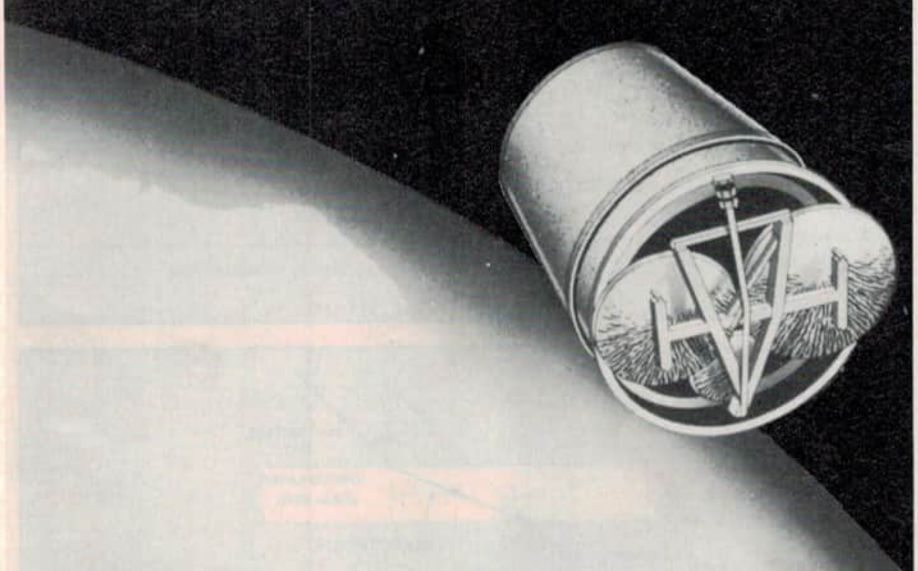
Last year the USA proposed the use of satellites as platforms for shortwave sound broadcasting transmitters. These could relieve the congestion and interference now experienced by ground-based international shortwave broadcasters. This article outlines the requirements and possible performance of such a system.

SATELLITES AND SHORTWAVE



John Kennewell
Learmonth Solar Observatory, Ionospheric Prediction Service

EVEN A CASUAL shortwave listener soon comes to realise that all is not well with sound broadcasting in the high frequency bands between 3-30 MHz. Tuning across the dial from 60 to 13 metres reveals many frequency channels that are occupied by more than one transmitter. The result is frequently an unintelligible cacophony of sound which destroys the utility of the channels for all. Although most of the transmissions attempt to convey meaning through modulation in one of the world's many languages, there is the occasional emission which deliberately aims to destroy the intelligibility of signals upon which it is imposed.

There are around 150 nations in the world today. Many of these have shortwave transmitting facilities. The high frequency band, through ionospheric reflections, offers an economical and accessible way of conveying to the rest of the world a country's unique philosophy on life. Whether it be an espousal of politics or religion, an attempt to teach a language or simple medical facts, ethnic entertainment, or merely a desire to present a view of world news compatible with one's ideology, HF broadcasting, given the right conditions, is an ideal medium for use by even the smallest nations. Unfortunately, conditions are not always ideal.

The ionosphere, upon which all long distance terrestrial transmissions depend, is a function of many variables. Its reflective properties depend on geographical position, time of day, season, transmission frequency, and the periodic and capricious nature of its sustainer, the Sun. To cope with all these variables, many national administrations operate multiple services, with several transmitters spread not only over adjacent bands, but also within the one band. Only in this manner can they be relatively assured of reasonable coverage of their desired target areas.

Unfortunately, there are simply just not enough HF broadcasting channels to accommodate this practice without significant mutual interference. This is particularly true during the years of low solar output; a time when the maximum useable frequencies (MUFs) on all transmissions paths are reduced.

To overcome some of these problems, what may be regarded as a rather radical area has recently been reconsidered. In a draft report to the CCIR (International Radio Consultative Committee) the USA last year proposed the use of satellites as platforms for shortwave sound broadcasting transmitters.

The idea is not entirely new, as feasibility studies of this concept were made in the early 1970s. However, at that stage it was decided that power requirements alone were too great to warrant further consideration. (It is generally regarded that a signal of 500 to 1000 μV at the input of a small portable receiver is required to give a good quality broadcast signal in a noisy urban environment. Such a signal level can only be achieved by a geosynchronous satellite transmitter with a power output of around one quarter of a megawatt!)

If the power requirements can be met there are some definite advantages (and also some disadvantages) to be had from a space-based shortwave transmitter. For it is the very frequencies that are useless for ground originated, ionospherically reflected signals that can propagate from space, through the ionosphere, to the ground. Consideration of exactly what these frequencies are involves both an interesting physical and geometrical problem, and also helps to quantify the discussion into actual system parameters.

Upon reflection

When a swept frequency signal is directed from the ground vertically toward the ionosphere there is found to be a critical frequency, f_c, which separates the behaviour of the signal. Frequencies below f_c are reflected by the layer in question, whereas frequencies above f_c pass essentially undeviated through the layer. If the signal is not vertically incident upon the ionospheric layer, the maximum frequency, f, that is then reflected is given by

$$f = \frac{f_c}{\cos i}$$

where i is the angle that the signal makes with the normal to the layer (Figure 1).

In actuality, the reflection of an oblique radio wave by the ionosphere is a gradual process; see Figure 2(a). As the wave penetrates into the ionised layer, it encounters a greater density of electrons which cause the wave to be progressively bent or refracted.

If the wave frequency is not too high this refraction leads to the equivalent of reflection. If, on the other hand, the frequency is too high to allow reflection, it will travel past the level of maximum ionization and be refracted back onto a path parallel to the original incident wave. This is shown for an idealised homogeneous ionospheric layer in Figure 2(b).

Note that the transmitted ray is parallel to, but slightly displaced from, the projection of the incident ray. This displacement increases with the thickness of the refracting layer and also as the wave frequency ▶

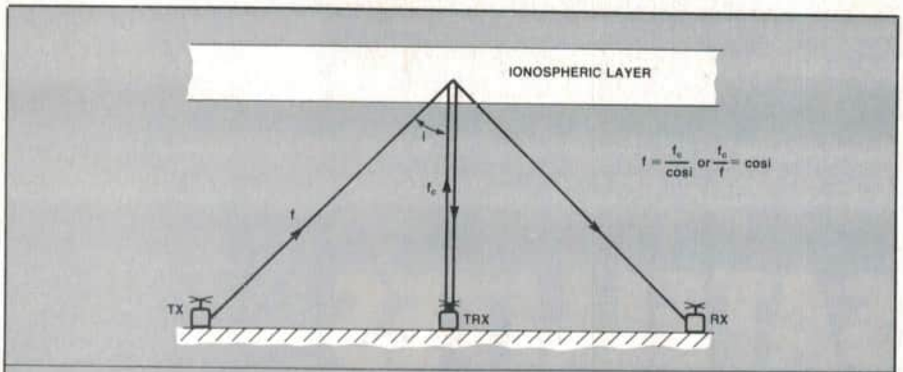


Figure 1. Relationship between an ionospheric layer critical frequency, f_c (vertical incidence), and the maximum frequency, f, that may be transmitted at an oblique incidence angle, i.

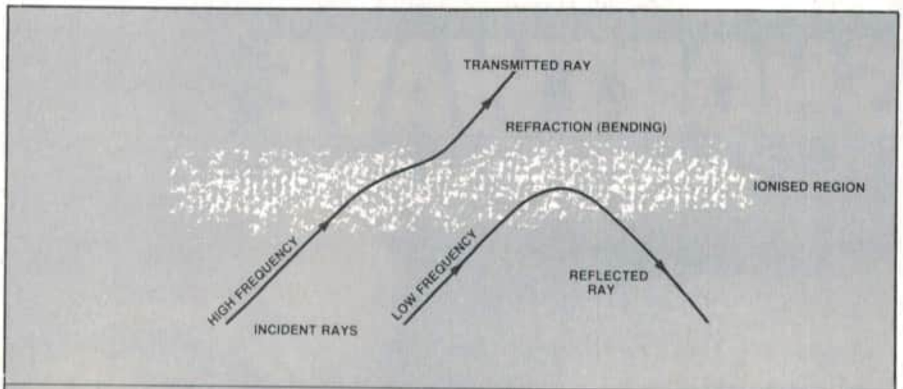


Figure 2a. Process of refraction within an ionospheric layer may lead to reflection (low frequencies) or transmission (high frequencies).

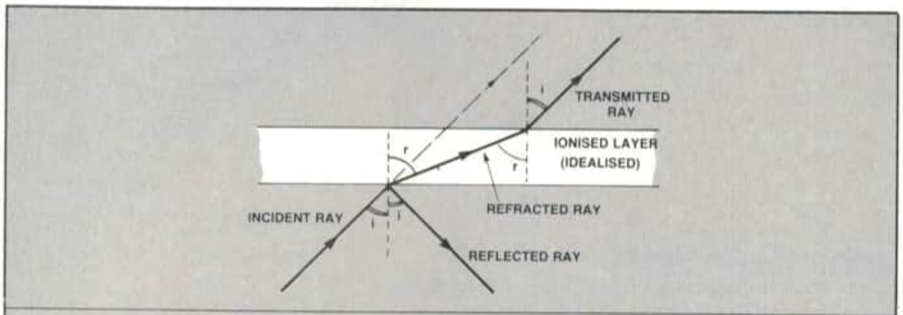


Figure 2b. An idealisation of the process showing the angles involved. Note that the transmitted ray is parallel to the incident ray, but displaced a small amount.

approaches the oblique critical frequency. It is assumed to be negligible (and in practice is negligible for high orbit satellites) in the following discussion.

The sky's the limit

In applying the above points to a signal propagating from a satellite to a receiver on the ground we simply invoke the principle of reciprocity. That is, if the transmitted frequency is too low, it will be reflected by the ionosphere back into space. The problem reduces to finding the minimum frequency, for any given path, above which the ionosphere wil appear transparent to the satellite signal. All discussions will consider only satellites located well above the ionosphere.

From the time of the first Sputnik, the Soviets have been using frequencies near 20 MHz, and more lately 15 MHz, for telemetry purposes. These satellites, however, usually orbit *within* the ionosphere and their signals are subject to ducting which can result in reception at their antipodal points. This phenomenon is quite complex to analyse, is not very predictable, and thus will not be considered here.

The geometry relating to the satellite-to-ground propagation path is shown in Figure 3. The transmitted signal strikes the ionospheric layer at point I with an incidence angle of i. It is this angle together with the critical layer frequency, f_c, at this point that will determine whether or not the satellite signal, f, will pass through to the ground station. Unfortunately, the angle i is not something that is readily available for computation. What is required is an expression relating the minimum transmittable frequency, f, to the distance 'a' that the ground station, E, lies (measured along the ground) from the satellite sub-point, P.

With the aid of some trigonometry this can be done to produce a satellite HF propagation equation. The equation, including its derivation (for those interested) is given in the accompanying box. To be useful, however, the implications of this equation need to be presented visually.

Before looking at some tabular results from the satellite HF propagation equation, it is useful to consider the portion of the earth's surface that can be seen by satellites at different heights. This is shown in Figure 4. The limit of satellite optical visibility is given both as an angle measure along a great circle path, and a ground distance along the same path. The actual distance appropriate to radio frequencies will be somewhat larger due to refraction effects by the Earth's surface. This can be generally simulated by using a value for the Earth's radius, R, that is 4/3 the actual value of 6370 km.

Some results

Using the values of Figure 4, the satellite equation may be used to compute minimum propagable frequencies within this area. This is done in Table 1 opposite.

All computations were done assuming that the principal ionospheric refracting layer was 300 km above the surface. This layer is the F2 layer in the daytime, and the F layer at night. It is interesting to note that although the lower altitude satellites have a more limited ground coverage, the lowest frequency that can be employed at the edges of the area serviced is 3.4 times the ionospheric critical frequency at the point of penetration. This number is independent of satellite height. It is obvious, of course, that a wider range of frequencies may be used by the higher satellites to service stations within the coverage area.

An alternative presentation of data is shown in Figure 5. This shows the minimum useable frequencies for a satellite in geosynchronous orbit, assuming a uniform ionospheric layer at 300 km with a critical frequency of 10 MHz across the hemisphere. This is, of course, not realistic but it serves as a first basis for discussion. In actuality the critical frequency is determined in a good part by the angle that the Sun's rays make with the point in question.

The critical frequency is thus highest for the point directly below the Sun and falls off on either side longitude-wise, and also with latitude. Ten MHz is a good approximate value for the maximum critical frequencies experienced at times of moderate sunspot number (e.g: $R=50$), rising to around 15 MHz or higher when the sunspot number is high (e.g: $R=150$). These values will occur directly at the sub-solar point.

If, at a particular time, this was to be coincident with the sub-satellite point, the contours of Figure 5 would be correct for the central part of the graph, but would be more open at greater angular distances. This change, reflecting the diminished critical frequencies away from the sub-solar point, would also mean that the last contour (around the limit of the visibility circle) would also be at a lower frequency (e.g: 20 MHz).

P is the satellite sub-point, the point on the Earth's surface immediately below the satellite
S is the satellite transmitting the signal (at height h)
E is the ground receiver
I is the point at which the signal passes through the principal ionospheric refracting layer (at height d)
R is the Earth's radius
The minimum frequency that can be propagated from S to E is given by $f = f_c \cos i$ where f_c is the ionospheric critical frequency at point I. Any frequency lower than this will be reflected back into space.

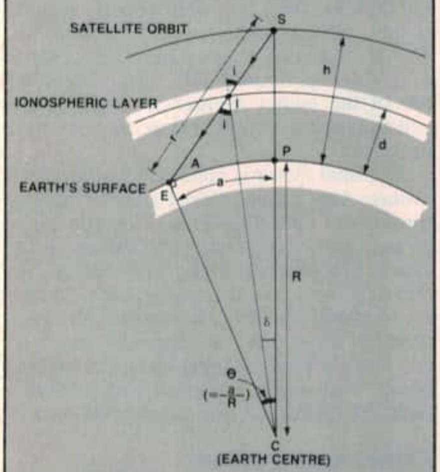


Figure 3. Geometry for signal propagation from a satellite to the ground.

SATELLITE HF PROPAGATION EQUATION

The geometry of the satellite propagation path is shown in Figure 3. The problem is to find the minimum frequency that will propagate from the satellite at **S** to a ground receiver at **E**, a distance **a** from the sub-satellite point **P**. The ionospheric critical frequency at **I** is f_c.

In triangle **CSI** the rule of sines yields $\frac{R+h}{\sin i} = \frac{R+d}{\sin \varnothing}$

Similarly, in triangle **CSE** we have $\frac{R}{\sin \varnothing} = \frac{r}{\sin \Theta}$

The above two equations combine to eliminate $\varnothing$ giving $r = \frac{R(R+h)\sin\Theta}{(R+d)\sin i}$

The cosine rule in **CSE** also gives $r^2 = (R+h)^2 + R^2 - 2R(R+h)\cos\Theta$

Equating the two expressions for r, the satellite to receiver distance, gives

$$\frac{\sin^2\Theta}{\sin^2 i} = \left(\frac{R+d}{R}\right)^2 + \left(\frac{R+d}{R+h}\right)^2 - \frac{2(R+d)^2}{R(R+h)}\cos\Theta$$

The incidence angle **i** is related to the frequencies f and f_c by the relation

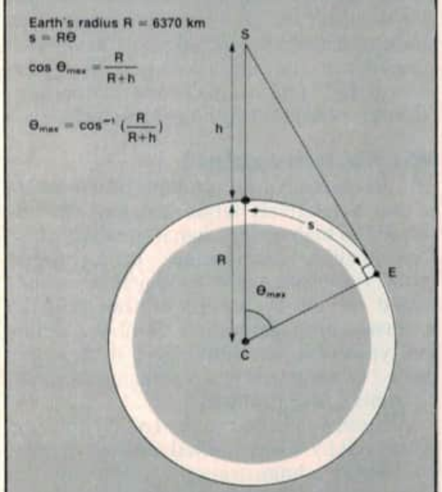
$$\sin^2 i = 1 - \cos^2 i = 1 - f_c^2/f_2$$

Thus the above equation can be rewritten:

$$\frac{f_c}{f} = \sqrt{1 - \frac{\sin^2\Theta}{\left(\frac{R+d}{R}\right)^2 \left\{1 + \left(\frac{R}{R+h}\right)^2 - \frac{2R}{R+h}\cos\Theta\right\}}}$$

Now $\Theta = a/R$ and if we let $\alpha = \frac{R}{R+h}$ and $\beta = \frac{R+d}{R}$ we have the final form of the satellite HF propagation equation:

$$\frac{f_c}{f} = \sqrt{1 - \frac{\sin^2(a/R)}{\beta^2(1+\alpha^2) - 2\alpha\beta^2\cos(a/R)}}$$



h(km)	Θ_{max}	s(km)
1000	30°	3200
2000	40°	4100
5000	56°	5300
10 000	67°	5900
18 000	75°	6150
36 000	81°	6300

Figure 4. The maximum portion of the earth's surface that can be seen by a satellite of altitude 'h' km.

Θ° \ h	1000 km	2000 km	5000 km	10 000 km	18 000 km	36 000 km
0	1.0	1.0	1.0	1.0	1.0	1.0
10	1.6	1.2	1.1	1.0	1.0	1.0
20	2.8	1.9	1.3	1.2	1.1	1.1
30	3.4	2.8	1.7	1.4	1.3	1.2
40		3.4	2.4	1.8	1.5	1.4
50			3.2	2.4	1.9	1.7
56			3.4	2.8	2.3	1.9
60				3.1	2.6	2.2
67				3.4	3.1	2.6
70					3.3	2.8
75					3.4	3.2
80						3.4
81						3.4

Limit of visibility

TABLE 1. f/f_c values for various satellite heights h (km) and various ground ranges Θ°. (Θ is the angular distance from the receiver to the sub-satellite point). The ionospheric layer is assumed to be at 300 km.

If the sub-solar point was to the left of the diagram, the central left contours would be unchanged but the ones to the right would be spread further apart. An egg-shaped pattern would result with the 'fat' end of the egg on the right or eastern side of the hemisphere.

Factors affecting

Apart from shielding of the satellite signal by the F-layer there are three other factors to be considered. The first one of these is the phenomenon of sporadic-E. At certain times of the day during the summer months, patchy intense regions of ionization form at around the 100 km level. This sporadic-E can occasionally have such a high critical frequency that, in effect, blankets the F-layer from ground signals (Figure 6).

A satellite signal that passed through the F-layer might very infrequently be blocked by such ionization. Because of the irregularity of this phenomenon, it has been estimated that, in the worst case, a 26 MHz signal transmitted from geosynchronous ►

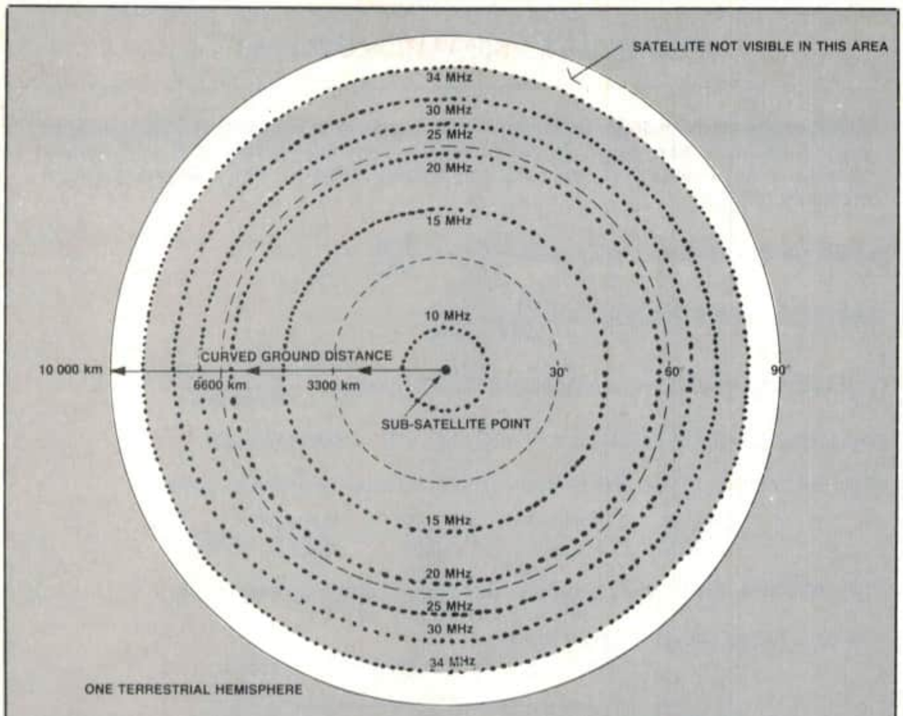


Figure 5. Contours showing the minimum frequencies that will propagate from a geosynchrous satellite (h = 36 000 km) to a ground station. The ionosphere is assumed to be homogeneous with a critical frequency of 10 MHz (this would certainly not be the case in actuality).

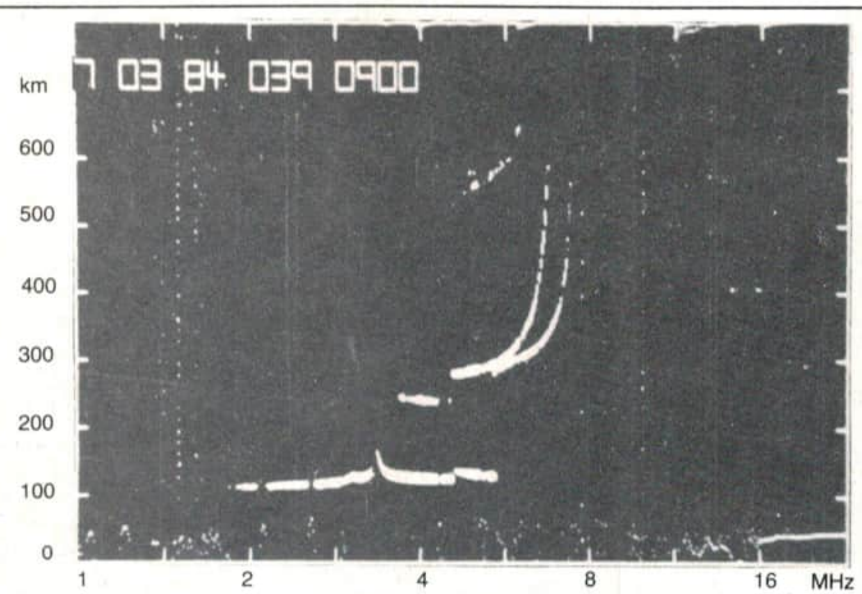


Figure 6. An ionogram showing 'blanketing' type sporadic-E. The higher F region echo at 300 km is obscured by the lower sporadic-E layer at about 5 MHz.

Operating frequency	26 MHz
Modulation type	AM (A3)
Transmitter power	16 kW
Antenna gain	40 dB
Antenna size (diameter)	500 metres
Antenna beamwidth	2°
Signal received by a receiver with a simple dipole or whip antenna	500-1000 μV
Received power	−78 dBW
Receiver noise environment	−122 dBW
Signal-to-noise ratio	44 dB

Figure 7. USA proposal specifications for a geosynchronous sound HF broadcasting system.

orbit should be shielded for no more than 5% of the time. This would only apply at certain times of the year in selected geographical locations.

The second factor to be considered is shortwave fade due to solar flares. This normally occurs to a signal propagating from a ground-based transmitter via the ionosphere. When a large solar flare occurs, the increased X-ray emission from the sun causes intense ionization of the D-region (50-80 km). This can then absorb any HF signal attempting passage through it. Exactly the same phenomenon will occur to a satellite generated HF signal, although use of the higher frequencies i.e: 26 MHz will reduce the extent and frequency of the problem.

The final factor is concerned with polarisation. Any plane-polarised signal passing through an ionised layer will have its plane of polarisation rotated. This rotation can amount to several hundred revolutions for the signals we are discussing, and changes continuously. If the antennas used by both satellite transmitter and ground receivers are linearly polarised significant fading will occur. The simple solution is to employ a circularly polarised transmitting antenna.

Proposed system

Figure 7 gives details of the satellite HF broadcasting system which has been proposed by the USA. It is significant to note that it still requires 16 kW of transmitting power. The required antenna gain also limits (because of its directionality) the area that can be serviced. The massive size of the antenna, one-half a kilometre in diameter, would appear to be not a feasible proposition to construct. However, new lightweight space constuction techniques, developed primarily in connection with potential space solar power stations, have now made it possible to consider such large structures.

Within the realm?

The advantage of HF satellite broadcasting is that a relatively large area may be serviced by a high quality signal, relatively free from fading. Interference from ground based (ionospherically propagated) signals would be nil by nature of the mutually exclusive propagation modes of the different types of signal (although a strategically placed satellite, using a similar orbit, could certainly cause problems).

Although the concept could only be employed by a very limited number of present world administrations, these are the same nations currently emit the greatest number of HF transmissions. Translation of these transmissions to higher, and currently little used, HF broadcast bands would do much to alleviate current shortwave congestion.

Studies conducted show that the 26 MHz band should be useable for satellite broadcasting at most times. When solar output is low, the useable bands might extend through 21 and 17 MHz to as low at 15 MHz. One may well ask why the 88-108 MHz FM broadcast band has not been suggested for a similar type of satellite broadcast service. Although there are some associated technical problems, I believe the answer lies more in the political arena. ●