Space Research in Upper Atmospheric Physics for the Americas

Ronald F.Woodman
Instituto Geofísico del Perú

Introduction

During the first years of the space age, right after the launching of the Sputnik, Space Research was synonymous of Upper Atmospheric Research. Man, for the first time, was capable of placing instruments in an environment which was not reachable before. This was also true, a few decades before, when the first high altitude rockets where being developed and tested in the U.S. In fact, most of the unfired V2 rockets which were confiscated from Germany, after the second world war, were modified and used to carry instruments to study the upper atmosphere.

Nowadays, space research and space related activities are much broader in scope, with satellite communications, remote sensing and interplanetary exploration taking the largest portion. Nevertheless, upper atmospheric research continues to play and an important role within all of the space activities. It is, and we believe it will be of special importance in the space activities of the less developed nations of the third world, mainly because of the cost and level of the technologies involved.

Upper atmospheric research has always been justified from, both, an academic and practical point of view. Academically, as in other fields of science, we feel compelled to study and understand the world in which we live. It is part of human kind's superior nature. We could never be satisfied with observing the beauty of the northern auroral lights without understanding the natural mechanisms that produce them. From a practical point of view, for many years we were using the reflecting properties of the ionosphere for our radio communications, with demanding needs to predict its variable behaviour. This, we could not do, unless we had a clear understanding of the physical processes involved in its formation and variability. Its practical importance has increased with the use of space for important applications, like satellite communications and remote sensing. The upper atmosphere is the environment where these satellites have to operate, including often their human crew, and we have to understand its influence and potential hazards.

I have been asked to speak about Ionospheric Physics as a part of the sciences related to space. I have extended myself to include other aspects of the more general field of Upper Atmospheric Physics. In either case, the field is much to extensive to be included in a single lecture. At universities, the field is covered in many full year graduate courses and requires many textbooks. I have to reduce the scope and talk only about a few topics. To do so, I have used the criteria of what would fit best the goals of the present conference, the geographical characteristics of the countries involved, their economical and technological development, the broad spectrum of interest and background of the audience and the limitations of my own field of knowledge.

I will limit the scope of my talk to some important upper atmospheric problems peculiar to the equatorial and subtropical latitudes –where most of the participating countries are located– and to the potential of different remote sensing techniques to the study of relevant problems. It should be mentioned that space research is not limited to the use of satellite or rockets for their study. Ground base remote sensing techniques are equally powerful, and at times closer to the technical and economical possibilities of our countries. In addition, we have two very large radars in the region, at Arecibo and Jicamarca, which are available to scientists of nationalities other than that were the instruments are located. I will also mention briefly some problems and possibilities of carrying upper atmospheric research with the use of rockets.

I may be the only one to speak about ground bas

not include MA – Ozone – greenhouse

Within the slogan of the conference we read " ...perspectivas de cooperacion para el desarrollo". Motivated by the word "desarrollo" (development), I can not help to fall on the temptation to say a few words and express a few ideas about it,before I get to the main subject of my talk.

Whenever we hear or use the word "development" we think implicitly of "economical development". In fact, I am afraid that many economists and politicians think the two are equivalent. We should not forget that equally important is cultural and social development. Here I am using cultural in its broadest sense, including scientific and technological culture. This requires some elaboration. Many of us think of culture only in its artistic and humanistic context. For instance, we may think that a cultured society is one that can appreciate the beauty of classical music or the paintings of the masters, or that a cultured individual is one erudite in history and literature (If he knows the lineage of Ramses the second, we are really impressed). Without resting importance to these aspects of culture, I personally believe, – and hope that many of you agree–, that the true culture of the twentieth century civilization, includes in large part its scientific and technological knowledge. If we consider ourselves superior and more civilized and developed than our ancestors one century ago, it is because of our vastly superior knowledge in science and technology. If we consider one country more developed than the other, it is because of its development in science and technology, not necessarily because of its higher per capita income. In fact, unconsciously, when we imagine a more civilized society in other worlds of the universe, we imagine it with a much higher scientific and technological development.

– Israel + Kuwait

Of course, economical development is equally important, and for some even more so. But –and, I believe, here there is consensus– this has come as a consequence of scientific and technological development. There is no other way (unless one is lucky enough to seat on top of the largest gold or oil deposits of the world). In addition to this fact, the point I try to make above is, that apart of the implications of science and technology on the economical, we should value it also by the value of culture per se, and of science and technology as part of it. That means that when we appraise the value of a particular scientific project, let it be in space science or other fields, we should not value it only by its direct economical implications; we should also value it by its scientific and cultural merits. We should ask our selves as well, if it will improve the quality of a professor at the university, or the level of preparation of graduate students working on it. In fact, as far as the economical aspects, we should take for granted that this improvement will eventually have, somehow, secondary economical implications to the society which this professor or student is part of.

Space science activities have certain advantages over other scientific activities which should be taken into account, specially in countries like ours. It is very demanding on other important fields of science, like: physics, mathematics electronic instrumentation and digital signal processing, which makes it appealing and stimulating to ambitious students. Furthermore, one can find topics of research, specially in the field of upper atmospheric physics using ground based techniques, which are not very demanding on funds, or on expensive instrumentation. In fact some African and Caribbean universities have taken advantage of this and started graduate schools in physics using ionospheric research topics as fields of research for their graduate students.

Ionospheric research topics of interest

With the above mentioned criteria, we will limit ourselves to discuss three research topics of interest to Latin American countries: The Equatorial Electrojet, Equatorial F–region irregularities and The Appleton Anomaly. There are other topics of interest, for instance those associated with the South Atlantic Magnetic Anomaly an ionospheric Antarctic phenomena, where South American countries could also make important contributions, but which we will not include here for the sake of keeping the scope within reasonable bounds.

The Equatorial Electrojet

At lower ionospheric E–region altitudes, centered above 100 km. of altitudes there exist a system of electric currents generated by the motion of that region of the ionosphere with highest ohmic conductivity, which receives the name of "dynamo currents". As in a dynamo, they are the consequence of the field induced by the motion of a good conducting region of the atmosphere transverse to the earth magnetic field. This system of currents is particularly intensive in a relative narrow belt of latitudes , centered at the magnetic equator, a few hundred kilometers wide at either side. This current belt is known as the "Equatorial Electrojet" and has been the subject of intensive research by developed and underdeveloped countries, ever since its discovery at the Huancayo Observatory, close to 75 years ago.

The concentration comes about as a consequence of the tensorial nature of the conductivities in the ionosphere, the horizontality of the earth magnetic field lines at the magnetic equator and the limited vertical extent of the conducting E–region. The East–West dynamo electric fields drive a vertical (Hall) current which can not go beyond the region of conductivity. As a consequence, electric charges are accumulated on the top and bottom of this region, creating a secondary electric field, which is larger than the original East–West field that originated it. This field, in turn, creates a large (Hall) current in the same direction as the original field.

The presence of this current belt can be detected by the relatively large variations it produces in the earth magnetic field. This was the way it was original discovered and studied. Later, it was discovered that this high current is responsible for the generation of very strong fluctuations in the electron density of the ionosphere at the same altitudes and latitudes. These fluctuations, in turn, are responsible for the scattering of radio waves in the HF and VHF range of the spectrum. Including the anomalous high frequency echoes received by ionosondes, referred to as Sporadic Equatorial E–region (Esq) echoes.

The electron density fluctuations, better known as equatorial E–region irregularities, are capable of scattering at relatively high frequencies in the VHF range. In fact, much of what is known about the nature and dynamics of this irregularities are consequence of studies made with VHF radars. Many of these studies have been performed using the high power transmitters of the Jicamarca Observatory, but the echoes are so strong that they can and have been received by much smaller transmitters and antennas. This makes the irregularities and interesting field of study for the many underdeveloped countries which are under the magnetic equator, since they present very challenging physical problems and they can be observed with relatively inexpensive equipment.

The nature of the research problems presented by these irregularities is not limited to the field of ionospheric physics. The irregularities are generated as a consequence of plasma instabilities. As such, they provide a natural laboratory for the study of this important field of plasma physics. The instabilities are related to similar instabilities associated with plasma confinement and nuclear fusion problems. Their study in the ionosphere has the advantage that the plasma is created –at no cost– by nature and that it does not have the additional complications of the laboratory plasma boundaries.

The altitude of this irregularities makes difficult the use of satellites for its study. But they can and have been observed by rockets. Several rockets have been launched from the equatorial ranges of India, Brasil and Peru.

The potential that the scattering properties of these irregularities have in the field of communications is of special interest. It is possible to communicate using VHF frequencies to points much beyond the horizon. Normally VHF communication is limited to the radio horizon of the transmitter, since they can not be reflected by the ionosphere; but, the scattering properties of the equatorial E–region irregularities permit communications between points slightly beyond 1000 km. apart, provided a region of the electrojet is somewhere in the vicinity of the mid point of the link. Radio contact has been established for several hours during the day, for instance, between Lima and La Paz, and voice communication between Lima–Juliaca and Lima–Iquitos (Peru). In both cases, relatively low power transmitters were used. These possibilities present many research problems to fully asses the communication potential and derive engineering parameters for the design of communication links.

Equatorial F-region irregularities.

Not only the E–region presents electron density irregularities at the magnetic equator, equally dramatic and interesting are the irregularities present at F–region altitudes, between 250 to more than 1000 km. altitude. These irregularities were first observed as spread traces instead of the thin traces normally obtained in the output records of an ionosonde (ionograms). Originally, it was not known what produced the estrange nature of the echoes, and the phenomena was known as "Spread–F", because of the spread nature of the traces. Even though, much has been learnt about its nature they are still referred at times as Spread–F irregularities.

The irregularities are also responsible for the quick fading of the signals received from satellites whenever they transverse a region affected by them. The phenomenon is known as satellite signal scintillation, since it is the radio analogy to the optical scintillation of stars as they transverse a turbulent atmosphere. Satellite signal scintillation is important because:

1. They affect the quality of the signals received from utilitarian satellites, and
2. They are used as remote sensing tool for the observation of the properties of the irregularities that produced them.

The equatorial F–region irregularities are actually deep fluctuations in electron density covering a wide spectrum of sizes. They include fluctuations with scale length of the order of a meter to scale lengths of the order of hundreds of kilometers. The fluctuations are very large. They have magnitudes comparable to its mean density. That is, the density can fluctuate to a very small fraction of its undisturbed value. The largest scales are responsible for the spread echoes in an ionogram and the scintillation of satellite signals. The smaller scales are capable of producing radar echoes at VHF. In fact, most of what is known about the physical mechanisms responsible for their creation has resulted from observations made with the incoherent scatter radar at Jicamarca. When Spread–F conditions are present, the echoes received at Jicamarca are several tens of decibels above the maximum incoherent scatter levels. So much stronger, that they could be received by smaller antennas and transmitters, although little has been done to exploit this possibility. All length scales are very much elongated along the magnetic field lines. For all practical purposes they present the same (fluctuating) density along the field line for a distance much longer than the transverse dimension of the fluctuation.

F–region irregularities cover a much wider range of latitudes than the E–region ones. They can be observed by all the countries in the South American continent. In fact, F–region irregularities exist at all latitudes, but those associated with the same type of equatorial mechanisms are confined within a belt of about ±20 of latitude centered in the Magnetic Equator.

There are many plasma processes involved in the creation of the instabilities, but the primary process responsible for the largest fluctuation length scale is believed to be a Raleigh–Taylor instability. The positive gradient of electron density in an undisturbed ionosphere is unstable for an interchange of magnetic field tubes. The higher tubes weigh more than the lower ones. Therefore, an interchange lowers their potential energy and is unstable, in a way no different than a bucket full of water when placed up side down. The heavier portions of the ionospheric plasma fall down and the lighter ones float up in kilometers size bubbles. The latter have been detected by rockets and satellites. At the sharp contours of these bubbles, other unstable mechanisms take place, creating the smaller scales in a cascading of sizes, all the way down to the smallest meter scale fluctuations.

Although the process responsible for the largest scale is well understood, the mechanisms responsible for the smallest scales are not. Thus, there are many opportunities for research in this field, specially in the theoretical aspects. The problem is of interest not only to ionospheric physicists, but to plasma physicists as well. More observations are required to give the necessary clues about the important physical processes that are taking place at the smaller scales. There are also problems related with its variability from day to day. Why are they present very strongly one day and not the next? The answer to this question requires observations of the background parameters at all the latitudes involved.

It should be mentioned that the irregularities are capable of deteriorating the quality of the radio waves that transverse them. They produce fading in the HF radio waves used for communications which reflect at the same affected altitudes. They also produce fading in the signals coming from satellites, including those used for communications and navigation. They are also capable of distorting the radar signals used for the remote sensing of the earth using radar techniques. These undesirable effects also motivate research in this area because its practical applications.

The Appleton Anomaly

In the absence of ionization transport, one would expect maximum ionization at latitudes were the incident flux of the ionizing radiation of the sun is a maximum. At the equinoxes, for instance, one would expect this maximum to occur at the geographical equator. In reality, this maximum occurs at latitudes approximately ±15° north and south of the magnetic equator. Since it does not conform with above simple minded expectation, it is referred to as an anomaly. It receives the name of Appleton anomaly in honor of Sir Edward Appleton, its discoverer. The "anomaly" is centered around ±15° of latitude but its dynamics involves a much wider latitude region which cover most of the South American and some Caribbean countries.

The excess of ionization at these latitudes is explained in terms of a net transport of ionization from the region of creation to higher latitudes. During the day the dynamo electric fields are such that they produce an upward vertical drift. Therefore, ionization produced .at altitudes of maximum production is first moved at higher altitudes. The ions are allowed to move relatively free along the magnetic field tube but are constrained to stay within the same tube —— drifting vertically. Magnetic field tubes at the equator are curved downwards. A tube, with maximum ionization production at the equator, will curve downward to regions of lower production at low altitudes and higher latitudes (approximately ±15°). But, the high mobility of the ions along the magnetic tube will tend to redistribute the ions to a diffusive equilibrium position, with higher density at lower altitudes, i.e. higher latitudes. The end result is the "anomalous" distribution with higher densities at higher latitudes north and south of the magnetic equator. The mechanism of redistribution is called the fountain effect, since any one particular ion moves up at the center —— as in a fountain—— and then falls down north and south of the equator , were they accumulate.

Optimum frequencies for radio communication depend on the value of ionization densities at the latitudes affected by the anomaly. These in turn depend on the solar activity and on the magnitude of the vertical drift at the equator and the magnitude of the velocity of the neutral winds in the north–south direction. The latter are responsible for any asymmetry in the redistribution of ionization. Solar activity is known from satellite measurements. The vertical drifts and north–south winds are measured at he Jicamarca Observatory. This capability makes the South American continent specially suited to do research in this area. Theoretical modelling would have the necessary inputs, and the results could be checked with observations in the region.

Apart of the importance the equatorial anomaly has in the communications field that we have mentioned above, it has also importance in understanding the variability in the occurrence of F–region irregularities. Both phenomena are closely coupled, and present challenging problems to investigate.

Ground Based Observations of the Upper Atmosphere

As we mentioned in the introduction, ground based techniques, for the study and observation of the upper atmosphere, have special significance in developing countries like ours. It is certainly less expensive than satellites and rockets, and involves technologies which are easily within our possibilities. It is also possible, of course, to design and construct rocket and payloads to be carried by rocket or satellites launched and maintained by more developed countries. Some South American countries have even the potential to launch high altitude rockets of their own manufacturing. But, I will not elaborate on this here, and keep my self within the (arbitrary) bounds mentioned before.

One of the simplest instruments capable of remote sensing the ionosphere are satellite beacon receivers. They consist mainly of an antenna, a receiver and a recorder to measure the amplitude of the signal received. They can be used to record satellite scintillations. Many papers have been produced using such simple instruments and made important contributions to the subject of F–region irregularities. With slightly more complicated equipment, one can perform polarization measurements or multiple frequency propagation studies which allow the measuring of total ionospheric electron content, including large scale structure, like F–region bubbles or the equatorial anomaly. Two spaced antennas permit the observation of irregularity size distribution and horizontal drifts.

A classical instrument for the study of the ionosphere is the ionosonde. It is effectively a HF radar with about one hundred watts of power which measures the delay, as a function of frequency, of the echoes reflected from the ionosphere. It permits the monitoring of the electron density as a function of altitude and time. During the IGY in 1957, many of these instruments were deployed in many underdeveloped countries permitting their participation in the program. Unfortunately as they have aged and become obsolete, they have not been replaced and the activity has faded away. They can be obtained commercially at a few thousands of dollars, or build using personal computers at a lower cost and as an interesting instrumental development project at a university.

A powerful –and jet conceptually simple instruments, which can be put together at electronic laboratories of many of our universities– are airglow detectors, usually built around a Fabry–Perot interferometer. The instrument is capable of measuring the intensity and the spectral position and shape of the weak light emitted by the ionospheric chemical reactions associated with the ionization recombination. The intensity depends on the density of background parameters. Sky sweeps are capable of detecting structure (Large scale F–region irregularities, equatorial anomaly, etc.). The spectral shape and position gives information about temperature, winds and ionization drifts.

Finally ——although these is not an exhaustive list—— we have a range of radars going from low power backscatter radars to the gigantic incoherent scatter radars, like the one at Jicamarca and Arecibo. Among the small ones we have the so called ST radars (for Stratosphere– troposphere radars) or wind profilers, which at equatorial latitudes can also be used to observe E–region and F–region irregularities. They are a widely recognized tool to observe the dynamics of the upper troposphere and lower stratosphere, by scattering from turbulent fluctuations in index of refraction, used as tracers of the motions at this altitudes. The altitudes include the Ozone layer, whose importance as a field of research is discussed at other presentation within this conference.The ionospheric irregularities are much more efficient scatters. The echoes received from them can be used to study the dynamics of the irregularities proper as well as the dynamics of the background in which they are imbibed. We next have the MST radars, with one order of magnitude bigger antennas and transmitting power levels, capable of receiving echoes from turbulent electron density fluctuations at mesospheric altitudes. Although their construction is relatively ambitious its technology is not . We should remember it was developed in Peru. The costs involved are higher but perhaps less expensive than a sequence of a balloon or rocket launchings with comparable potential.

An incoherent scatter radar belongs to the several million dollar instrument category. On the other hand, there are two already in existence in the region which are opened to scientific proposals from outside. Incoherent scatter radars are among the most powerful tools to study the upper atmosphere. They can obtain echoes from the minute fluctuations in electron density, caused by the discrete nature of matter. They are capable of characterizing the temporal behaviour of these fluctuations, and from it determining the state parameter values of the ionospheric plasma. Parameters that have been measured include : electron density, ion density of the main constituents, electron and ion temperatures, ion drift and hence the intensity of the electric field driving it , the velocity of the ions along the direction of the magnetic field, the population of suprathermal (foto) electrons, neutral densities and the direction of the magnetic field. All of these parameters can be measured as a function of time and altitude, up to one to several thousand kilometers of altitude. Of course, they are also capable of obtaining echoes from turbulent neutral density fluctuations as a sensitive MST radar, as well as from ionospheric irregularities.

The first incoherent scatter instruments were built during the first years of the 1960 decade, yet we are witnessing the design and construction of new installations, with not much difference in technology, except for the much higher capacity in digital control and processing. I would not be surprising if soon we will see the most developed Latin American countries follow building its own.

Relevant International Scientific Programs

I should close by mentioning that there are many scientific programs which include among their goals the research topics that we have mentioned here. We have the International Equatorial Electrojet Year (IEEY) proposed within IAGA by the Inter–Divisional–Commission on Developing Countries (IDCDC), and many of the working groups of the Solar Terrestrial Energy Program (STEP) coordinated by the Scientific Committee for Solar Terrestrial Physics (SCOSTEP), a very ambitious program for the 1990 decade. We also have many of current US initiated programs which are going to be incorporated within STEP like CEDAR and SUNDIAL. Any future collaborative space program within the Americas, should take into account the objectives of these programs, they should take advantage of their existence and the support they have already received from the participating countries and scientific bodies.

8/3/90