Presented at HF Radio 2003 - Ninth International IEE Conference on HF Radio
Systems and Techniques 23 - 26 June 2003, University of Bath, UK

THE ESTIMATION AND MEASUREMENT OF THE EFFICIENCY AND EFFECTIVENESS OF SMALL ANTENNAS IN AN ENVIRONMENT

M.J. Underhill

University of Surrey, U.K.

SUMMARY

There are several possible definitions of antenna efficiency and effectiveness. That there are so many has caused much confusion and misunderstanding. Agreed and standardised definitions are now needed for all of these.

Ground constants and the height of the antenna above ground affect both the antenna pattern and the environmental losses. Antenna efficiency and effectiveness are consequently affected. Thus there cannot be a single measure for either antenna efficiency or antenna effectiveness.

The Chu-Wheeler Q-bandwidth criterion is shown not to apply to all small antennas. It is therefore unscientific to use this criterion as a means for putting a bound on the efficiency of any small antenna.

A new Q-bandwidth measurement method for most small antennas allows all the significant radiation resistance and loss components affecting the antenna input impedance to be separated and measured. The various antenna efficiencies can be determined from measurements made at the antenna terminals.

Small antennas close to ground can, and usually will, launch strong ground-waves. Ground currents can also radiate strongly out to about a wavelength. This is in addition to any ground reflections. The far field antenna pattern is determined by the combination of these effects. Thus any far field efficiency measures of an antenna close to ground can be seriously in error (up to 8-20dB) as a consequence.

Local environmental losses do not in general appear in full in the antenna resistance measured at the antenna terminals. However ground constants can be estimated from antenna resistance measurements assuming sufficient uniformity of the ground. The estimated ground constants can then be used to predict the local losses, and the magnitudes of the ground and surface waves.

A new line-integral potential method is used to give first estimates of some of these effects. Better calibration of these will come both from interpreting existing results and from making further measurements to confirm or contradict past measurements.

ANTENNA EFFICIENCY AND EFFECTIVENESS.

The requirement for any antenna in its designated environment is to have sufficient ‘gain’ in the required directions of transmission and reception. As a working definition we propose that ‘antenna rf effectiveness’ in a particular environment should be defined to be the same as the ‘antenna gain’ in that environment.

Antenna ‘effectiveness’ can also be related to size; a small antenna can be said to be more ‘effective’ than a large antenna even if the gains are equal in the required direction. ‘Cost effectiveness’ is then an appropriate measure.

The ‘gain’ of an antenna is always its ‘directivity’ times its ‘(total) efficiency’. The ‘directivity’ of an antenna is established entirely by its ‘pattern’ (in its given environment). For a 100% efficient antenna the ‘gain pattern’ and the ‘directivity pattern’ are identical both in shape and magnitude. Hence the ‘(total) efficiency’ of an antenna is the ratio of the ‘gain’ in a given direction to the ‘directivity’ in the same direction. This provides the means for defining and measuring the ‘total efficiency’ of an antenna.

DEFINITIONS OF ANTENNA EFFICIENCY

The fundamental definition of efficiency is “power-out divided by power-in”. The difference between power-out and power-in is what is dissipated as heat in the antenna structure itself or in the local environment.

For the directivity, gain and erp (equivalent radiated power) of an antenna, power not going in the desired direction can conceptually be considered as a loss: thus for (Cassegrain) reflector antennas the term ‘aperture efficiency’ is often used.

‘Reciprocity’ means that the efficiency, antenna pattern, and gain of an antenna on transmit is the same as on receive.

Figure 1 shows that losses occur in different places in a radio transmission system and that is why it is necessary to have more than one definition of antenna efficiency.

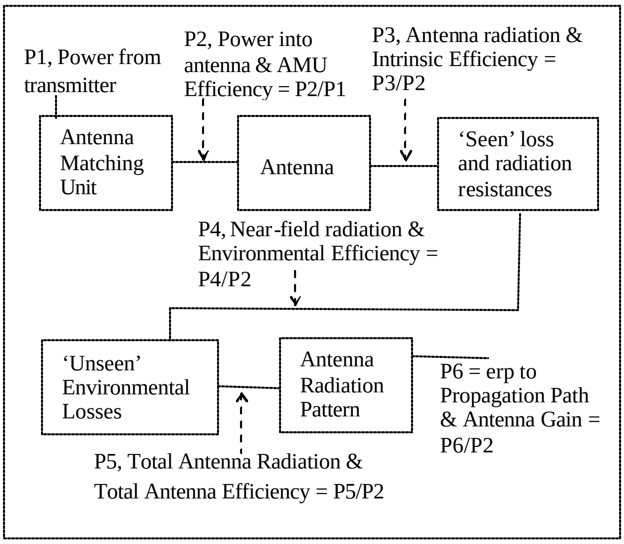

Figure 1 Various losses and antenna efficiencies

As shown in Figure 1, losses can occur in an antenna matching unit (AMU), in the conducting surfaces of the antenna (skin-effect losses), and distributed in the local environment. The local environment losses are divided into a part which is 'seen' at the antenna input and a part that is not. In principle it is possible to derive the 'unseen' part from the 'seen' part if the relationship between the two can be established (as discussed below). Otherwise the 'unseen' part has to be found from the overall antenna efficiency less all the directly measured losses.

The various possible options for antenna efficiency definition can be seen by inspection of Figure 1. The input power P1 from the transmitter is progressively reduced to be P2, P3, P4, and P5. P6 is a power density in a given direction so that P6/P5 is the 'directivity' of the antenna in that direction and P6/P2 is the antenna gain as shown.

From this figure we can see that there are fifteen possible choices of efficiency definition of which ten are antenna efficiency definitions because they include the efficiency P3/P2 for the antenna itself. It is not surprising that there is considerable confusion over the meaning of antenna efficiency. There is a strong need for some agreed standard definitions that are based on what can be measured and validated. The IEEE-Std 145-1993 on antenna efficiency is clearly inadequate and is causing confusion and misunderstanding.

Three very important definitions are 'intrinsic efficiency = P3/P2', 'total antenna efficiency = P5/P2' and 'antenna gain = P6/P2. 'Intrinsic efficiency' is important because it is little affected by the environment and is essentially the efficiency of the antenna in free space. It represents the rf that just escapes the surface of the antenna and has not been dissipated as heat in the antenna conductor surfaces.

THE CHU-WHEELER Q-BANDWIDTH CRITERION

In free space small antennas have high Q, low bandwidth are not in general as efficient as large antennas. But this fact has been greatly exaggerated.

The Chu-Wheeler $Q = (ka)^{-3}$ limit has in the past been used to estimate the best possible Q that a small antenna can have if contained inside a sphere of radius a [1]. Both measurements and new theory confirm that there are several small antennas that contradict the Chu-Wheeler criterion [2 and 3] in some cases by up to two or three orders of magnitude. In fact measurements show that with suitably careful design a small (transmitting) antenna can have an 'intrinsic' efficiency of 80 to 90% or more. This claim is hotly contested by those who believe that the Chu-Wheeler criterion applies to all small antennas or that the present state-of-the-art 'Method of Moments' simulations correctly represent the behaviour of small antennas [4].

We should realise that Chu never claimed that his particular antenna mode expansion applied to every antenna [1]; it was others who did so later [3]. Also scientific logic requires that if one counter-example can be shown, the criterion is invalidated. One example taken from [3] is repeated here. It is based on the existence of the 'dipole mode' of a tuned folded dipole. It invalidates the Chu-Wheeler criterion.

Example- Tuned folded dipole. Fig 2 shows (a) the equivalent circuit of (b) the small tuned folded dipole.

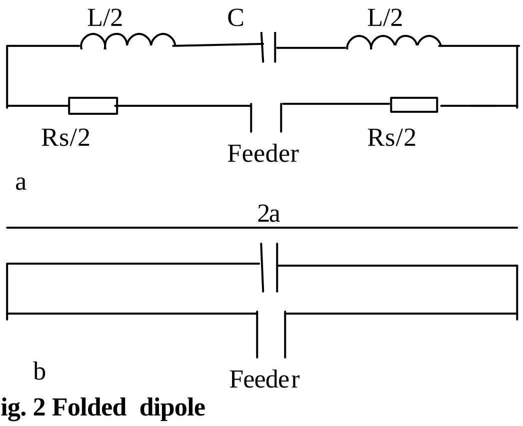

Fig. 2 Folded dipole
a Equivalent circuit
b Small tuned folded dipole.

For convenience, the dipole is taken to have conductor thickness and spacing so that it can be analysed as two pieces of 300 ohm twin transmission line. With sufficient conductor coupling, the dipole mode currents are equal and in phase in the lower and upper conductors so that a four-to-one impedance step-up applies to the small dipole mode radiation resistance. Then with $k = 2p / l$ for small ka we find:

$$X_L = 2\times 300 \tan(ka)\ @\ 600\,ka \qquad (1)$$

$$R_{rad}\ @\ 4\times 20\,(ka)^2 = 80\ (ka)^2 \qquad (2)$$

$$Q_1 = X_L/R_{rad} = 7.5/ka = 1.2\ l/a \qquad (3)$$

The Q factor for this model tends towards $7.5/ka = 1.2\ l/a$. Thus for this ‘practical’ semi-empirical model, based on no new theory, the Chu-Wheeler limit does not apply for antenna sizes where $ka < 1/\sqrt{(7.5)}$.

The Chu-Wheeler limit is therefore emphatically contradicted and as a result cannot be used as a means for putting a bound on the efficiency of any small antenna. In fact it should be replaced by $Q = (ka)^{-1}$, or thereabouts, as shown by the examples given in [3].

THE NEW Q-BANDWIDTH MEASUREMENT TECHNIQUE FOR SMALL ANTENNAS

For a small tuned loop the new Q measurement technique explained in [2] allows the intrinsic loss and radiation resistances together with some of the environmental losses to be determined. The method requires that there is one dominant resonance for the antenna so that the measurement of Q by bandwidth is sufficiently accurate. The method also requires that Q-bandwidth measurements are made over a wide tuning range for the antenna. A wide frequency band gives the best possible separation of the various resistance components that vary with frequency. Those components that vary similarly with frequency can be separated by varying the height, size, orientation or position of the antenna in the environment.

As an example, the method can separately identify the coupling of a loop with its image in the ground. This mode can have a considerably enhanced radiation resistance component when the height of the loop above ground (or the frequency) is increased.

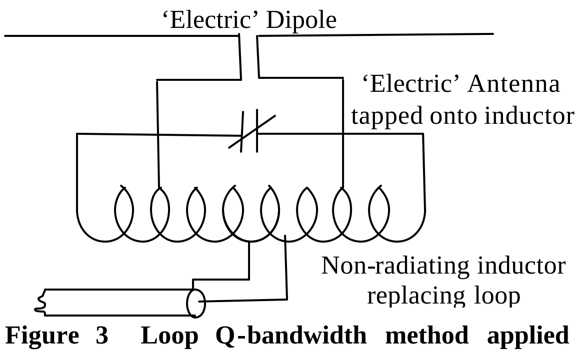

Figure 3 Loop Q-bandwidth method applied to ‘electric’ antenna

What was not shown in the original paper [2] is that the tuned loop method can also be used to measure short ‘electric’ antennas such as dipoles and monopoles. The loop is replaced by a high Q inductor that ideally should not be allowed to radiate. The ‘electric’ antenna is tapped across the inductor at suitable points depending on the impedance of the ‘electric’ antenna over the frequency range of interest. Then the Q-bandwidth measurements can be taken as before over a wide range of frequencies.

Thus the new Q measurement method can give a detailed picture of the various loss and radiation resistance components for practically any small antenna when in place in its operating environment. The new method is sufficiently sensitive to detect and separate near-field and close in environmental losses. It is important to emphasize that the measurements are made at the antenna terminals. The further advantage is that the measurements can estimate ground constants and so further out losses may also be predicted (with a suitable near-field propagation and loss model).

ANTENNA PATTERNS AND PROPAGATION MODES

The ground permittivity and conductivity change an antenna pattern substantially if the antenna is close to the ground. A change in the antenna pattern which does not also introduce loss does not change the efficiency of the antenna, but the directivity and gain will then be changed.

There are three processes that can change the antenna pattern substantially. First, the antenna pattern is changed by ground reflections, as is well known.

Secondly, there is propagation above and along the ground that where the energy is predominantly above the ground surface.

Thirdly, there are ground currents giving a ‘sub-surface wave’ where the energy is predominantly in the ground surface, but can leak out over a distance.

A new proposition investigated below is that out to just beyond the near-field distance of about $\lambda/2\pi$ the ground currents can radiate very substantially. A vertical monopole will ‘inject’ large ground currents that flow in the direction away from the antenna and these will radiate well from the region close to the antenna. A horizontal (small) loop, or dipole, close to ground, will also create substantial ground currents, but these are at right angles to the direction of propagation away from the antenna and radiate less well. A vertical small loop couples less well to the ground and so although the currents are in the favoured direction the ground radiation can be predicted to be smaller.

These various effects are shown in Figure 4. The proposed mechanism for the near-field ground losses is also shown.

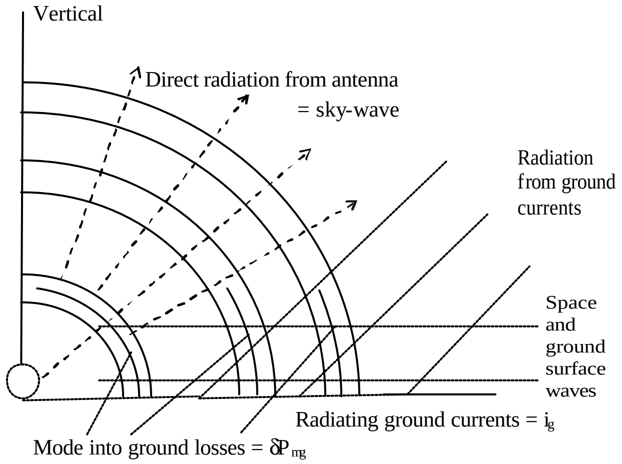

Figure 4 Radiation modes of a small antenna over ground (loop at bottom left)

The assumption is that near the antenna we can analyse the radiation fields as if they were coming from the mouth of a waveguide horn with a vertical horn angle of 90° and a horizontal horn angle of 360°. The horn modes are taken to be significant out to one wavelength or so and thereafter the radiation is regarded as being separated into sky-waves and surface waves.

ANTENNA PATTERNS AND PROPAGATION LOSSES OVER GROUND

For an HF antenna the ground-wave gain is easy to measure gain and far-field efficiency with a properly calibrated field probe and a known input power $P_{in} = P_2$ to the antenna. With an aircraft (and GPS) the sky-wave gain at different elevation angles and in different directions can be found. The total antenna pattern can then be plotted, and integrated to give the total radiated power $P_{tot} = P_5$. From this the overall far-field antenna efficiency can be calculated as $P_{tot}/P_{in}. = P_5/P_2$.

The antenna pattern over ground cannot be assumed. The present state of knowledge about how a real ground affects the antenna pattern is not yet good enough either in theory or in simulations.

It is certainly totally wrong to assume that the antenna pattern over ground is half-isotropic or half a 'doughnut' pattern. It is also wrong to allocate near-field environmental losses to the intrinsic loss of the antenna itself. For an antenna close to ground the 'half-isotropic' assumption and the omission of near-field losses creates serious systematic errors in the far-field antenna gain (efficiency?) method as described in the standard IEEE Std 149-1979.

Subsequent further out ground and surface wave losses and antenna pattern changes from the interaction of the ground wave, surface wave, and the sky-wave, are regarded as propagation effects that are not part of the overall antenna efficiency, but can be considered as contributory factors to antenna 'effectiveness'. As an example Figure 5 shows F.M. Kabbary's ground wave results from a broadcast CFA(Crossed-Field Antenna) in Egypt [5].

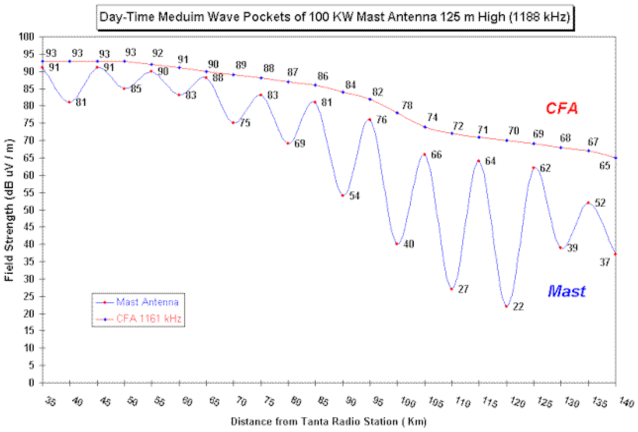

Figure 5. Comparison of ground-waves from CFA and 125m mast.

Here we can see that the much smaller (9m height) CFA has about 4dB gain over the larger (125m height) conventional antenna. Furthermore we can see that the conventional antenna generates an unfortunate interference pattern that can be explained if the ground wave travels about 2.5% slower than the 'space-wave' over the ground. Alternatively if the space-wave is 'trapped' between the ionospheric D-layer and the ground (as a waveguide mode) as Kabbary suggests, it will travel slower than the ground-wave. Either way an interference pattern can be created. Time delay measurements will be able to decide which explanation actually holds.

Being lower in height the CFA obviously generates less space-wave and more ground-wave over the particular ground conditions at Tanta in Egypt. It should be noted that elsewhere in the world simulated results, together with a very few reported measured results, show a -8dB relative loss for the CFA [6]. A small broadcast antenna that radiates a larger ground-wave than a much larger conventional antenna, in a particular environment, can certainly be said to be more 'effective' in that environment.

An interesting effect observed and explained originally by Millington [8] shows that the ground constants can have a dramatic and in some cases counter-intuitive effect on the measured signal just above ground. Figure 6 is taken from the QinetiQ website http://www.cpar.qinetiq.com/grapple.html (on 16-03-03). This describes the ground-wave simulation programme 'Grapple', which models Millington's measurements. Here you can see either a 20dB increase (counter-intuitive) as the ground-wave signal passes from land to sea or a 20dB decrease as the ground-wave passes from sea to land. This is an example where ground conditions introduce a step of 20dB in field strength measurements. It is therefore not unreasonable to expect that ground conditions under a small antenna could introduce field strength measurement errors of up to 20dB.

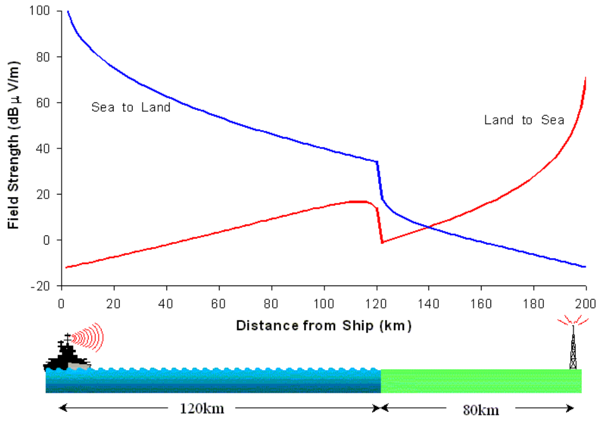

Figure 6 Ground-wave propagation over a land-sea boundary (from QinetiQ's 'Grapple')

We can therefore conclude that failure to take account of the modification of antenna pattern and the presence of near-field (ground) losses in a correct manner renders the IEEE Std 179 an inappropriate efficiency measurement method for antennas near ground. The misunderstanding and misuse of this method has perhaps served to perpetuate the myth that small antennas cannot be made efficient by good design.

ESTIMATION OF GROUND RESISTANCES .

We propose that even loss free ground under an antenna can (re-)radiate, and not just reflect, power coupled into it by an antenna. Furthermore the evidence is that the antenna then launches surface and ground waves that are only lightly coupled to each other and to the sky-wave. The measured far-field pattern is a combination of all these paths depending on the direction and angle above ground. Each of these processes makes at least some contribution to the input resistance of the antenna.

The novel line integral (quasi-scalar bi-vector) potential method was used in [3] to predict some of the tuned loop radiation modes. Essentially this method derives and uses a Green's function that has additional near-field terms; it therefore gives a much more accurate representation of near field and close-in effects. (If it were used in computer simulations such as NEC, such simulation programs could well be considerably improved!)

The new method also determines whether a current radiates or not. If there is a real part of the product of the current and the line integral potential, radiation is predicted to occur. If there is no real part there is no radiation. The imaginary part of the product is the local stored energy.

It shows that the proposition that "it is the current that radiates" is not always true. As an example it can explain the operation of the Goubau line or 'G-string', where the current on the line clearly does not radiate [8]. See Figure 6. The classical explanation is that there is a current in the enamel or paint on the line that cancels out the conductor current. If that were so, a few watts into the line would burn off the enamel! The new line integral potential method can show that there is enough displacement current in the space around the wire to prevent radiation from the wire conduction currents. It also shows that away from any discontinuities the line potential gradient is out of phase with the line current and so no radiation occurs.

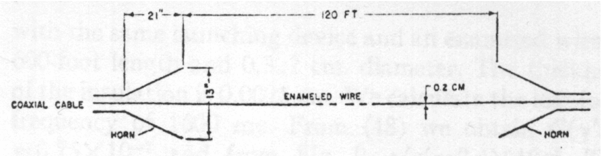

FIG. 11. Sketch of a surface wave transmission line.

Enamel coat on wire 0.005cm (= 50micron), εr = 3, $\tan\delta = 8\times10^{-3}$ - not thick enough to carry space-surface wave

Figure 7 The Goubau 'G-line' low-loss single wire surface-wave transmission line.

Over a PEC (Perfectly Electrically Conductive) ground the new method shows that that the ground surface currents at a distance from the antenna do not radiate strongly (as a Goubau line?). Then the surface power spreads out as 1/r corresponding to the fields spreading out as $1/r^{1/2}$. However close to the antenna ground currents current will spread out as 1/r implying that the surface power is spreading out as $1/r^2$. By using the new line-integral potential we find that these close-in currents radiate significantly.

The ground current has three origins: direct injection, as for a monopole above ground; induction of a surface wave; and reflection from the ground to create an 'image' of the antenna. The distribution of the current is different for these three cases. The first of these cases is simple to analyse. The second and third of these can be approximated by the assumption that the current is injected on an equivalent radius r_i at a small distance from the antenna depending on its height above ground.

Impedance of and surface velocity on PEC Ground-Plane

The surface impedance of a lossless Perfectly Electrically Conducting (PEC) ground plane to the first approximation is $Z_0 = 120p$ or 377Ω/ (ohms per square).

A better value is found as a consequence of using the new potential method. Taking the coupling m between a short wire and a long wire to be asymptotically $m = 1/2p$, we find $e_r = 1 + m$ and $m_r = 1 - m$. This gives $Z_c = 120p \times \ddot{O}\{(1-m)/(1+m)\} = Z_0 \times 0.852 = 321W$.

The additional coupling of the surface to itself also slightly lowers the surface wave group velocity (and thus increases the phase velocity). The group velocity becomes $v_g = c_{em}\,\ddot{O}(e_r m_r) = c_{em} \times \ddot{O}(1 - m^2) = 0.987\,c_{em}$.

This 1.3% reduction is about a half of what was predicted above from Kabbary's measurements. It indicates that the assumed coupling factor is on the low side, and so further measurement or theory is needed.

At a centre point of an infinite ground-plane the input impedance is found to be $Z_c/2\pi = 51.1\Omega$, corresponding to a radial wave propagating out to infinity.

A resonant ground plane of finite size will reduce this impedance value towards zero only on the assumption that there is no ground-plane radiation.

The power at the centre divides in resistance ratio into the space wave launched by the vertical and the surface wave of the ground-plane. For short verticals any radiation from the ground-plane can dominate the overall antenna pattern.

Radiation Resistance of PEC Ground-Plane

A centrally injected radial current sets up a travelling wave magnetic field over the ground-plane surface. The potential at the centre is then found by taking the line integral from infinity to a radius. The real part of the current-potential is the radiated power density at that point. A further integration gives the total power radiated (for the original current input). The radiation resistance is then computed. With the particular coupling assumptions thought to be relevant the radiation resistance for this process on its own is predicted to be between 51.1 and $51.1 \div 4 = 12.8\Omega$, depending on the assumptions made.

Note that this distributed resistive loss is in series with the distributed inductance of an equivalent transmission line. Such losses modify the transmission line input impedance by an amount that is considerably smaller than a simple combination of the resistive loss and the transmission line impedance. Transmission line theory also gives the transmission loss that then occurs [9]. More accurate predictions await further measurements of the input impedances for various ground conditions.

Note that a $\lambda/8$ vertical monopole has a predicted radiation resistance of 12.4Ω, comparable with 12.8Ω. In this case the two antenna patterns would combine with equal strength. The predicted radiation pattern of radiation from ground currents alone has nulls vertically and on the horizon.

ADDITIONAL CONCLUSIONS

The following conclusions are additional to those already in the summary.

Some methods of measuring antenna efficiency over ground are very inaccurate and can have large systematic errors. The new proposed Q-bandwidth measurement method can give accurate results for 'intrinsic' efficiency, and can in principle give an indirect estimate of total ground losses.

Ground effects and environmental losses can give large differences between the 'intrinsic' antenna efficiency of the antenna itself and the IEEE standard definition of antenna efficiency. The 'intrinsic' antenna efficiency has the advantage that it is reasonably independent of the environment.

Ground-wave interference patterns (reported by Kabbary) can be 'explained' if the surface-wave velocity is 2.5% lower than the space-wave velocity.

REFERENCES

1. CHU, L.J.: 'Physical Limitations in Omni-directional Antennas', J Appl. Phys, Vol 19, December 1948, pp1163-1175
2. UNDERHILL, M.J., and HARPER, M., 'Simple circuit model of small tuned loop antenna including observable environmental effects', IEE Electronics Letters, Vol.38, No.18, pp. 1006-1008
3. UNDERHILL, M.J., and HARPER, M.: ''Small antenna input impedances that contradict the Chu-Wheeler Q criterion", not yet published but available at www.ee.surrey.ac.uk/showstaff? harper (or underhill)
4. HAWKER, P., 'Technical Topics' regularly in RSGB RadCom. See Ap 2003, Dec 2002, Nov 2002, May 1999, Feb 1999, Feb 1998, Nov 1997,
5. KABBARY, F.M., 'Daytime Wave Pockets of Medium Wave Mast Antennas' Antennex on-line Issue #59, March 2002
6. TRAINOTTI, V., 'Short medium-frequency AM antennas', IEEE Trans on Broadcasting, September 2001, pp263-284
7. MILLINGTON, G. and ISTED, G.A., 'Ground wave propagation over an inhomogeneous earth. Part 2 Experimental evidence and practical implications', Proc. IEE, 1950, **97**, pp. 209-222
8. GOUBAU, G. "Surface Waves and Their Application to Transmission Lines", J.A.P., Vol. 21, Nov., 1950, pp 1119-1128
9. COLLIN, R.E., 'Antennas and radiowave propagation', McGraw-Hill, 1985, ISBN 0-07-Y66156-1. See section 6.10 on ELF propagation