rome station was replaced by its most recent version, the transportable DPS4 Digisonde 256; this instrument is able to scale the ionograms and the corresponding electron density vertical profiles automatically and also offers a series of new measurements such as thermospheric winds, etc.

I think that other colleagues could write better than me about modern ionosondes and their improved uses: excellent examples can be found in this special issue.

Let me briefly recall some other methods of ionospheric measurements for some important purposes, different from normal sounding.

a) *Active radio methods* – Concerning «ionospheric sounding», four other techniques of active measurements deserve mention.

The first technique is *fixed-frequency continuous vertical sounding*, performed by a simplified normal vertical ionosonde, continuously operating at a fixed frequency. This technique, which is of incomparable utility in monitoring the dynamical phenomena of the ionospheric layers, was introduced in a systematic manner in 1948-1950.

The second technique is *oblique ionospheric sounding*, both *monostatic* (or *by ground back-scattering*) and *bistatic*.

Ground back-scattering sounding uses a fixed-frequency or a normal variable-frequency pulse ionosonde, but with horizontal and not vertical antennas. It was introduced around 1953. During the period 1957-1961 I was responsible for some ionospheric experiments regarding the International Geophysical Year 1957-1958 conducted at the famous experimental radio station of Torrechiaruccia, at Santa Marinella on the coast north of Rome, where G. Marconi performed his last experiments on the detection of moving objects by microwaves. I gained a good experience in the technique of ground back-scattering, with a home-made 2 kW fixed-frequency (18.6 or 22.3 MHz) pulse ionosonde connected to a rotating Yagi antenna (1 revolution per minute). For these experiments I also used, in connection with the above mentioned equipment, a «Panorama» ionosonde, realized in 1954 in Germany by K. Bibl, which, in my opinion, has been the most advanced type among all analogic ionosondes. I find it surprising that sounding by ground back-scattering has been so little used, when one considers on the one hand its great simplicity and flexibility and, on the other, its remarkable capability in monitoring the state of the ionosphere to some thousand of kilometers all around, and especially in following large-scale dynamical phenomena, such as travelling disturbances and ionospheric storms.

Several years after ground back-scattering another technique was introduced, namely *bistatic oblique sounding*, where the transmitting part of a normal ionosonde is placed at one of the terminals of a chosen circuit and the receiving part is placed at the other terminal, an accurate synchronization between the two parts being necessary. Such a system is very advantageous from a technical point of view, because one can appreciate directly the situation concerning the actual modes of propagation along the circuit, and in particular the effective MUF, the frequency intervals where single-mode propagation occurs, and so on. In fact, it is frequently used in communication systems when reliability is the main requirement, as it typically occurs in military communications. Several scientific experiments have also been performed by this technique. For instance, a successful experiment at very long distance was conducted in 1993 by Bruno Zolesi and Cesidio Bianchi at ING from Rome to Xin Xiang (Henan), China. Nevertheless, in my opinion its cost-benefit ratio from a geophysical point of view is, all being considered, less favourable than that of ground backscattering sounding.

The third major technique is *vertical incoherent back-scattering sounding* (or *Thompson scattering*), based on a ionosonde similar to a normal pulse vertical ionosonde but operating at a higher fixed frequency (of the order of 50 MHz or more) and higher peak power (some MW). At these frequencies, no reflection occurs in the ionosphere as it does at much lower frequencies, but a non directional scattering: a very small part of the incident radio power is incoherently back-scattered towards the transmitting antenna by single electrons and ions of the ionospheric plasma. When analyzing the spectrum of the back-scattered pulses the density and temperature of electrons and of the prevalent ionic species can be deduced, while the altitude is deduced from the echo delay and the possible vertical velocity of the back-scattering zone is deduced from the Doppler frequency shift of the echo.

The fourth major technique is the normal pulse sounding by special ionosondes aboard a satellite orbiting at an altitude higher than the absolute maximum of ionization, namely in the so-called *topside ionosphere* (altitudes from about 300 to 1000 km). The first measurements of this kind were performed in 1964-1965 with the Canadian satellite *Alouette*.

b) *Passive radio methods* – Among these methods (measurement of the angle of arrival, Doppler shift, polarization, etc. of radio signals), the most important is the measurement of the Faraday rotation of the plane of polarization of VHF radio signals from satellites, which gives the integrated electron density (the so-called *Total Electron Con-*

tent, TEC) along a column of unitary section centered on the trajectory of the signals from the satellite and then along lines practically crossing the denser part of the ionosphere. The first measurements of TEC were performed in 1960 from a low orbiting satellite and in 1963 from a geostationary satellite.

c) *Direct measurements by instruments on space vehicles* – Since about 1960 direct *in situ* measurements of chemical and physical quantities in the ionosphere have been made by instruments carried by space vehicles. The most interesting instrument is perhaps a simple electrostatic device, the *Langmuir probe*, which measures the number density and temperature of electrons and ions.

Table I summarizes the principal ionospheric measurements.

Table I. Principal aeronomic measurements.

Domain	Measurement	Measured
NEUTRAL ATMOSPHERE		
	Direct measurements	
Astronautic heights	Flight-time mass spectrometers	Composition, density, T
	Ionic sondes (*f.i.* Langmuir)	Composition, density, T
	Indirect measurements	
≤ 1000 km	Atmospheric drag	Density
≤ 300 km	Spectrometry of atmospheric luminescences	Composition, density
≤ 120 km	Artificial clouds	Movements
80÷120 km	Radar observation of meteors	Movements
≤ 100 km	Acoustical sounding	Movements
≤ 100 km	Laser sounding	Movements, composition
IONOSPHERE		
	Direct measurements	
Astronautic heights	Ionic sondes (*f.i.* Langmuir)	Composition, density, T
	Indirect radio measurements	
≤ 1000 km	Incoherent vertical sounding	N_e, T_e, N_i, T_i
≤ 1000 km	TEC measurement	TEC
≥ 300 km	Downwards sounding from satellites	$N(h)$ topside ionosph.
≤ 300 km	Upwards vertical sounding	$N(h)$ bottomside ionosph.
≤ 300 km	Oblique sounding	General monitoring
≈ 100 km	LF and MF partial reflection	N_e, ν_e

Symbols: *e*, electrons; *h*, height; *i*, ions; *N*, concentration; *T*, temperature; ν_e, electron collisional frequency.

The main characteristics of the «normal» ionosphere

The data progressively arising from normal operation in each ionospheric station were soon organized in a manner fairly similar to that used for geomagnetic data: the hourly data of each day were assembled in monthly tables, the 28 ÷ 31 values for each hour providing their monthly median value. This procedure was generally adopted already in the late Forties.

Reference to median (better than mean) values for a relatively long period of time (a month) is very useful to cancel most of the great ionospheric variability with small (some hours) and medium (few days) periods, which constitutes a great drawback in studying ionospheric morphology on a large space-time scale. Therefore, it was natural to introduce as the *normal* (or *mean* or *reference* or *standard*) *ionosphere* the ionosphere just represented, for each layer, by numerical tables or graphs of hourly median monthly values of the critical frequency f_0 (or, equivalently, maximum electron density N) and minimum virtual height h', which are arranged in terms of the hour of the day, or the year or the *sunspot number R*, a dimensionless number, varying between 0 and about 200 with a quasi period of about 11 years, which is the most popular index of the solar activity (figs. 4 to 7).

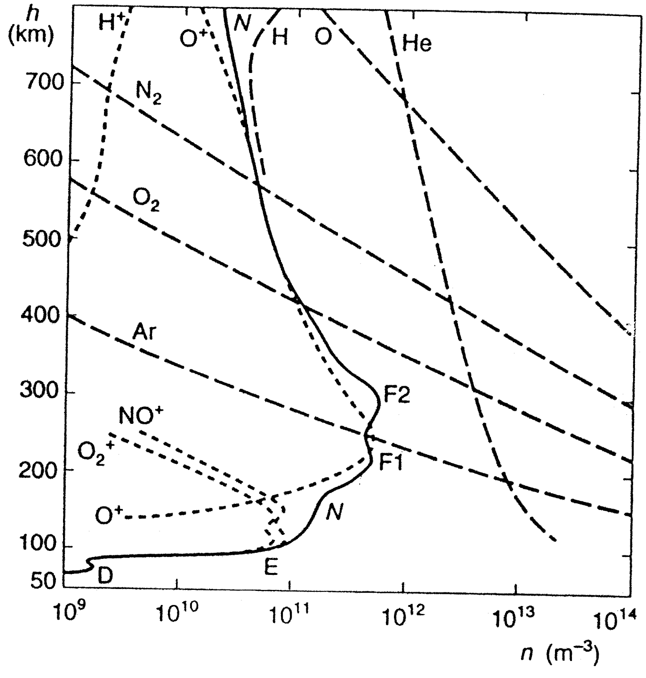

Fig. 4. Mean values of the number density n of the principal neutral and ionic species in the ionosphere (N is for free electrons) around local noon in summertime during a period of medium solar activity at mid-latitude. The symbols D, E, F, F_1, F_2 indicate the normal ionospheric regions and layers. (Adapted from P. Dominici, *Ionosfera*, in *Enciclopedia delle Scienze Fisiche*, III, 298-312, Roma 1993).

It is important to remember that the ionospheric knowledge mainly derives from measurements of electron density, *i.e.* that most of what we have called «ionosphere» to date should be more precisely called *electronosphere*. It is necessary to assume that, as a first but good approximation, the electronosphere depicted by the normal ionosphere sufficiently well represents the *global ionosphere*, namely that the electron density is a good index of the total atmospheric ionization, both electronic and ionic, at ionospheric heights.

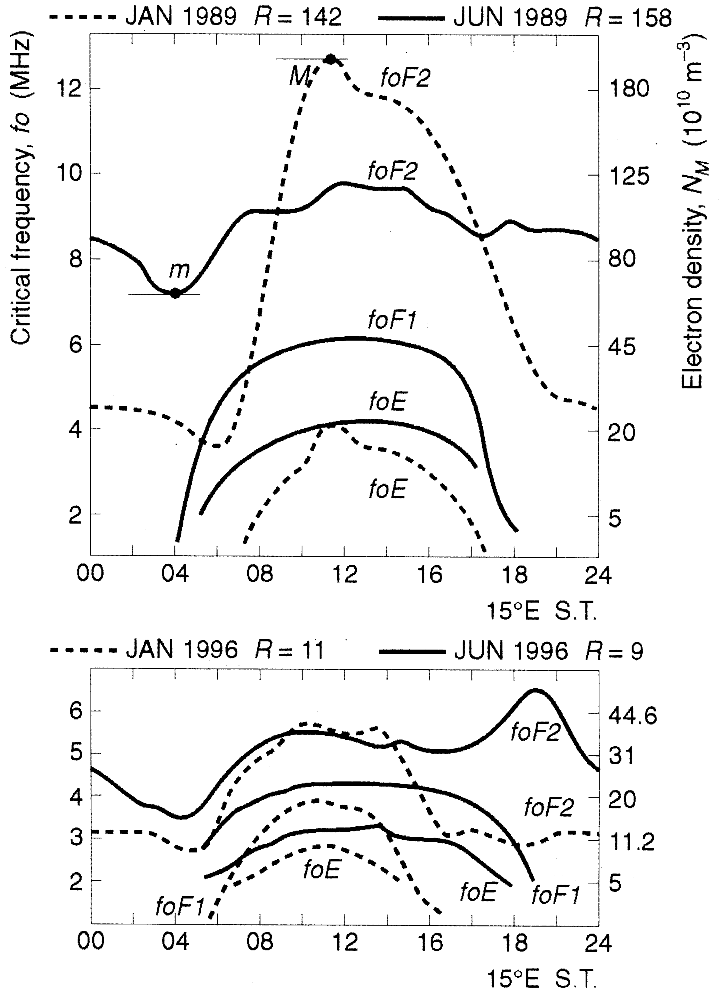

Fig. 5. Hourly behaviour at Rome of monthly median values of critical frequency f_0, and maximum electron density N_M of the E, F_1, and F_2 layers for the months of January (broken lines) and June (continuous lines), both in a year of high (1989) and low (1996) solar activity. (Courtesy of B. Zolesi, ING).

The principal mean characteristics of the normal ionosphere at middle latitudes are:

a) The vertical distribution of the electron density N never exhibits a non monotonic character (fig. 4).

a1) *Regular regions and layers*: depending on latitude and solar activity, during the daytime in medium solar activity periods N increases from $\approx 10^6$ m^{-3} at the lower limit of the ionosphere (50 km) to $\approx 10^9$ m^{-3} in the *D region* ($50 \div 90$ km), whose prevalent ions are NO$^+$ and O$_2^+$ produced by Extreme UltraViolet (EUV) solar radiations, solar X rays in the wavelenghth interval $0.1 \div 10$ nm and Cosmic Rays (CR), further increases to $\approx 10^{10}$ m^{-3} in the *E region* ($90 \div 140$ km), whose prevalent ions and ionizing agents (CR apart) are the same as in D region, and finally increases to $\approx 10^{12}$ m^{-3} in the *F region* ($140 \div 400$ km), whose prevalent ions are NO$^+$, O$_2^+$ and O$^+$ produced mainly by solar X rays in the wavelength interval $10 \div 100$ nm and UltraViolet (UV) solar radiation. Above this maximum, N monotonically decreases and the prevalent ions become the lighter ones, as, in order of increasing height, O$^+$, He$^+$ and H$^+$. Because of the high density of neutral particles, diffusion is almost negligible and ionic recombination prevails over capture as the electron loss process in D, E and lower F regions, while diffusion is important and capture prevails in upper F.

a2) *The F region seasonal splitting*: in summer and equinoctial months during years of medium and low solar activity the F region shows two maxima of N in daytime hours (figs. 4 and 5); the lower one ($140 \div 280$ km, $N_M \approx 5\ 10^{11}$ m^{-3}), with the heavier ions (NO$^+$, O$_2^+$) is the called F_1 *layer* and the other ($280 \div 1000$ km), the F_2 *layer*, with lighter ions (O$^+$, He$^+$, H$^+$), is the most ionized part of the whole ionosphere ($N_M \approx 10^{12}$ m^{-3}). This seasonal splitting of the F region was recognized in 1933.

a3) *Irregular layers and sporadic E*: some sporadic layers are sometimes recognizable on the normal ionospheric structure. The most important are some occasional layers occurring between E and F_1/F_2 layers at sunrise and especially those occurring, often with very high ionization (sometimes exceeding the F ionization), at the height of the maximum ionization of the E region ($110 \div 140$ km) which have different origins and, as previously recalled, are known, as a whole, as *sporadic E*.

b) *Solar control on D and E regions and the F_1 layer*: according to Chapman's theory, there is a strong solar control on atmospheric ionization: electron density is greater in the daytime than at night, in summer than in winter (but non for the F_2 layer, as we show later on), in years of high solar activity than during those of low solar activity (fig. 6).

c) *F region anomalies*: there are three noticeable anomalies in the behaviour of the F_2 layer with respect to the F_1 layer and E region, and precisely (see, for an example, fig. 5).

c1) *F_2 seasonal anomaly*: N_M occurs in the daytime of winter months instead of summer months.

c2) *F_2 daily anomaly*: the daily maxima of ionization often occur just before and after the local noon in wintertime and near the local sunset in summertime.

c3) *F_2 night anomaly*: in summertime large ionization exists at nighttime, which is not much lower than that in daytime.

d) *Ionosphere and neutral atmosphere*: as previously pointed out, the most important part of the ionosphere both from a geophysical and a radioengineering point of view, namely the *bottomside ionosphere* (to the absolute maximum in the F region: $50 \div 300$ km), lies in a much denser neutral atmosphere (at the altitude of the absolute N_M in the F region there are about a thousand molecules or atoms for each free electron or ion).

Table II summarizes the main characteristics of the permanent ionospheric regions and layers at middle latitudes.

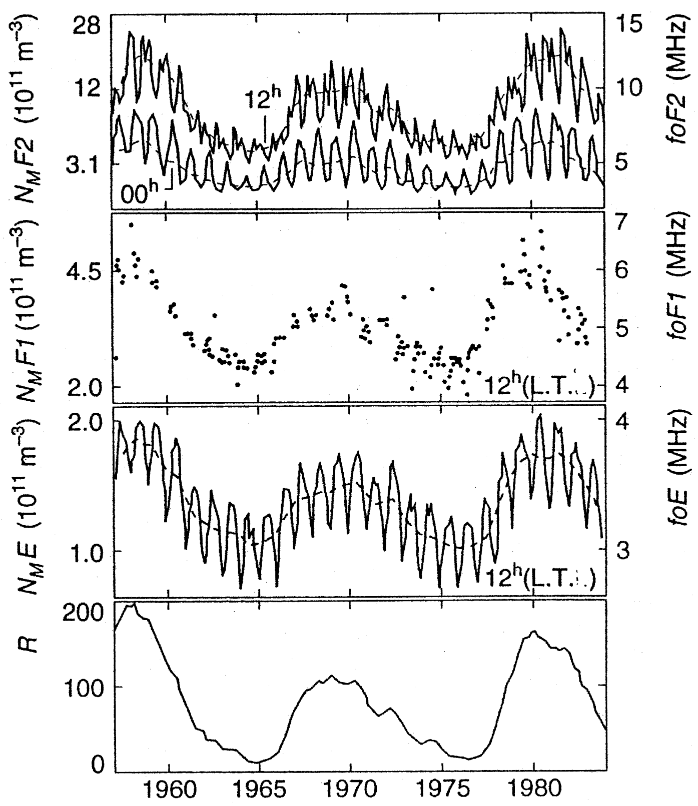

Fig. 6. Long-term behaviour at Rome of the monthly median values and of their running 12-month means (broken lines) of f_0E and f_0F_1 critical frequencies at 12 h of local time (L.T.) and of f_0F_2 critical frequencies at 00 h and 12 h of local time during the period 1957-1984; at the bottom, the behaviour of the monthly sunspot number *R* (from P. Dominici and B. Zolesi, *Il Nuovo Cimento*, part C, X, 191-208, 1987).

The *polar ionosphere* is not normally included in the normal ionosphere, because at high latitudes the ionosphere exhibits very special peculiarities, even in its median characteristics. The most important of these is a general great instability in ionospheric parameters, arising from the strong interactions of ionospheric particles with the so-called *solar wind*, a fast flow from the Sun of mainly high energy protons (H^+), α-particles (He^+) and electrons that the geomagnetic field deviates towards the polar caps. Consequently some important phenomena occur, among which the magnificent *polar aurorae* and the generation of systems of electrical currents, coupling the ionosphere to the geomagnetic field and magnetosphere (*polar electrojet*). Because of these couplings, a major precipitation of electrons occurs in the polar ionosphere from the surrounding magnetosphere. The electron scattering from the mid-latitude ionosphere, induces the *polar night anomaly*, *i.e.* the presence of a consistent night electron density in wintertime, when direct solar ionizing radiation is absent for very long periods.

Table II. Scheme of the normal ionosphere.

Region	*D*	*E*	*F*	
Layer	*D*	*E*	F_1	F_2
Height interval (km)	$50\div90$	$90\div140$	$140\div250$	$250\div1000$
Prevalent ionic species	$NO^+\div O_2^+$	$NO^+\div O_2^+$	$NO^+\div O_2^+$	$O^+\div He^+\div H^+$
N_{max} : daytime (m^{-3})	10^9	10^{10}	$5\ 10^{11}$	10^{12}
nighttime) (m^{-3})	10^8	10^9		10^{11}
Material density (m^{-3})	10^{22}	10^{17}	10^{16}	$10^{15}\div10^{13}$
e preval. photoproduc. by	EUV, hard X, CR	EUV,hard X	UV, soft X	UV
e positive diffusion	Negligible	Negligible	Weak	Strong
e loss process	Recombination	Recombination	Recombination	Capture
(negat. diffus. apart)	($\alpha \approx 10^{-12} m^3 s^{-1}$)	($\alpha \approx 10^{-14} m^3 s^{-1}$)	($\alpha \approx 10^{-14} m^3 s^{-1}$)	($\beta \approx 10^{-4} s^{-1}$)
Dependence on site and time		Simple (Chapman-like)		Complex

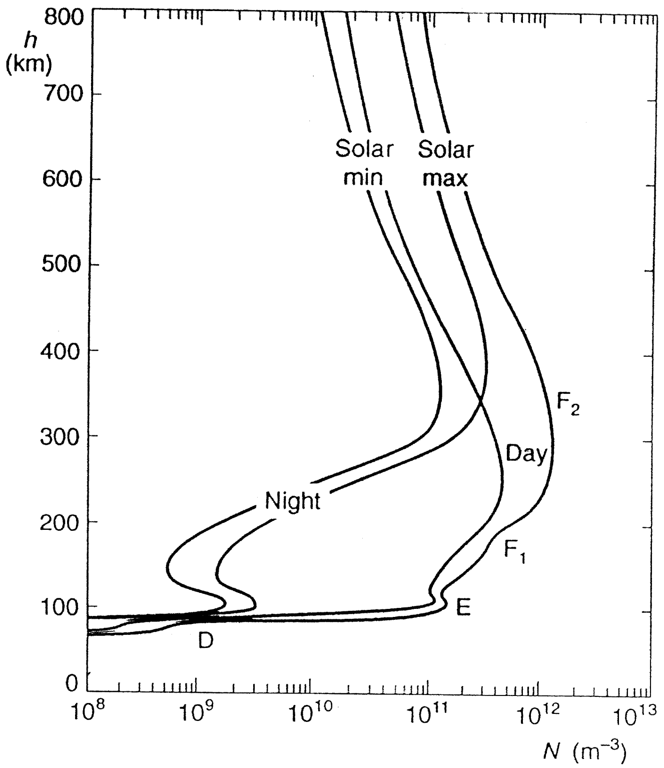

Fig. 7. Typical mid-latitude electron density *N* profiles at maximum and minimum solar activity for both daytime and nighttime conditions (from M.H. Rees, *Physics and Chemistry of the Upper Atmosphere*, Cambridge, U.K., 1989).

The real ionosphere

The actual daily behaviour of ionospheric characteristics (*f.i.*, *N*/time graphs as in fig. 8.1, or *N*/height/time tridimensional graphs as in fig. 8.2 and 8.3) shows that there are wide differences from day to day and from median behaviour. So, there are other pecularities to be studied and explained, besides the previously mentioned anomalies recognizable in the normal ionosphere.

These pecularities arise both from the irregularities of the ionizing solar radiation and large-scale ionospheric dynamics. Among the irregularities of the solar activity, particularly important are those related to solar bursts. These produce strong irregularities in solar wind and the geomagnetic field (geomagnetic storms and disturbances of various kind), with important effects on the distribution of the charged particles in the ionosphere (ionospheric storms, SID, SWF, etc.).

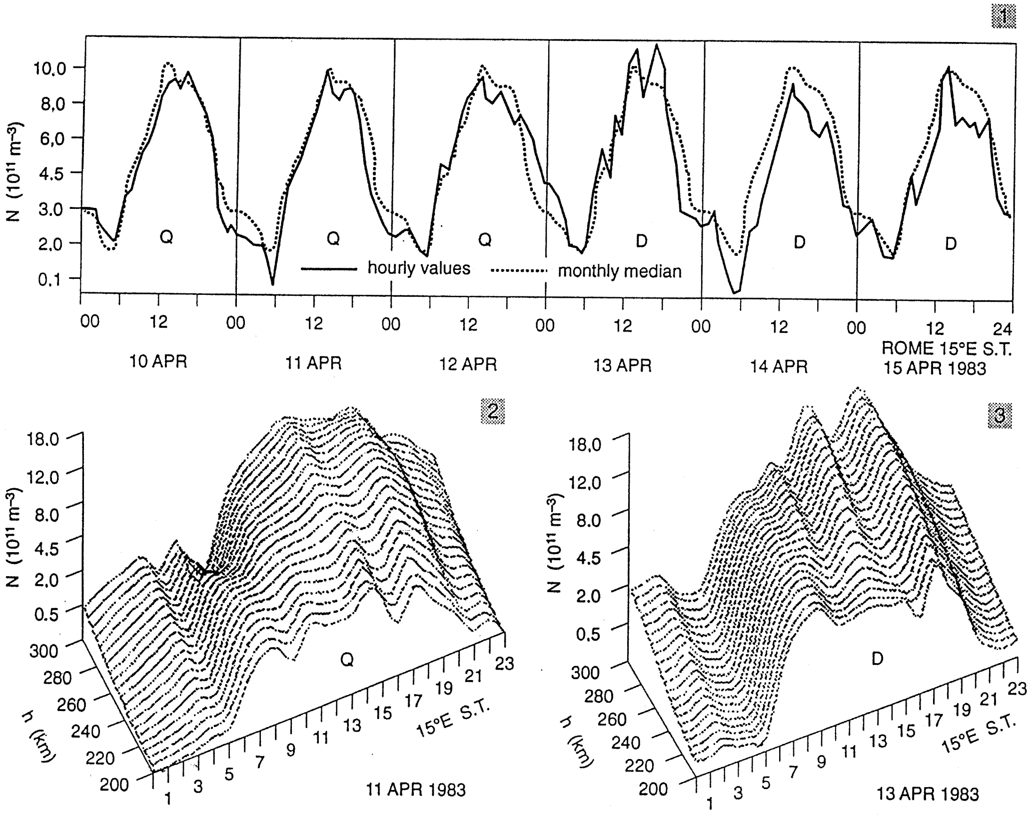

Fig. 8. (1) Behaviour of hourly values of $h'F$ and f_0F_2 at Rome on 6 successive days including magnetically quiet (*Q*) and disturbed (*D*) days. (2 and 3) Constant-height/time profiles of the penetration frequency f_pF_2 during a quiet and a disturbed day. (From P. Dominici, Lj.R. Cander, and B. Zolesi, *Annali di Geofisica*, 40 (5), 1171-1178, 1997).

It is sad to say that, generally speaking, the problems associated with these «anomalies» and «differences» of the real ionosphere in relation to the theoretical scheme of Chapman's theory and to the experimental model represented by the normal ionosphere (or by any of many models at different space scales that have been deduced) have not received attention comparable with their great importance from a geophysical point of view. I have the impression that at that time, *i.e.* in the Fifties and Sixties, apart from some important exceptions, the main efforts were directed to serious morphological research to increase the long- and medium-term accuracy of the just mentioned empirical models: it was common to call all departures from the normal ionosphere simply «disturbances», «fluctuations» and so on.

This trend was in itself a good thing because it largely satisfied a real necessity of short-wave radio communications especially in designing future circuits. However, in its conclusion it was not too productive in physically investigating the above «fluctuations» and devise a more realistic theory of the ionosphere.

Broadly speaking this was the situation of ionospheric physics when, from the Sixties, the rôle of the ionosphere in radio communications became much less important.

The reduced technical interest in ionospheric radio communications and aeronomic revision in the Seventies

In the early Sixties the need to entrust to carrier radio waves new kinds of signals, with frequeny spectra thousands of times wider than that of the earlier telegraphic, telephonic or AM broadcasting ones (few kHz) became more and more important in radio communications: as time progressed, FM broadcasting, multiplexed telephonic channels, TV signals, digital data signals, reaching several MHz in bandwidth. So, the radio carriers were progressively forced to move from the HF (3 ÷ 30 MHz) of short and very short waves to the VHF (30 ÷ 300 MHz) of metric waves and finally to the EHF (300 ÷ 3000 MHz) of microwaves, in various kinds of atmospheric propagation between terrestrial stations and then also with intermediate space stations.

Microwaves are now almost exclusively used in radio applications and the ionosphere has only a marginal importance, *i.e.* as regards the «scintillation» of the signals due to ionospheric small-scale irregularities, on the atmospheric propagation of microwaves (even including the crossing of the ionosphere when satellite stations are used). In fact, the use of short waves by ionospheric propagation is now essentially restricted, apart some from military services, to long distance AM broadcasting and to communications with ships and airplanes in particular conditions. So, the practical importance of the earlier close interaction between ionospheric physics and commercial radio communications has been considerably reduced. For instance, a comparison between the «ionospheric» contents of the IAGA, URSI, IEEE, etc. international meetings some thirty or twenty years ago and in recent years is very indicative in this respect.

In my opinion, the weakening of this interaction has been, all things considered, not completely negative, because ionospheric physicists have gained an improved capability of freely reconsidering their knowledge from a more correct point of view, that is to study ionospheric phenomena as they really are, namely «atmospheric» and not «radio» phenomena.

It soon appeared very clear that the scheme of the electronosphere was in general no longer acceptable.

The radio reflecting ionosphere is really an atmospheric region where a much less dense plasma exists formed by free electrons and positive/negative ions among neutral molecules of the atmospheric gases. The interaction of charged and neutral particles cannot be represented simply by collisional momentum transfer between electrons and molecules, but also by other electron-ion, electron-molecule, ion-molecule, ion-ion collisional and non collisional chemical-physical processes. A particularly important phenomenon is the electron and ion diffusion arising by scattering from such processes, which influences the rates both of apparent production (positive diffusion) and loss (negative diffusion) of electrons and each ionic species of interest, particuarly, as already stated, in the upper part of the *F* region.

This new perspective can be considered the particularization in ionospheric physics of a more general philosophy which pervaded all of geophysics in the mid Seventies, based on the definitive recognition as a whole of the Earth's field, terrestrial phenomena, and related sciences.

To overcome the traditional division of Earth sciences into those regarding the solid Earth (like geology, seismology, etc.), those regarding the terrestrial waters (hydrology, oceanology, etc.) and those regarding the terrestrial atmosphere (meteorology, aerology, physics of the ionosphere, etc.), and to consider terrestrial phenomena together with some aspects of planetary space physics and solar astrophysics, people interested in geophysics began to recognize that some kind of «globalization» was a necessary condition to correctly study many difficult geophysical problems which appeared unsolvable when considered in a less general and narrower field. For instance, only from such a «global» viewpoint is it possible to study the geomagnetic field, which originates from direct strong interactions among the plasma in the Earth's nucleus, the radial thermal gradient in the Earth's interior, terrestrial rotation, the magnetic properties of crustal rocks, the dynamic and electrical properties of the upper atmosphere and ionizing solar radiation.

This new approach in atmospheric physics is known as the *aeronomic point of view*. As regards ionospheric physics, it means critically reviewing our knowledge on the frequent, strong and complicated interactions among the relatively dense neutral atmosphere at, below, and above the ionospheric heights, with its different chemical and ionic (included electronic) populations, the electromagnetic and corpuscular radiation from the Sun, the Earth's magnetic and gravitational field, and the Earth's movements, especially its rotation.

The actual theoretical statement considers the continuity eq. (1) for electrons and each ionic species which can be assumed to be relevant, but including as an agent of production or loss of particles a *diffusion rate* d_i, that is the number of particles of the species i that go inwards towards (positive diffusion) or outwards (negative diffusion) from the unit volume per unit time. Its general expression can be assumed (not considering the sign) as

$$d_i = \nabla(N_i \mathbf{v}_i), \qquad (6)$$

N_i and $\mathbf{v}_i$ being the concentration and the speed of the particles of the species i respectively (the form assumed by (6) in some interesting cases are too complicated to be shown here).

Table III shows a synthetic pattern of the principal particle processes in the permanent ionospheric regions.

Table III. Principal particle processes in the ionosphere.

Process	Species	Scheme	Notes
*Processes of electron/ion production**			
1. Photoionization	+ (e,m^+)	$m + f \rightarrow m^+ + e$	If by solar photons, great de-
2. *e* photodetachment from m^-	+ (e)	$m^- + f \rightarrow m + e$	pendence from site and time
3. *e* detachment from m^-	+ (e)	$A^- + B \rightarrow AB + e$	High material density
4. *e* radiative capture	+ (m^-)	$m + e \rightarrow m^- + f$	
5. *e* non radiative capture	+ (m^-)	$m + m' + e \rightarrow m^- + m'$	*High material density*
cosmic rays			Height 50÷90 km
6. Ionization by meteors	+ $(e,m^\pm)$	Complex scheme	Height ≈ 100 km
solar wind			High latitude
*Processes of electron/ion loss**			
7. Radiative recombination	– (e,m^+)	Inverse of 1	
8. Dissoc. radiative recombin.	– (e,m^+)	$AB^+ + e \rightarrow A + B + f$	
9. *e* capture	– (m^-)	Inverse of 2, 3	Atmospheric luminescences
10. *e* detachment from m_i^-	– (m^-)	Inverse of 4, 5	

Symbols: A, B, atomic species; *e*, electron; *f*, photon; *m*, neutral molecule or atom; m^+, positive ion; m^-, negative ion; m', intermediary molecule; * the diffusion being not considered.

Some relevant themes of current ionospheric research

As previously observed, it was useful in the past to call «anomalies» all the behaviours of ionospheric structures which disagree the consequences of Chapman's theory, as if the experimental facts had to adapt themselves to this theory and not, as it should be, the theory had to adapt to them.

The situation is quite different from the current aeronomic point of view. The question is: a) to have a world-wide measurement organization able to give in a synoptic manner all the relevant data on the chemical-physical state of both neutral and charged particles of primary importance in the upper atmosphere (dynamic conditions, concentration, kinetic temperature, coefficient of production and loss, etc. at each height for each of these relevant species) and then b) to elaborate a suitable theoretical apparatus.

These tasks are undoubtedly difficult, but unlike some other geophysical fields where difficulties of a fundamental character exist, there are less important difficulties, such as those regarding better measurements both in quantity and quality, and better analytical representations.

a) *The measurements: the ignorosphere* – I will not consider here any question concerning the improvements in measurement techniques in progress. I will point out that a basic, but not insurmountable difficulty, consists in the fact that the synoptic measurements mainly regard the electronosphere, while those regarding the ionic species and the neutral atmosphere are very sporadic both in time and space.

Particularly unsatisfactory is the situation of the «electronospheric» measurements themselves regarding both the polar regions and the *D* region, *i.e.* the height interval

50 ÷ 90 km, even though the problems differ. In fact, as to what concerns the polar regions, or *auroral zones*, the question pertains to the above mentioned improvement of the world-wide network of ionospheric observatories. As to what concerns the measurements in the above ionospheric zone the latter is above the greatest height covered by the aerological sounding (≈ 30 km) and below the lowest heights of the low-orbiting satellites (≈ 300 km). On the other hand, as to what concerns radio measurements in such a region, the electron density is so low that normal vertical sounding and many of the other radio methods are very difficult to employ: at LF frequencies a quasi metallic reflection occurs, that gives poor physical information on the reflection zone, and at HF a very high absorption occurs. In point of fact, we could use for this atmospheric zone the term *ignorosphere*.

It is of fundamental importance that a special place has to be reserved to the auroral zones and the cited ignorosphere in the general task of improving the actual measurements. A primary rôle in this improvement should be assumed by the interested international scientific associations, like IAGA and URSI.

b) *The revision of Chapman's theory* – To date, the fundamental theory of the ionosphere still remains the old Chapman theory, although its basic hypotheses may be questionable.

First, the hypothesis that the state and the properties of the ionosphere are well represented by the concentration and events of free electrons is only acceptable as a first, albeit good, approximation in many cases of great practical interest, but cannot be accepted when considering phenomena beyond the radio propagation. For instance, one of the phenomena where the ion content plays a rôle of equal importance with respect to that of free electrons is the so-called *equatorial electrojet*, an electric current flowing in the Earth's equatorial plane at a height of about 120 km (*i.e.* in the *E* region) which produces the regular daily variation of the geomagnetic field. This current is produced by complicated interactions among neutral molecules, free electrons and ions in the ionosphere, vertical thermal gradients in the whole atmosphere, Earth's rotation and geomagnetic field.

As to what concerns the other hypotheses, isothermy appears to be acceptable only in narrow height intervals, while chemical homogeneity, or, more precisely, the prevalence at each height of only one definite chemical species, is less acceptable and the assumption of photoionization and ionic recombination as exclusive electron production and loss processes appears to be too restrictive. Finally, the hypothesis of hydrostatic equilibrium is absolutely unacceptable for the ionosphere, at least for its bottom part, where, as shown hereafter, ample and persistent movements are present.

Let me recall something personal in this respect. Many years ago, in 1954, F. Mariani – a dear friend of mine who then gave up ionospheric heights in favour of magnetospheric ones – and myself attempted to remove Chapman's assumption of chemical homogeneity in order to explain the *F* region splitting and F_2 layer seasonal anomaly. Our suggestion concerned the overlapping of two Chapman's layers from different chemical molecules and with different electron loss processes (ionic recombination and capture). In winter daytime there is a great overlap, the two ionizations accumulate giving an appearance of only one dense layer, while in summertime the lighter molecules separate themselves upwards from the heavier ones and two layers (F_1, F_2) appear (fig. 9.1 and 9.2). As it is well known, about twenty years later a formal and more complete explanation was given, substantially agreeing with the above results. This theory is based on the equilibrium between the ionization processes of N_2 and O_2 molecules according to the different mass and electron decay processes of the result-

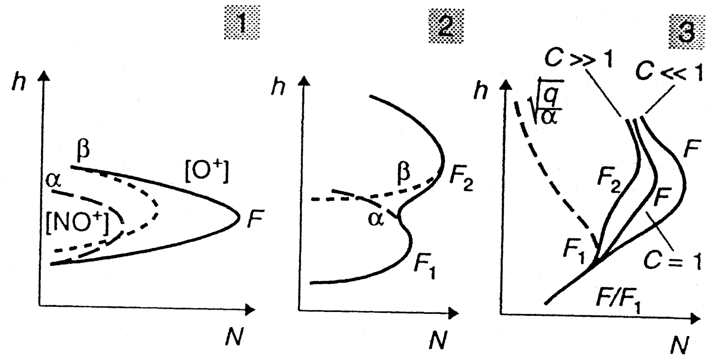

Fig. 9. (1, 2) A morphological model (P. Dominici and F. Mariani, 1954) to explain the seasonal splitting of the *F* region based on a strong (wintertime: 1) or weak (summertime: 2) superposition of two Chapman-like layers from different chemical species and different electron loss (α and β are respectively the recombination and capture coefficients). (3) A subsequent formal model on the same matter (see the text).

ing ions (ionic recombination with a coefficient α where the heavier ions N^+ and NO^+ prevail, capture with a coefficient β where the lighter ions O^+ prevail). As fig. 9.3 shows, the appearance as a one-layer or two-layers structure depends on the value of the parameter $C = \beta^2/(\alpha q)$, q being Chapman's photoproduction function (3).

Returning now to the principal question, in my opinion the efforts that have been made to generalize Chapman's theory starting from suitably revised basic hypotheses, that we can call the *theoretical approach* to the ionospheric phenomenology, have been fairly inconclusive to date if compared with the results obtained by another procedure, that I call the *pragmatic approach*.

c) *The pragmatic approach to ionospheric phenomena* – This approach consists simply in studying the ionospheric appearances as they are in themselves, not considering any previous theoretical model.

As an example, I propose to consider a typical $f_0F_2(t)$ graph, as in fig. 5, or the equivalent $N(t)$ graph. If it is reasonable to assume the attachement to neutral molecules as the only process of electron loss, according to (1) and (5) we can write for each point i of the graph:

$$(\underline{dN_i}/\underline{dt})_i = p_i + \beta_i \underline{N_i} \pm d_i = 0, \tag{7}$$

where the underlining indicates values that are given by the graph itself. In particular, in a point of maximum (M) is $(dN/dt)_M = 0$, $p_M = q_M$, q being Chapman's solar photoionization function (3), the electron diffusion in this condition is reasonably negligible and then $q_M = \beta_M \underline{N_M}$; at this point, the value of q_M or, alternatively, of β_M can be obtained according to which of these two quantities is known by other measurements or considerations. Similarly, in a point of minimum (m) is $(dN/dt)_m = q_m = 0$, $p_m = d_m = \beta_M \underline{N_M}$ (the electron production is due only to positive diffusion). Similar procedures can be applied to ionograms in real heights, $N(h)$, and the results can be integrated with those obtained by $N(t)$ graphs. There have been some studies of this kind, also in recent years.

Let me make another personal notation, in recalling that in the Fifties, and also later, I profitably used methods like these to deduce values of capture (F_2) and recombination (F_1, E) coefficients from ionograms taken during solar total eclipses.

It is for me a surprising fact that this approach, perhaps not too elegant but certainly fruitful, has received relatively poor attention in the past. However, I have the impression that it has experienced a surge of new life when new perspectives recently appeared on some large dynamical phenomena in the ionosphere.

d) *Large-scale ionospheric dynamics* – The existence of ample and more or less regular movements of isoelectronic surfaces in the ionosphere was first recognized and systematically studied the end of the Forties, and later, mainly by continuous recording of ionospheric echoes by fixed-frequency vertical pulse ionosondes. In time, other methods of observation were used, focussing on some regional campaigns performed in recent years, like observations of moving irregularities in fast vertical ionograms, Doppler shift measurements of HF signals and incoherently back-scattered radar signals, TEC measurements, etc. In particular, Doppler techniques have been and are still implemented at ING by C. Bianchi and his collaborators. The data as a whole allowed the growth of the special section of ionospheric physics devoted to so-called *Travelling Ionospheric Disturbances* (TIDs).

In the Eighties these and other similar experimental data were critically analysed and led to the recognition that many of these TIDs had a wavelike nature, being the ionospheric appearance, as transversal waves, of longitudinal atmospheric waves. The latter can be divided into *acoustic waves*, *i.e.* ordinary compressional waves, and *gravity waves*, which are characterized by an almost vertical wave vector, because of the modification impressed by the particular dependence of the sound speed in the atmosphere on the wave surfaces at some distance from the origin. The atmospheric waves are normally indicated as *Acoustic-Gravity Waves* (AGWs) and their ionospheric appearances as *Ionospheric Waves* (IWs).

Figure 10 shows the appearance of local oscillations due to IWs in the daily behaviour of hourly values of f_0F_2 at some European stations and, with greater evidence, in the corresponding «oscillatory component», that is given by the percentage differences $D\%$ (–1, +1) between the actual values and 3-values running mean values.

The wave parameters of IWs (speed, period, wavelenght, direction) can be deduced by means of the usual methods, but it is always necessary to remember that: 1) these waves are greatly attenuated (only three to six cycles are generally recognizable on amplitude/time or amplitude/space representations); 2) because they suffer complex interference phenomena, the general wave patterns are very often rather confused. Some difficulties thus ensue and even the current sophisticated methods of wave analysis are only in part able to remove them. In any case, when speaking about IWs the term «period» must be correctly understood as «quasi-period» of the few clearly recognizable oscillatory cycles on a record of some ionospheric parameters or on a suitable space-time representation of them, or as the «most frequent period» from a Fourier analysis, or an equivalent one, and the term «speed» indicates the horizontal velocity of the wavelike behaviour which can be deduced from the above ionospheric data.

IWs range in period from a few minutes to some hours, in speed from 100 to 1000 m/s and in amplitude (in terms of electron density) up to some 10^{10} m^{-3}.

In the past, they were usually classified in relation to their origin (more precisely, the origin of the primary AGWs) mainly in *tropospheric IWs* (arising from meterological activity, lightnings, earthquakes, volcanic activity, etc.) and *auroral IWs* (from instabilities

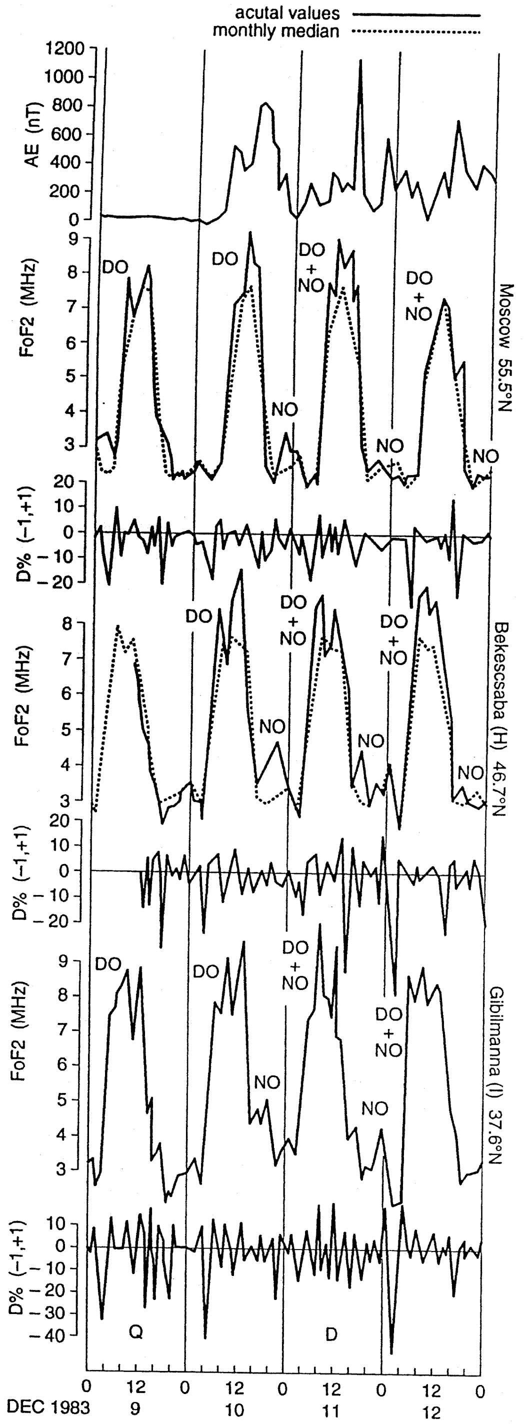

Fig. 10. Hourly values during some days of the auroral electrojet index *AE* and, at some European ionospheric stations, of f_0F_2 and of the percentage difference *D%* (−1, +1) between actual values of f_0F_2 and values from a symmetrical 3-values running mean. Daytime after-sunrise oscillations (*DO*), probably due to ionospheric waves excited by the solar terminator, are always present, even on magnetically quiet days (*Q*), while nighttime oscillations (*NO*) are related to a disturbance of the auroral electrojet (beginning on 10 December) and interfere with daytime oscillations (*DO* + *NO*). The oscillation amplitude seems to increase as the latitude decreases. (Adapted from P. Dominici, B. Zolesi, and Lj.R. Cander, *Physica Scripta*, 37, 516-522, 1988).

of the auroral electrojet, a roughly circular electric current in the polar caps produced by the interaction of the solar wind with the geomagnetic field and Earth's rotation). In recent years, another more useful classification has been introduced, with reference to the wave characteristics, especially the period and the direction of the wave vector. So, IWs are distinguished into: *long-period IWs* (from a few hours to several hours), corresponding to the previous auroral IWs, mainly equatorwards from polar caps with a speed of 100 ÷ 600 m/s and large amplitude; *short-period IWs* (from a few minutes to about an hour), corresponding to the previous tropospheric IWs, omnidirectional with about the speed of sound and small amplitude; *medium-period IWs* (1 ÷ 3 h), roughly equatorwards with speed of 100 ÷ 300 m/s and noticeable amplitude.

The latter medium-period IWs exhibit an almost daily systematicity, that is a large independence on the actual ionospheric conditions, and especially on geomagnetic conditions: this invariance is absent in other kinds of IWs. Figure 11 is very indicative of this feature, and, moreover, shows that IWs are more evident in wintertime than in summertime. A useful circumstance concerning these waves is that the prevalent period is around two hours and hence hourly measurements, like those that have been produced and are still produced by the world-wide network of ionosonde stations, are useful for the purpose.

Let me make another (and last) personal notation. In the last fifteen years I have studied just these waves almost exclusively from the existing hourly data of European ionosonde stations and specific data published by some authors. In the late Eighties I reached the same conclusions as those independently formulated by V.M. Somsikov, who has studied this kind of waves for a longer time than me: the medium-period IWs arise from the sudden thermal transition occurring in the lower atmosphere at the passage of the solar terminator. I use the term *Solar Terminator Ionospheric Wave* (STIW) to signify that it is possible that some medium-period IWs could have another origin. STIWs are well recognizable because, as I have found, their direction is about normal to the surface of the nearest solar terminator, with a consequent typical daily rotation of the direction of the wave vector observed at middle-high latitude stations in the northern hemisphere, namely from SE in the morning to S at local noon and then to SW in the afternoon. This kind of origin also explains their cited seasonal effect.

e) *Ionospheric modelling and predictions* – Among the traditional themes of ionospheric research there is still a noticeable interest in the modelling of the parameters of ionospheric layers. This aims not only to obtain a rational pattern of a complicated structure like the ionosphere, but also to predict the operative parameters in HF ionospheric radio communications, both for the management and design of these communications, and the consequent allocation of future services on the available radio frequencies by the *ad hoc* international and national agencies.

At the beginning, until about the second world war, ionospheric predictions were limited, for evident reasons of a first approximation, to the monthly median values of critical frequency and minimum virtual height of E, F_1, and F_2 layers, from which some rapid procedures allowed the relevant operative parameters to be deduced like maximum usable frequencies, and some empirical criteria overcame difficulties arising from strong ionospheric disturbances.

In time, some *prediction formulae* have been proposed, especially for the characteristics that exhibit a definite regularity, as, generally speaking, those regarding E and F_1. For instance, Bruno Zolesi and I have deduced at ING formulae of such a kind for the monthly median values of E and F_1 critical frequencies in Central Italy around local

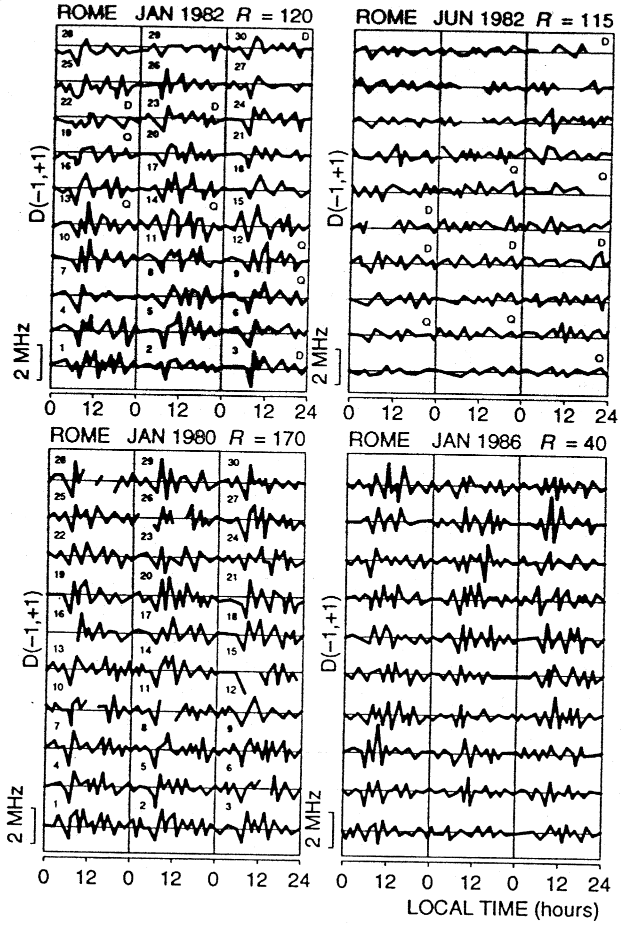

Fig. 11. Local oscillations of $D(-1, +1)$ regarding hourly values of f_0F_2 at Rome during January and June 1982, and, at a different sunspot number R, during January 1980 and 1986. Q and D indicate respectively the international magnetic 5 quiet and 5 disturbed days. The always present daytime oscillations seem to be more ample in wintertime than in summertime, and independent of both geomagnetic and solar activity. (Adapted from P. Dominici, B. Zolesi, and Lj.R. Cander, *Physica Scripta*, 37, 516-522, 1988).

noon from the data obtained at the ionosonde station in Rome during the period 1938-1988:

$$f_0E = 3.18\ [(1 + 8.83\ 10^{-3}\ R) \cos\chi]^{0.222}\ \text{MHz}, \qquad (8)$$

$$f_0F_1 = 4.00\ [(1 + 1.36\ 10^{-2}\ R) \cos\chi]^{0.196}\ \text{MHz}, \qquad (9)$$

χ being, as already said, the zenith distance of the Sun, variable with the time of year and the hour of day, and R the sunspot number.

For the F/F_2 characteristics, that are the most important from the point of view of both geophysicists and radio communicators, this numerical modelling is not easy and it is necessary to use graphical or nomographic models or computerized numerical models.

Ionospheric prediction in terms of monthly median values rapidly showed its limitations when it became more and more important to predict parameters that should be valid at the moment, *i.e.* not for the fictious normal ionosphere but for the real, particularly if perturbed, ionosphere.

During the last twenty years, some methods have been proposed and used to realize this much more useful kind of prediction. In the last ten years the European Union has promoted and sponsored two important research projects in this direction. The first, PRIME (Prediction Retrospective Mapping over Europe) led by Peter Bradley, is devoted to devising models and software programs to apply in the European region. The second, IITS (Improved quality of service in Ionospheric Telecommunication System planning and operation), led by Rudi Hanbaba, regards the possible extension of the prediction area to the high latitude and near eastern regions around Europe.

Some of such methods consist in introducing in the previous «median» predictions some corrective coefficients deduced from suitable actual or near actual measurements, like a contemporary or very recent vertical or oblique ionogram. This philosophy rapidly reached its final stage, that is the instantaneous measurement of the optimum working radio frequency and the mimimum transmitted power for a given signal-to-noise ratio at the receiving terminal, obtained by a bistatic oblique sounder diplexed on the radio circuit of interest, together with a kind of panoramic receiver able to show the degree of occupancy of all the radio channels in the band of interest. Such a method, even expensive, obviously solves in the best manner the problem of optimizing the management of radio circuits but, equally obviously, removes the concept of prediction itself. Nevertheless the prediction has an intrinsic great importance from a scientific point of view. In fact, it is a formidable means to validate the models.

So, returning to the ionospheric predictions themselves, a conclusion coming from the experienced facts is that it should be convenient to adapt to the ionospheric field the technique of modern meteorological forecasting. That means devising a mathematical model formed by the differential equations expressing the continuity of the concentrations of free electrons and relevant ionic species at each height in the ionosphere and by the equations giving the corresponding values of measurable quantities (critical frequencies and so on), entering into it by the measured value of these quantities at sounding stations of a suitable regional or world-wide network.

As to what regards this network, the situation appears to be more favourable than in meteorology because it is common knowledge that the data from a certain ionospheric sounding station should be generally significant for the ionosphere across a zone until several tens of kilometers around. The actual network of vertical sounding stations will be sufficient if it is integrated by some other similar stations covering some still uncovered geographical areas and by some oblique sounders covering the oceans. To realize that appears to be only a question of international cooperation and, moreover, the rôle of the cited international scientific associations will be of primary importance.

On the contrary, the situation is much less satisfactory as regards the above mathematical model. I shall limit myself to indicating only three problems, which are of a very different nature.

The first problem concerns the still poor knowledge we have on *ionospheric variability*. We know that it arises in part from the general atmospheric variability and in part from the short-term variability of the ionizing solar radiation both electromagnetic and corpuscular (solar wind). Among the «atmospheric induced» variabilities we only know something about the wavelike irregularities related to the cited atmospheric-gravity waves. About «solar induced» ionospheric variability, I have the impression that its morphological connection to the variability of solar data has not yet completely appeared. Moreover, it cannot be excluded that ionospheric variability could also have other, even minor, origins. It is a happy circumstance that, as can be deduced from the abstracts of recent ionospheric meetings, much attention is now given to these questions.

The second problem regards the structure of the equations in the model, and precisely whether they can have the usual linear forms or if it is necessary to consider, at least in some cases, non linear forms in order to give the correct importance to the nonlinear phenomena that have been recognized in the ionosphere. The current opinion is that this complication can be considered only in relatively few cases, but there are no conclusive studies at this regard.

The third problem fortunately has not a current issue and in fact it projects itself in an indefinite, even perhaps near, future. It deals with the same problem which recently appeared in meteorological forecasting, where it limits to about 7 days the time domain of reliable predictions. The question is that in meteorological phenomena the ultimate protagonists are the free molecules and atoms of the atmospheric gases, whose dynamics has a chaotic nature, though it produces macroscopic structures (clouds, currents, eddies, etc.), whose dynamics has, on the contrary, a largely deterministic nature. In short, the result is that the development of these phenomena can be, in principle, highly dependent on their initial conditions, that is usually called *deterministic chaos*. The situation in the ionosphere is similar, with the difference that there are also or exclusively (in the top ionosphere) charged particles (free electrons and ions), it being necessary therefore to consider, as additional interactions, the electrodynamic ones with the geomagnetic field. Moreover, the particle concentration is much less than in the meteorological field and the importance of the chaotic component is proportionally greater. In both cases, to define in such terms a more realistic mathematical model to be used for forecasting appears nowdays a question still concerning that part of mathematical analysis dealing with the deterministic chaos.

Conclusions

The flowing in my mind of so many memories on the evolution of ionospheric knowledge, to which I have participated only in small part but always as an interested witness, has made me write at length. I apologize to all those who have been so patient in following me up to now.

The main sensation I have at this final instant is still one of pure enthusiasm – the rational translation of what at the beginning I called «love» – in comparison with a scientific discipline still so fascinating and rich in things to be investigated. Over and over I felt – and the same occurs now – I was hovering in the cool and dark majesty of the hundreds of kilometers of the ionosperic heights observing what happens around me and, at the same time, being delighted by the wonderful vision offered by the blue and variegated beauty of our old mother Earth: isn't this beautiful? I consider myself really lucky, because the first steps of my life have brought me to this point.

But now, after this fine rest, it is time to return to my road. I wish that even its last part is pleasant because of the esteem and honour shown to me by so many colleagues and other people!