THE EXTRAORDINARY WAVE MODE: NEGLECTED IN CURRENT PRACTICAL LITERATURE FOR HF NVIS COMMUNICATIONS

M. C. Walden G0IJZ*

*181 Coleridge Road, Cambridge, CB1 3PW, United Kingdom. e-mail: marcuswalden@ieee.org

Keywords: Extraordinary wave, HF NVIS communications.

Abstract

Current practical literature for HF NVIS communication places significant emphasis on *foF2* as being the maximum frequency for vertical propagation. This, however, fails to consider the extraordinary wave. This paper presents the analysis of 5 MHz beacon data, showing the relevance of the extraordinary wave in the MUF calculation for NVIS propagation. The results are in full agreement with established scientific theory and ionospheric propagation prediction methods, the detail of which the HF NVIS user community may not be aware of.

1 Introduction

Near-vertical incidence skywave (NVIS) propagation allows HF ionospheric communication over relatively short distances, typically up to about 400-500 km, using frequencies in the range 2-10 MHz. This technique is of relevance to military and humanitarian organisations, as well as amateur radio operators, particularly during emergency situations where the normal power and communications infrastructure may have failed. This technique primarily makes use of waves transmitted at high angles from the ground, such that terrain obstructions (e.g. mountains) have little or no influence on signal strengths. However, appropriate choice of operating frequency is important for effective NVIS communication [3].

NVIS propagation is predominantly via the F2 region and, therefore, knowledge of the maximum useable frequency (MUF) supported by this region is crucial for effective operation. Current practical literature (and much more that is available on the internet) places emphasis on *foF2* (the ordinary wave or o-wave critical frequency) as the maximum practical operating frequency that does not penetrate the ionosphere at vertical incidence [4,10]. Unfortunately, this guideline fails to recognise the existence of the extraordinary wave (x-wave). Additionally, the optimum working frequency (OWF) or frequency of optimum traffic (FOT) is generally defined as being about 85% of the MUF, which encourages operation at lower frequencies still [2].

This paper presents the analysis of automatic beacon data at 5 MHz showing that, for short NVIS links, the x-wave critical frequency *fxF2* defines the maximum operating frequency at vertical incidence. The significance of this is a broader choice of frequencies available to the HF NVIS user, particularly at mid-latitude locations and during sunspot minima when critical frequencies tend to be low. This paper is aimed at HF users who are located primarily in mid-latitude locations (defined roughly as being between geomagnetic latitudes 30° and 60°, north or south). Historical aspects, prediction methods, absorption and polarisation issues are also discussed, with consideration of other regions.

2 The 5 MHz Experiment

Since 2002, the UK MoD and Ofcom have allowed UK radio amateurs access to a number of 5 MHz channels in order to undertake propagation experiments. The Radio Society of Great Britain (RSGB) launched ‘The 5 MHz Experiment’ with the aim of encouraging radio amateurs to undertake propagation and antenna studies. The 5 MHz band is important for NVIS propagation at mid-latitudes, particularly during a solar-cycle minimum when the MUF fails to reach 7 MHz during the day and absorption might be high at 3.5 MHz. As part of the 5 MHz Experiment, a network of three beacons has been established in the UK. These are located at Chilton, Westmorland and the Orkney Isles (callsigns GB3RAL, GB3WES and GB3ORK respectively). Each transmitter is frequency- and time-locked to GPS, transmitting in turn every 15 minutes on 5.290 MHz. The conducted power is 10 W and the antennas are either horizontal or inverted-V dipoles [12,13].

3 Analysis of Automatic Beacon Data

This analysis considers the reception of the Chilton beacon GB3RAL at amateur radio stations G3WKL and G4ZFQ. Table 1 lists the location, great-circle distance and bearing from GB3RAL to each of them.

Station	Location	Distance	Bearing
G3WKL	Newport Pagnell	70 km	35°
G4ZFQ	Cowes	92 km	180°

Table 1: Receiving station details relative to GB3RAL

The receiving stations use relatively simple direct-conversion (zero-IF) receivers that do not employ automatic gain control

(AGC). The low-frequency output is sampled by a PC soundcard and signal processing is performed using software that measures the peak beacon signal and average noise levels at the soundcard. If no beacon signals are received, then the peak noise level is measured. Although the receiving equipment is not calibrated, the relative signal levels at each station are consistent and provide the basis for this analysis. The received beacon signal levels are correlated with the critical frequencies (as measured by the Chilton ionosonde) that have been adjusted for oblique propagation using the secant law:

$$f_{obl} = f_{vert}\sec(\varphi) = \frac{f_{vert}}{\cos(\varphi)}, \quad (1)$$

where f_{obl} is the maximum frequency supported for oblique propagation at an angle of incidence φ for a maximum vertical (critical) frequency f_{vert}. The secant law assumes a plane earth and ionosphere and can also be written as:

$$f_{obl} = f_{vert}\sqrt{1+\left(\frac{d}{2h'}\right)^2}, \quad (2)$$

where d is the ground distance between transmitter and receiver and h' is the virtual height of the ionosphere at the point of reflection [2]. For simplicity, the height of peak electron density *hmF2* has been used instead of h', since the virtual heights for the o-waves and x-waves are, in general, different. The resulting error in f_{obl} increases rapidly for lower reflection heights and larger ground distances. However, for the shorter NVIS ranges considered here (d < 100 km) and for propagation via the F2 region, this error is small. Also, the ionogram parameter *fxI* is initially assumed to equal the x-wave critical frequency *fxF2*, although it actually refers to the maximum F-region reflection frequency. It will be shown later that this appears to be reasonable for the Chilton data.

3.1 Signal Variation Against Chilton Critical Frequencies

Figure 1 shows the variation in signal-to-noise ratio of GB3RAL received by G3WKL over 24 hours on the 4th September 2007, as well as the Chilton *foF2* and *fxI* adjusted using the secant law. At this low point of the solar cycle, *foF2*sec(φ) fails to exceed the beacon frequency for the majority of the day, yet the beacon is received at relatively strong signal levels, which contradicts the generally accepted practical guideline that *foF2* defines the maximum vertical operating frequency. By contrast, there is much better correlation with *fxI*sec(φ). Between about 1600-1700 UTC, there is a drop in signal-to-noise ratio when *fxI*sec(φ) falls to or below 5.290 MHz. Figure 2 shows the GB3RAL signal-to-noise ratio against Chilton *fxI*sec(φ) as received by G3WKL, whilst Figure 3 shows the GB3RAL signal-to-noise ratio against *fxI*sec(φ) at G4ZFQ, both during September 2007. The onset of signal reception occurs only when *fxI*sec(φ) exceeds the beacon transmit frequency. There appears to be a secondary step-increase in signal-to-noise ratio in Figure 3 when *fxI*sec(φ) exceeds about 6 MHz. This is approximately 0.7 MHz above the beacon frequency and, therefore, this feature might indicate the reception of both the o- and x-waves at G4ZFQ. The effect is less obvious in Figure 2. The strong signals received before *foF2*sec(φ) exceeds 5.290 MHz and the similarly strong signals afterwards indicate that the x-wave experiences similar losses to the o-wave at 5 MHz. These results show that the x-wave is important to NVIS propagation at mid-latitudes and, therefore, *fxF2* defines the maximum operating frequency at vertical incidence and not *foF2*. This is particularly relevant at sunspot minima when the percentage frequency difference between the o- and x-wave critical frequencies becomes quite significant. For example, Figure 4 shows a Chilton ionogram from 1300 UTC on 1st December 2008, where the measured values for *foF2* and *fxI* were 5.075 MHz and 5.750 MHz respectively. The difference in frequency is equivalent to about 13% of *foF2*.

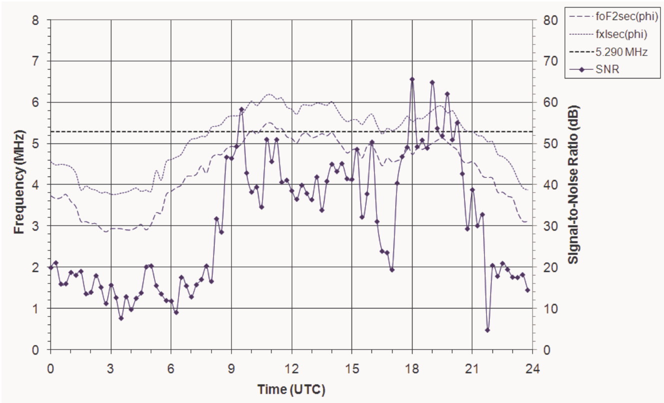

Figure 1: GB3RAL signal-to-noise ratio at G3WKL and Chilton *foF2*sec(φ) and *fxI*sec(φ) on 4th September 2007

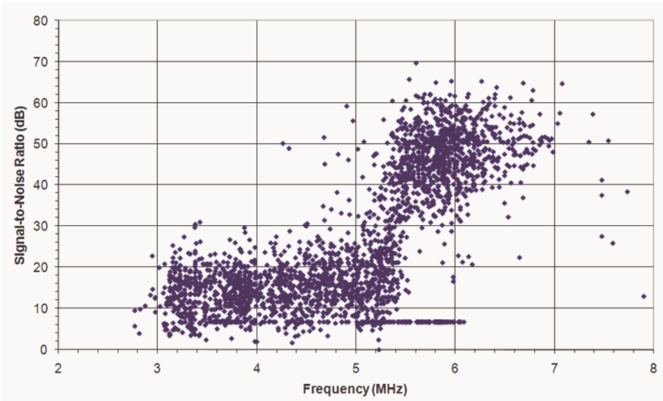

Figure 2: GB3RAL signal-to-noise ratio at G3WKL against Chilton *fxI*sec(φ) during September 2007

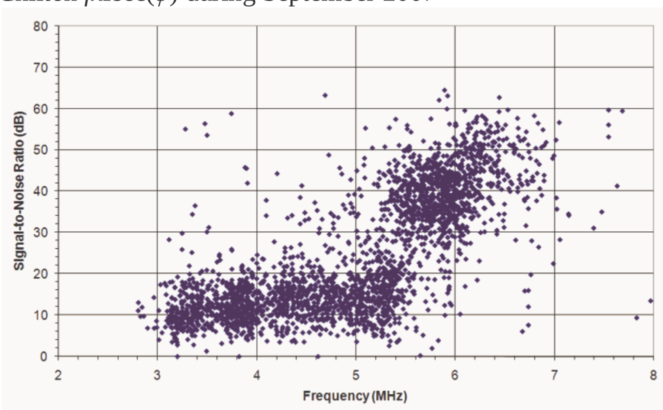

Figure 3: GB3RAL signal-to-noise ratio at G4ZFQ against Chilton *fxI*sec(φ) during September 2007

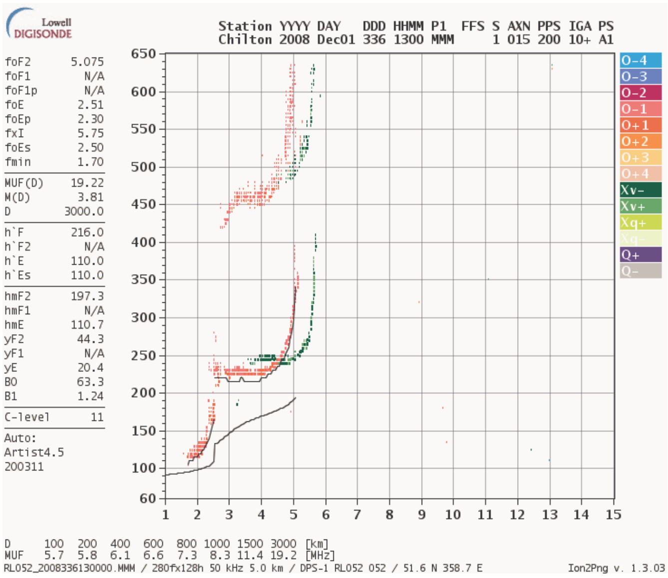

Figure 4: Chilton ionogram at 1300 UTC, 1st December 2008

3.2 Estimation of Electron Gyrofrequency from Chilton Critical Frequencies

The relationship between the o-wave and x-wave critical frequencies is given by:

$$foF2^2 = fxF2^2 - fxF2 f_H, \quad (3)$$

where f_H is the electron gyrofrequency. Equation (3) reduces to the following frequently-used approximation if *foF2* and *fxF2* are much larger than f_H [2]:

$$fxF2 - foF2 \approx \frac{f_H}{2}. \quad (4)$$

In principle, it is therefore possible to estimate the electron gyrofrequency from the measured critical frequencies. Figure 5 shows a histogram of the estimated electron gyrofrequency derived using Equations (3) and (4), with *foF2* and *fxI* (in place of *fxF2*) measured by the Chilton ionosonde during September 2007.

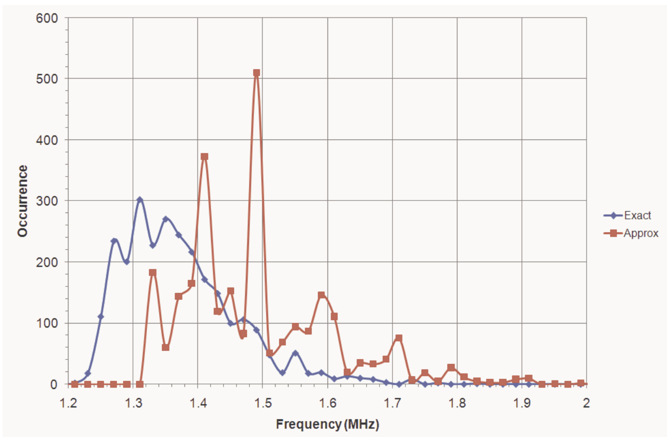

Figure 5: Estimated electron gyrofrequency using measured Chilton *foF2* and *fxI* during September 2007

In general, the use of Equation (3) (exact) shows good agreement with the theoretical electron gyrofrequency over Chilton, which varies from about 1300 kHz at an altitude of 100 km down to about 1190 kHz at 300 km (calculated using the 10th International Geomagnetic Reference Field or IGRF-10). This supports the assumption that *fxI* approximates the x-wave critical frequency *fxF2*. The histogram covers a range of frequencies with numerous peaks, which may be attributed to (i) variation of reflection height, (ii) slight spread F and/or (iii) measurement resolution and errors.

4 Discussion

This section offers further discussion on the relevance of the x-wave for HF NVIS links with the focus on mid-latitude locations.

4.1 Magnetoionic Theory

The Appleton Equation (also known as the Appleton-Hartree or Appleton-Lasson Equation) defines the refractive index of an ionised medium. Together with the magnetoionic polarisation equation, wave propagation through the ionosphere can be characterised [2]. To paraphrase Hunsucker and Hargreaves [5], "it is virtually impossible for an ordinary mortal to make much sense" of the Appleton equation and the magnetoionic polarisation equation "in their full glory". Indeed, it could be argued that the HF user does not need to. However, knowledge of some salient points may prove beneficial. The equations identify the existence of two characteristic waves in the ionosphere: the o-wave and the x-wave (which are evident in typical ionograms, as seen in Figure 4), whilst Equation (4) indicates that the x-wave critical frequency exceeds that of the o-wave by about half the electron gyrofrequency. Therefore, real-time frequency-planning decisions could be made on the basis of measured *foF2* and knowledge of the local electron gyrofrequency. Figure 6 shows a global map of electron gyrofrequency at an altitude of 300 km (obtained using Proplab Pro software with IGRF-10 model).

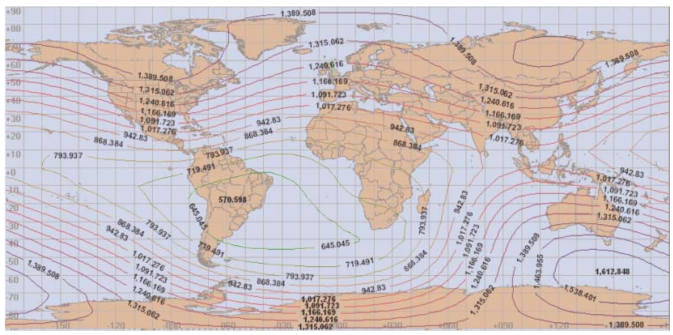
Figure 6: Map of electron gyrofrequency at an altitude of 300 km (Proplab Pro)

4.2 Historical Consideration of the Extraordinary Wave

Sir Edward Appleton was awarded the Nobel Prize in Physics in 1947, for his research into the ionosphere carried out during the 1920s and 1930s [1]. During the Second World War, HF usage intensified for short and medium range operational purposes and, consequently, a number of organisations developed prediction methods, making allowance for the higher MUF due to the x-wave. The French Service de Prévision Ionosphérique Militaire (SPIM) simply added 0.7 MHz to the o-wave MUF (at the time, the electron gyrofrequency was considered to be about 1.4 MHz), whereas the US Central Radio Propagation Laboratory, National Bureau of Standards (CRPL, NBS) applied nomograms directly to the x-wave component [11]. In the UK, the Interservices Ionospheric Bureau (ISIB) applied a correction to the MUF calculation that was dependent on the operating frequency and the MUF factor *M*. For low frequencies and short distances, 0.7 MHz was added [14].

4.3 Propagation Prediction Software

Propagation prediction software enables the HF user to select optimum operating frequencies without the need to use complicated charts and nomograms. Unfortunately, the automation of prediction methods may have led to the existence of the x-wave being obscured from the user. Example software includes REC533 [6], VOACAP [7] and ASAPS, all of which can calculate the MUF for a given link at a given time. REC533 is an ITU method that considers the x-wave in the basic MUF calculation by adding half the electron gyrofrequency for zero ground distance. The contribution reduces for increasing distances and the variation of electron gyrofrequency with location has also been accounted for. ASAPS (based on REC533) uses fxF2 for NVIS MUF calculations. VOACAP is the result of many years of scientific and engineering development within the US Government, dating back to work by the NBS during the Second World War [7]. As mentioned, the MUF calculation was based on the x-wave component [11]. This may not be obvious to the HF user when planning optimum frequency usage for HF NVIS links using software-based methods.

The term MUF actually relates to the monthly median maximum operating frequency at a given time and therefore effective operation can occur 'above the MUF' on some days of the month. The FOT/OWF is a 'safe' frequency that should usually work, although it may not necessarily be optimum, but leads to operation on lower frequencies, when, in fact, higher frequencies could also be viable. This should be considered when frequency scan lists are determined for automatic link establishment (ALE) systems [8]. The assumption that *foF2* equates to the MUF together with the FOT guideline limits the choice of frequencies available to the HF operator, particularly at sunspot minima.

4.4 Absorption

Within the ionosphere, the collision of electrons with neutral molecules and ionised particles results in absorption or loss of energy of the radio wave. Two types of absorption can be defined: non-deviative and deviative absorption. For non-deviative absorption, the x-wave experiences greater absorption than the o-wave, particularly if the operating frequency is close to the electron gyrofrequency [2]. This effect is clearly observed in ionograms, where substantial absorption of the x-wave occurs below about 4 MHz during the daytime and only the o-wave return is seen (see Figure 4). During the night-time, when D-region absorption is reduced, the x-wave may be observed above about 3 MHz. In general, the o-wave is present from the lowest ionosonde frequency (about 1.6 MHz). Deviative absorption occurs when the operating frequency (at vertical incidence) is close to the critical frequency. This effect can be seen in Figure 2 and Figure 3 as a ramp-up (albeit rapid), rather than a step change in signal strength as $fxI\sec(\varphi)$ exceeds the beacon frequency.

Although excessive absorption at lower frequencies renders the x-wave ineffective, above about 3-4 MHz, the x-wave can provide signals comparable in strength to those of the o-wave, as is evident in Figure 2 and Figure 3 and, therefore, its inclusion in MUF calculations/predictions for vertical and NVIS propagation is justified.

4.5 Wave Polarisation at Vertical Incidence

This section considers vertical incidence, although it should be noted that wave polarisation depends on geomagnetic latitudes and angles of incidence. As a wave enters the ionosphere, it is separated into the two characteristic waves, each following different paths. At mid-latitude locations, these waves are elliptically polarised with opposite senses, being highly elliptical at medium frequencies (< 2 MHz) and almost circularly polarised at higher frequencies (the wave polarisations are circular at a magnetic dip pole). For vertical propagation at mid-latitudes, the power is divided approximately evenly between each wave [2]. Assuming operation at frequencies where absorption is similar for the o- and x-waves, then received power levels will be comparable. Polarisation fading might also be present if the circularly-polarised waves are received by a linearly-polarised antenna.

At the magnetic dip equator, the two waves are linearly polarised. An antenna generating an electric field that is north-south aligned only excites an o-wave. Conversely, if the transmitted electric field is aligned east-west, then only the x-wave is excited. Both waves are excited equally if the electric field is at 45° to the dip equator [2]. During the Vietnam War, the US military used antennas oriented in a magnetic north-south direction for greater received levels [3]. This doctrine results in excitation of the o-wave only, which, the author speculates, was used to minimise the risk to communications when operating on frequencies where the x-wave could experience significant absorption losses. At or close to the magnetic dip equator, critical frequencies are generally much higher than those at mid-latitude locations, such that the frequency difference between the o- and x-waves is less significant.

4.6 Noise Levels

At HF, external noise sources (atmospheric, galactic and man-made) tend to limit system sensitivity rather than receiver noise figure. Generally, noise levels increase as the operating frequency is lowered, although man-made noise may be 'coloured' depending on the source. The FOT guideline may lead to operation on noisier frequencies, particularly at mid-latitude locations during sunspot minima when critical frequencies are low, although it should be noted that path loss reduces with lowering frequency, which might offset the increased noise.

4.7 Longer NVIS Ranges and Beyond

For the longer NVIS links, the angle of incidence can be quite large (e.g. $\varphi = 45°$ when $h' = 250$ km and $d = 500$ km) and are technically not near-to-vertical. As the angle of incidence becomes more oblique, propagation via the E and F1 regions may occur, depending on location, critical frequencies and region heights (even for links within the 400-500 km NVIS range). The low MUF at mid-latitude locations during sunspot minima combined with the FOT guideline leads to operation at frequencies where these other propagation modes are more likely to occur. Under these circumstances, signals tend to be received at lower elevation angles where typical NVIS antennas exhibit reduced gain and, therefore, it would be preferable to use higher frequencies (if available) to ensure propagation via the F2 region. Antennas suited to NVIS propagation tend to be used at low heights; for example, resonant dipoles or loops less than a quarter wavelength above ground. Maximum radiation is at high angles, which tends to make NVIS antennas less effective for long-distance ionospheric propagation.

Whilst Equations (3) and (4) give the difference between the o-wave and x-wave critical frequencies, for waves reflected from the same equivalent height at oblique incidence, the difference is dependent on the direction of propagation. For greater angles of incidence and, therefore, the higher operating frequencies supported (compared with vertical incidence), the frequency difference can be almost negligible for east-west (or west-east) directions, whilst at the magnetic dip equator, the difference can be approximately equal to the electron gyrofrequency for north-south (or south-north) paths [2]. For the prediction of long-distance paths (non-NVIS), the day-to-day variation of the predicted maximum operating frequency at a given time is far greater than the difference between the o-wave and x-wave maximum frequencies, such that there is little benefit to the predictions by considering the x-wave [9].

Concerning wave polarisation at oblique incidence, the analysis is somewhat more complex than for vertical incidence. The resultant coupling of energy to the receive antenna depends on both transmit and receive antenna polarisations, path directions and locations, which may even give rise to non-reciprocal propagation under certain conditions [2]. However, this topic is beyond the scope of this paper and the reader is encouraged to consult the more comprehensive ionospheric physics literature.

5 Summary

This paper has presented an analysis of 5 MHz beacon data and shows that the x-wave critical frequency *fxF2* (in general, the ionogram parameter *fxI* can be used at mid-latitude locations) defines the MUF for vertical propagation and not the o-wave critical frequency *foF2*, as is documented in some practical literature for HF NVIS communications. These results are in agreement with established ionospheric theory and prediction methods, which, however, may not be obvious to the HF user community. This paper has also discussed other considerations important to NVIS communications.

Acknowledgements

Peter Martinez (G3PLX) is gratefully acknowledged for his beacon-monitoring software and for making the author aware of the x-wave. John Gould (G3WKL) and Alan Reeves (G4ZFQ) are thanked for providing real-time beacon data and allowing it to be used in this paper. The RSGB and the 5 MHz Working Group are thanked for maintaining the database and allowing it to be readily available for propagation studies. Les Barclay (G3HTF) and George Lane are kindly thanked for providing some historical insight. The MoD and Ofcom are gratefully acknowledged for giving UK radio amateurs access to a number of 5 MHz channels, without which the author's own investigations would not have occurred. The Rutherford-Appleton Laboratory is thanked for allowing the use of the Chilton ionograms and ionosonde data.

References

[1] E. V. Appleton. "The Ionosphere", Nobel Lecture, (1947).
[2] K. Davies. Ionospheric Radio, Peter Peregrinus, (1990).
[3] D. M. Fiedler, E. J. Farmer. Near-Vertical Incidence Skywave Communications: Theory, Techniques and Validation, World Radio, (1996).
[4] J. M. Goodman. HF Communication: Science and Technology, Van Nostrand Reinhold, (1992).
[5] R. D. Hunsucker, J. K. Hargreaves. The High-latitude Ionosphere and its Effects on Radio Propagation, Cambridge University Press, (2002).
[6] ITU-R Recommendation P.533-9. "Method for the prediction of the performance of HF circuits", (2007).
[7] G. Lane. "Signal-to-Noise Predictions Using VOACAP – A User's Guide", Rockwell-Collins, (2001).
[8] G. Lane. "Improved Guidelines for Automatic Link Establishment Operations", Ionospheric Effects Symposium, Alexandria VA, USA (3-5 May 2005).
[9] G. Lane. Private communication, (Sept. 2008).
[10] N. Maslin. HF Communications: A System Approach, Pitman, (1987).
[11] K. Rawer. The Ionosphere: Its Significance for Geophysics and Radio Communications, F. Ungar Pub. Co., (1956).
[12] A. Talbot. "Design and construction of the 5 MHz beacons GB3RAL, GB3WES and GB3ORK – part 1", *Radio Communications*, **81**, pp. 85-87, (June 2005).
[13] A. Talbot. "Design and construction of the 5 MHz beacons GB3RAL, GB3WES and GB3ORK – part 2", *Radio Communications*, **81**, pp. 88-89, (July 2005).
[14] K. W. Tremellen, J. W. Cox. "The Influence of Wave Propagation on the Planning of Short-wave Communication", *Journal of the IEE*, **94**, part IIIA, pp. 200-219, (1947).