Proceedings of 20th International Television Symposium 1997
Montreux, Switzerland, 13-17 June

Paper

Explaining some of the magic of COFDM

J H Stott (BBC)

1. INTRODUCTION

Coded Orthogonal Frequency Division Multiplexing (COFDM) [1, 2] has been specified for digital broadcasting systems for both audio -- Digital Audio Broadcasting (DAB) [3] and (terrestrial) television -- Digital Video Broadcasting (DVB-T) [4, 5, 6]. COFDM is particularly well matched to these applications, since it is very tolerant of the effects of multipath (provided a suitable guard interval is used). Indeed, it is not limited to 'natural' multipath as it can also be used in so-called Single-Frequency Networks (SFNs) in which all transmitters radiate the same signal on the same frequency. A receiver may thus receive signals from several transmitters, normally with different delays and thus forming a kind of 'unnatural' additional multipath. Provided the range of delays of the multipath (natural or 'unnatural') does not exceed the designed tolerance of the system (slightly greater than the guard interval) all the received-signal components contribute usefully.

Multipath (natural and unnatural) can alternatively be viewed in the frequency domain as a frequency selective channel response. Another frequency-dependent effect for which COFDM offers real benefit is the presence of isolated narrow-band interfering signals within the signal bandwidth. Note that conventional analogue television signals (NTSC/PAL/ SECAM) essentially behave like narrow-band interferers to COFDM.

COFDM copes with both these frequency-dependent effects as a result of the use of forward error coding. However, rather more is involved than simply adding coding -- the 'C' -- to an uncoded OFDM system. The coding and decoding is integrated in a way which is specially tailored to frequency-dependent channels and brings much better performance than might be thought based on a casual inspection.

This paper attempts to provide a simple explanation of the COFDM 'magic' by which this is achieved.

2. UNCODED OFDM

2.1 What is OFDM?

OFDM spreads the data to be transmitted over a large number of carriers -- typically more than a thousand*. The data rate to be conveyed by each of these carriers is correspondingly reduced. It follows that the symbol length is in turn extended. These modulation symbols on each of the carriers are arranged to occur simultaneously.

The carriers have a common, precisely-chosen frequency spacing. This is the inverse of the duration, called the *active symbol period*, over which the receiver will examine the signal, performing the equivalent of an 'integrate-and-dump' demodulation. This choice of carrier spacing ensures *orthogonality* (the 'O' of OFDM) of the carriers -- the demodulator for one carrier does not 'see' the modulation of the others, so there is no crosstalk between carriers, even though there is no explicit filtering and their spectra overlap.

A further refinement adds the concept of a *guard interval*. Each modulation symbol is transmitted for a total symbol period which is longer than the active symbol period by a period called the guard interval. This means that the receiver will experience neither inter-symbol nor inter-carrier interference provided that any echoes present in the signal have a delay which does not exceed the guard interval. Naturally, the addition of the guard interval reduces the data capacity by an amount dependent on its length. The concept of a guard interval could in principle be applied to a single-carrier system, but the loss of data capacity would normally be prohibitive. With OFDM it is possible to protect against echoes with prolonged delay, simply by choosing a sufficient number of carriers that the guard interval need not form too great a fraction of the active symbol period. Both DAB and DVB-T have a guard interval which is no greater than 1/4 of the active symbol period, but can protect against echo delays of the order of 200 μs (depending on the mode chosen).

Fortunately the apparently very complex processes of modulating (and demodulating) thousands of carriers simultaneously are equivalent to Discrete Fourier Transform operations, for which efficient Fast Fourier Transform (FFT) algorithms exist. Thus integrated circuit implementations of OFDM demodulators are feasible for affordable mass-produced receivers.

Refs.1 & 2 give more detailed descriptions of the basis of OFDM operation.

2.2. Effects of frequency-dependency on uncoded OFDM

It is possible to analyse the effects of frequency-dependent channels on an *uncoded* OFDM system. The signal-to-noise ratio (SNR) of each carrier is noted, the corresponding bit-error ratios (BERs) for each carrier are determined as a consequence of the SNRs and finally the BER for the whole data signal obtained by averaging the BERs of all the carriers used. This could in principle be performed for any selective channel and whatever modulation scheme is used on every carrier. Figure 1 illustrates an arbitrary selective channel.

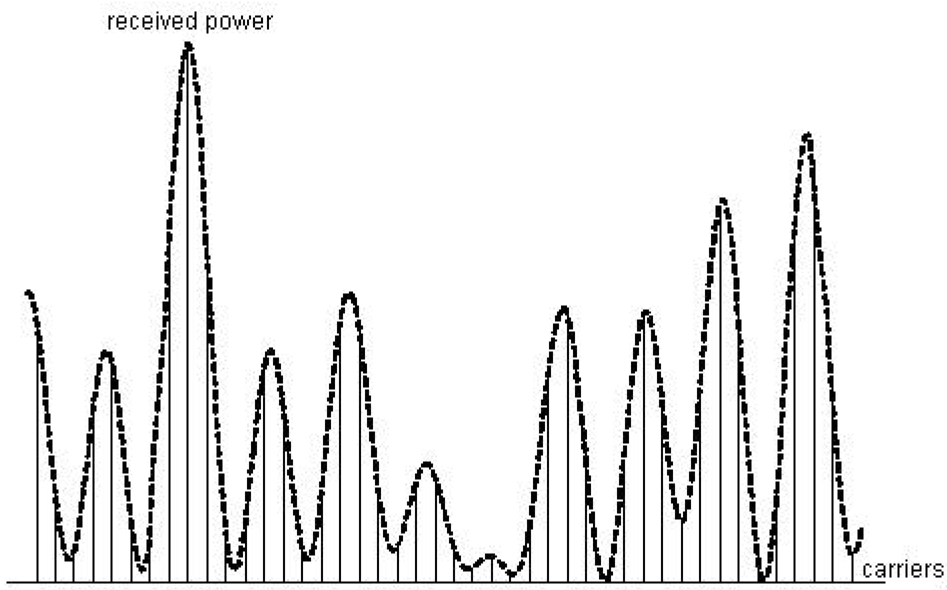

Fig. 1 - The effects of an arbitrary selective channel on the carriers of an OFDM signal.

This example represents the response of a channel having four arbitrarily selected path delays and attenuations. The dotted line represents the power frequency response of the channel.

However, we can avoid even that complexity for our present purpose, since a simple, brief examination of the effects yields useful results.

If there is one carrier, amongst *N* carriers in all, which is badly affected by interference, then the 'symbol' error ratio, SER, (where 'symbol' denotes the group of bits** carried by one carrier within one OFDM symbol) will be of the order of 1 in *N*, *even with infinite SNR*.

Similarly, if there is frequency selectivity, some carriers will be boosted and some attenuated, the SNRs and SERs for each carrier varying accordingly. Clearly, if there is a 0dB echo of delay such that every *m*th carrier is completely extinguished, then the SER will be of the order of 1 in *m* even at infinite SNR.

Such a simple analysis already shows that for any reasonable number of carriers, CW interference affecting one carrier is less of a problem than a 0dB echo -- for which an SER of 1/4 could occur with an echo delay of 1/4 of the active symbol period***. Clearly, uncoded OFDM is not satisfactory for use in such extremely selective channels.

It will be equally obvious that simply adding hard-decision-based coding to this uncoded system will not be a sufficient cure -- especially for the case of a 0dB echo which nulls (say) one carrier in 4.

The solution is the use of convolutional coding with soft-decision decoding, *properly integrated* with the OFDM system.

3. SOFT-DECISION DECODING

Let us first take the example of a simple single-carrier system. (The extension to OFDM will follow in the next section).

Consider a 2-level signal+. One bit can be transmitted per symbol, with say a '0' being sent as -1V and a '1' as +1V.

At a receiver, assuming that the gain is correct, we should expect to receive a signal always in the vicinity of either -1V or +1V, depending on whether a '0' or a '1' was transmitted, the departure from the exact values ±1V being caused by the inevitable noise added in transmission.

A simple receiver might operate according to the rule that negative signals should be decoded as '0' and positive ones as '1'. This is an example of a *hard decision*, with 0V as the *decision boundary*. If the instantaneous amplitude of the noise were never to exceed ±1V then this simple receiver would make no mistakes. But noise usually has a continuous distribution such as Gaussian and may occasionally have a large amplitude, although with lower probability than for smaller values. Thus if say +0.5V is received, it most probably means that a '1' was transmitted, but there is a smaller yet still finite probability that actually '0' was sent. Common sense suggests that if a large amplitude signal is received we can be more confident in the hard decision than when the amplitude is small.

This view of a degree of confidence is exploited in *soft-decision* Viterbi decoders. These maintain a history of many possible transmitted sequences, building up a view of their relative likelihoods and finally selecting the value '0' or '1' for each bit according to which has the *maximum likelihood*. For convenience, a Viterbi decoder *adds log-likelihoods* (rather than multiplying probabilities) to accumulate the likelihood of each possible sequence. It can be shown [7] that in the case of BPSK or QPSK the appropriate log-likelihood measure or *metric* of the certainty of each decision is indeed simply proportional to the distance from the decision boundary. (The slope of this linear relationship itself also depends directly on the signal-to-noise ratio, to which we shall return in the next Section). Thus the Viterbi decoder is fed with a *soft decision* comprising both the hard decision (the sign of the signal) together with a measure of the amplitude of the received signal.

With other rectangular-constellation modulation systems, such as 16-QAM or 64-QAM, each axis carries more than one bit, often with Gray coding. At the receiver, a soft decision can be made separately for each received bit. The metric functions are now more complicated than for QPSK, being different for each bit, but the principle of the decoder exploiting knowledge of the expected reliability of each bit remains.

4. SOFT-DECISION DECODING PROPERLY APPLIED TO COFDM

Metrics for COFDM are slightly more complicated. We start from the understanding that the soft-decision information is a measure of the confidence to be placed in the accompanying hard decision.

When data are modulated onto a single carrier in a time-invariant system then a *priori* all data symbols suffer from the same noise power on average; the soft-decision information simply needs to take note of the random symbol-by-symbol variations that this noise causes.

When data are modulated onto multiple carriers, as in COFDM, the various carriers will have different signal-to-noise ratios. For example, a carrier which falls into a notch in the frequency response will comprise mostly noise; one in a peak will suffer much less. It follows that in addition to the symbol-by-symbol variations there is another factor to take account of in the soft decisions: data conveyed by carriers having a high SNR are *a priori* more reliable than those conveyed by carriers having low SNR. This extra *a priori information is usually known as channel-state information* (CSI).

The channel-state information concept similarly embraces *interference* which can affect carriers selectively, just as noise does.

Including channel-state information in the generation of soft decisions is the key to the unique performance of COFDM in the presence of frequency-selective fading and interference.

5. SIMPLE EXAMPLE: A 0 dB ECHO

It is not obvious that the arrangement described in the previous Section can work perfectly satisfactorily when there is a 0dB echo of long delay, e.g. one causing frequent complete nulls. A simple explanation of this particular example may help.

5.1 Simple example channel

Let us consider a simple example in which there is a 0dB echo of such a delay (and phase) as to cause a complete null on one carrier in 4. Figure 2 illustrates the effect of this selective channel: 1 carrier in 4 is nulled out, while 1 carrier in 4 is actually boosted, and the remaining 2 are unaffected. Note that received *power* is shown, to which the SNRs of the carriers will be proportional if the receiver noise is flat, as is usual. (The response of such a channel is often portrayed with the vertical axis representing *voltage* rather than power. In that case the curve has a rectified-cosine form, with sharp 'teeth' at the minima, in contrast to the 'raised-cosine' form illustrated here). The 'mean power' marked is the mean of all carriers. It is equal to the total received power (via both paths) shared equally between all carriers.

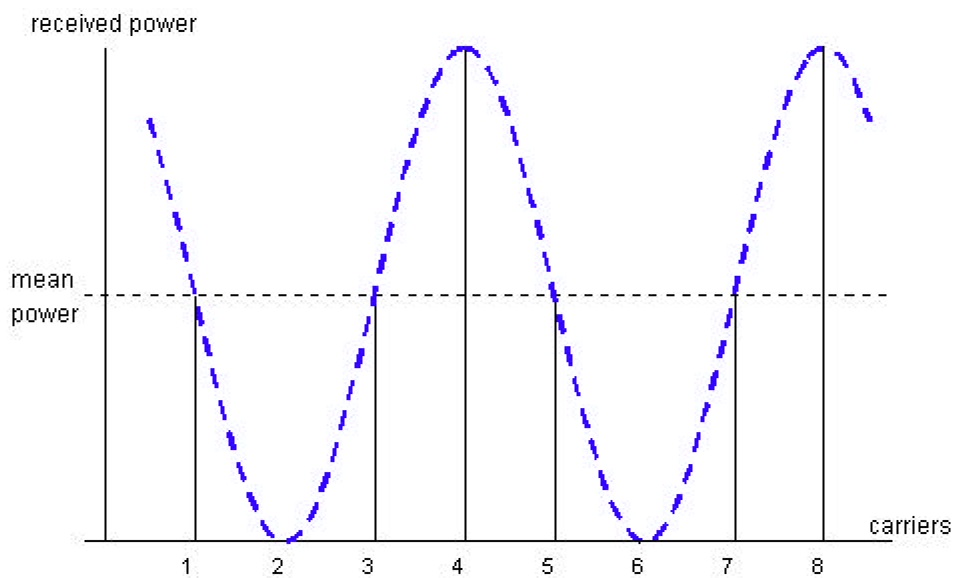

Fig. 2 - The effect of a channel with a single 0 dB echo of long delay, such that exactly 1 carrier in 4 is nulled out.

Although few COFDM carriers are illustrated, the pattern repeats cyclically for all of them. The dotted line represents the power frequency response of the channel.

5.2 Effect on uncoded OFDM

As noted above, the SNRs of the carriers will be proportional to the received powers, and, as Figure 2 shows, 1 in 4 is zero. As already explained in §2.2, this means that, even when the overall SNR is arbitrarily large, the overall SER will be of the order of 1 in 4. Thus uncoded OFDM is of no practical use with this channel. Furthermore, the addition of error correction using hard-decision decoding would make little impact on this great an SER.

5.3 How using CSI helps

Now consider the application of channel-state information when coding is added to the OFDM system.

Recall that the Viterbi metrics are weighted according to the SNR of the corresponding carriers. Clearly, the bits from the nulled carriers are effectively flagged as having 'no confidence'. This is essentially the same thing as an *erasure* -- the Viterbi decoder in effect just records that it has no information about these bits.

Now there is another well-known case of regularly occurring erasures, namely *punctured codes*. Typically, convolutional codes intrinsically have code rates expressed as simple fractions like 1/2 or 1/3. When a code having higher rate (less redundancy) is needed then one of these lower-rate 'mother' codes is *punctured*, that is to say certain of the coded bits are just not transmitted, according to a regular pattern known to the receiver. At the receiver 'dummy bits' are re-inserted to replace the omitted ones, but are marked as erasures -- bits having zero confidence -- so that the Viterbi decoder treats them accordingly. Punctured codes obviously are less powerful than the mother code, but there is an acceptable steady trade-off between performance and code rate as the degree of puncturing is increased. They are sometimes called *rate-compatible punctured codes* (RCPC).

Suppose we take a rate-1/2 code and puncture it by removing 1 bit in 4. The rate-1/2 code produces 2 coded bits for every 1 uncoded bit, and thus 4 coded bits for every 2 uncoded bits. If we puncture 1 in 4 of these coded bits then we clearly finish by transmitting 3 coded bits for every 2 uncoded bits. In other words we have generated a rate-2/3 code. Indeed, this is exactly how the rate-2/3 option of DVB-T is made.

Now return to our simple COFDM example in which 1 carrier in 4 is nulled out by the channel -- but the corresponding bits are effectively flagged as erasures thanks to the application of channel-state information. 2 out of 3 of the remaining carriers are received at the same SNR as that of the overall channel, while 1 is actually boosted, having an improved SNR. Suppose that rate-1/2 coding is used for the COFDM signal. It follows that the SNR performance of COFDM with this *selective channel* should be very slightly better (because 1 carrier in 4 is boosted) than that for a single-carrier (SC) system using the corresponding punctured rate-2/3 code in a *flat* channel. In other words, the effect of this very selective channel on COFDM can be directly estimated from knowledge of the behaviour of puncturing the same code when used in a SC system through a flat channel.

This explains how the penalty in required CNR for a COFDM system subject to 0dB echoes may be quite small, provided a relatively powerful inner code is used together with the application of channel-state information. It also explains how the code rate plays the dominant part, while the choice of constellation is of lesser importance. If we consider the application of puncturing to a rate-1/2 code in a SC system, the change in required CNR between unpunctured and punctured codes remains similar, increasing only very slightly as higher-order constellations are used.

5.4 Varying the delay of the echo

So far we have considered a very special example so as to make it easy to explain by invoking the close analogy with the use of code puncturing. But what of other delay values?

If the relative delay of the echo is rather shorter than we just considered, then the notches in the channel's frequency response will be broader, affecting many adjacent carriers. This means that the coded data we transmit should not simply be assigned to the OFDM carriers in order, since at the receiver this would cause the Viterbi soft-decision decoder to be fed with clusters of unreliable bits. This is known to cause serious loss of performance, which we avoid by *interleaving* the coded data before assigning them to OFDM carriers at the modulator. A corresponding de-interleaver is used at the receiver before decoding. In this way the cluster of errors occurring when adjacent carriers fail simultaneously (as when there is a broad notch in the frequency response of the channel) is broken up, enabling the Viterbi decoder to perform better. As just described, the process could be called *frequency interleaving*; where significant temporal fading is expected as well (e.g. in mobile operation) then the coded data may also be re-distributed over time, providing *time interleaving*. Time interleaving is often used in single-carrier systems, DVB-T uses just frequency interleaving, while DAB (specifically designed for the difficult conditions of mobile operation) uses both.

In the DVB-T system, the coding and interleaving just described are the inner code and inner interleaving; a further stage of outer interleaving and outer (Reed-Solomon) coding is added to complete the error correction arrangements [4, 6].

Unfortunately, the combined processes of interleaving, selective channel and Viterbi decoding are not readily amenable to analysis. The only practicable way to quantify the performance is by software simulation or practical hardware experiments. Extensive software simulation was performed to help choose the inner interleaving algorithm for DVB-T. In this way it was ensured that the performance of the system is as little dependent as possible on the fine detail of the channel response. For example, a range of simulations was performed for a channel containing a 0dB echo whose delay was varied over the entire permitted range, with remarkably uniform results -- even using 64-QAM.

5.5 Significance for SFNs

Our simple example of a 0dB echo often crops up when considering SFNs. If two synchronised COFDM transmitters operate on a common frequency there will some- where be locations where the two signals will be received at equal strength (and with a relative delay, depending on the geometry of the situation, which we assume to be within the system limits). An obvious question is: does reception suffer or benefit from this situation?

Clearly, compared with receiving either transmitter alone, the total-received-signal-to-noise power ratio (CNR) is doubled, i.e. increased by 3 dB, expressed in familiar decibel notation. However, the presence of the two transmissions makes reception *selective* rather than *flat* (as we might hope to have with a single transmission, without 'natural' echoes). This increases the CNR required to achieve the same BER, in a way which depends on the error-correcting code in use.

We have already seen a qualitative argument how this increase in CNR requirement may be closely related to the performance of punctured codes. Simulation shows that the increase in CNR requirement between flat and 0dB-echo channels is just below 3dB for a rate-1/2 code, while it is greater for higher-rate codes which have already been punctured. Practical experience supports the order of 3dB for a rate-1/2 code, while for rate-2/3 the increase is of the order of 6dB [9].

It follows that with rate-1/2 coding, receiving two signals of equal strength, in place of either one alone, increases the received CNR by 3dB while also increasing the CNR required for satisfactory reception (in what is now a highly-selective channel) by about the same amount. The performance is thus unchanged by adding the second path.

Since for most practical purposes the case of the 0dB echo appears to be more or less the worst one++, this is very encouraging for planning and developing SFNs.

Practical SFNs are likely to evolve rather than be switched on, fully-formed overnight. Clearly, it would be unwelcome if receivers which initially perform satisfactorily subsequently lose service when more transmitters are added. We can see that, in principle, such a problem should be avoided if a rate-1/2 code is used, and more generally should prove to be limited in extent provided a 'strong' code (i.e. one whose rate is not increased much beyond 1/2) is used. This would give those broadcasters choosing to use SFNs the freedom to commission transmitters in sequence. Furthermore it would allow them the flexibility to add filler transmitters to cover those areas found in practice to be unserved by the planned network, subject to keeping the relative path-delays within system limits.

6. CONCLUSIONS

COFDM is a modulation scheme which is especially tailored to work well with selective channels and isolated CW (or analogue TV) interferers. The forward error-correction coding -- the 'C' in COFDM -- is the key ingredient. However, the desired results are only achieved when the coding is closely integrated with the OFDM system.

The 'COFDM magic' is achieved by the use of *channel-state information* (CSI). In the presence of CW interferers and/or a selective channel, some OFDM carriers will be worse affected than others. This state of affairs can be recognised by the receiver, which builds up channel-state information about the reliability of each carrier and uses this to supplement the soft decision information used by the Viterbi decoder. This achieves a substantial increase in performance compared with an uncoded OFDM system or one which makes no use of CSI.

The performance of uncoded OFDM systems is amenable to analysis but the results do not give useful guidance about the performance of COFDM with CSI. Indeed they would suggest (quite falsely) that performance in selective channels would be very poor.

Coded systems with CSI can only be assessed in general by full software simulation or tests with real hardware. Nevertheless, a qualitative argument has been presented here for a testing (yet realistic) scenario -- a 0dB echo -- in order to explain how, by the use of CSI, COFDM copes well with it.

Since the effect of a selective channel is shown to be similar to the effect of puncturing the error-correcting code, it follows that 'strong' (unpunctured) codes will give the best COFDM performance on selective channels. This may lead to the choice of different combinations of code-rate and modulation-constellation than would be made for the less-demanding flat channel.

7. ACKNOWLEDGEMENTS

The author wishes to thank the many colleagues, within the BBC and within collaborating organisations throughout Europe, who have helped him to develop his understanding of COFDM.

BBC participation in developing the DVB-T specification and subsequent tests of it has been partly funded by the European Commission through the RACE dTTb and ACTS VALIDATE projects.

8. REFERENCES

1. European Broadcasting Union (EBU), 1988. Advanced digital techniques for UHF satellite sound broadcasting. Collected papers on concepts for sound broadcasting into the 21st century.
2. MADDOCKS, M.C.D., 1993. An introduction to digital modulation and OFDM techniques. BBC Research Department Report No. RD 1993/10.
3. ETS 300 401 (1994): Radio broadcast systems; Digital Audio Broadcasting (DAB) to mobile, portable and fixed receivers.
4. ETS 300 744 (1997): Digital broadcasting systems for television, sound and data services; framing structure, channel coding and modulation for digital terrestrial television.
5. STOTT, J.H., 1996. The DVB terrestrial (DVB-T) specification and its implementation in a practical modem. Proceedings of the 1996 International Broadcasting Convention, IEE Conference Publication No. 428, pp. 255-260, September.
6. MØLLER, L.G., 1997. COFDM and the choice of parameters for DVB-T. Proceedings of 20th International Television Symposium, Montreux.
7. OLIPHANT, A., 1997. Validate -- verifying the European specification for digital terrestrial TV and preparing for the launch of services. Proceedings of 20th International Television Symposium, Montreux.
8. PROAKIS, J.G., 1989. Digital communications, 2nd edn. McGraw-Hill International.
9. MORELLO, A., BLANCHIETTI, G., BENZI, C., SACCO, B., and TABONE, M., 1997. Performance assessment of a DVB-T television system. Proceedings of 20th International Television Symposium, Montreux.

*	For example, DVB-T has options for using either 1705 or 6817 carriers.	Return to Text
**	The number of bits carried by each modulation symbol will depend in the usual way on the choice of modulation constellation, e.g. 2 bit/symbol for QPSK, 4 bit/symbol for 16-QAM, and so on. Note that the net capacity may be reduced from these figures if forward error correction is added to the basic system.	Return to Text
***	Such an echo would otherwise be acceptable if the guard interval were chosen to be 1/4 of the active symbol period, as is possible in both DAB and DVB-T.	Return to Text
+	Strictly, the signal described is an example of Amplitude-Shift Keying (ASK) which could also be called 2-PAM (Pulse Amplitude Modulation) if used to modulate a carrier. When coherent demodulation is used, both BPSK (Binary Phase-Shift Keying) and one axis of QPSK (Quadrature Phase-Shift Keying) are essentially similar.	Return to Text
++	Certain more complicated combinations of path amplitudes and delays, with more than two paths, may be found to be more critical for specific COFDM systems, but the probability of them occurring in practice (and coinciding with desired reception locations) appears to be extremely small. Remember that locations satisfying the peculiar delay relationships are determined by geometry, while the amplitude relationships are primarily determined by details of terrain, antenna heights and local clutter.	Return to Text

Top

Back to
Published Papers