Appl. Phys. B 91, 189–194 (2008)
DOI: 10.1007/s00340-008-2961-2

Applied Physics B
Lasers and Optics

B. WOODWARD✉
H. SARI

Underwater speech communications with a modulated laser

Department of Electronic and Electrical Engineering, Loughborough University, Leicestershire LE11 3TU, UK

Received: 27 June 2007/Revised: 27 January 2008
Published online: 5 March 2008 •

ABSTRACT A novel speech communications system using a modulated laser beam has been developed for short-range applications in which high directionality is an exploitable feature. Although it was designed for certain underwater applications, such as speech communications between divers or between a diver and the surface, it may equally be used for air applications. With some modification it could be used for secure diver-to-diver communications in the situation where untethered divers are swimming close together and do not want their conversations monitored by intruders. Unlike underwater acoustic communications, where the transmitted speech may be received at ranges of hundreds of metres omnidirectionally, a laser communication link is very difficult to intercept and also obviates the need for cables that become snagged or broken. Further applications include the transmission of speech and data, including the short message service (SMS), from a fixed installation such as a sea-bed habitat; and data transmission to and from an autonomous underwater vehicle (AUV), particularly during docking manoeuvres. The performance of the system has been assessed subjectively by listening tests, which revealed that the speech was intelligible, although of poor quality due to the speech algorithm used.

PACS 42.55.Px; 42.60.By; 42.62.-b

1 Introduction

Diver communication systems have remained essentially unchanged for decades. Systems using hard-wired links have been in use for nearly a hundred years, while through-water acoustic systems have been in use for at least forty years. Despite the rapid exploitation of digital mobile telecommunications technology, as a result of advances in speech coding methods and digital signal processors, underwater speech communications systems have remained relatively archaic. The through-water acoustic systems available commercially nearly all use analogue technology and are clones of earlier versions working on the principle of single-sideband modulation. The aim of our research has been to design a digital speech communications system for divers, firstly using an acoustic channel with its inherent problems of multipath interference and noise [1, 2], and then using a free-space laser channel for certain clear-water scenarios. The system with a modulated laser source uses a digital signal processor (DSP) but is less complicated than its acoustic counterpart. It is restricted to relatively short ranges in clear water, but is capable of secure communications because in practice it is extremely difficult for an intruder to intercept the beam and decode the transmitted signals in most practical scenarios. The system is also capable of much higher data rates, for applications where this is necessary, than the corresponding acoustic configuration and is not susceptible to multipath interference and noise. This paper describes the detailed design and implementation of a prototype system and its initial testing. It is envisaged that the design may be extended to an underwater local area network (LAN), such as an Ethernet [3]. While this can be easily set up and at a lower cost than using optical fibres, it is recognised that the properties of the through-water channel can significantly affect the performance of such a system.

✉ Fax: +44-1509-227014, E-mail: b.woodward1@lboro.ac.uk

2 Transmission considerations

2.1 *Attenuation of light in water*

The propagation of light in water has been studied by many researchers [3–12]. Their studies include visible spectrum optical communications and distance sensing [8], the use of laser diodes and light-emitting diodes and the effects of extraneous light [9], the design of a laser underwater camera image enhancer for mine warfare applications [10], signal processing for a laser underwater target detection system [11], and underwater docking of autonomous underwater vehicles using optical terminal guidance [12]. As far as can be ascertained, there have been no previous studies of underwater speech communications by optical means.

The attenuation of light by water is caused by two independent mechanisms, scattering and absorption, which affect the amplitude, phase, and arrival angle of the light beam. Scattering is a random process that changes the directions of individual photons without altering their properties, while absorption is a thermodynamically irreversible process that transforms the energy of photons into thermal energy. This is the major absorption mechanism in the sea and varies considerably with wavelength. Minimum absorption occurs at the wavelengths of blue–green light, i.e. 430–570 nm (0.43–0.57 μm), as shown in Fig. 1 for distilled water. For sea water the minimum absorption is about 0.0001 cm^{-1} at 500 nm, allowing light to penetrate to depths of 600 m or greater [13, 14]; the value rises to about 0.01 cm^{-1} at 300 nm and 750 nm, but the absorption spectrum and exact minimum vary depending on the purity of the water [15]. At ultraviolet wavelengths the attenuation is dominated by small-particle scattering, while at infrared wavelengths it is caused by absorption. Particle scattering depends on the ratio $2\pi r/\lambda$, where r is the effective spherical particle radius and λ is the wavelength. For small particles, defined as $r < 0.1$ μm, the Rayleigh scattering condition applies, i.e. $2\pi r \leq \lambda$, such that the total scattering diminishes as r decreases. As the ratio of large particles to small particles increases, the Rayleigh scattering contribution increases. If the particle sizes are large compared to the wavelength, i.e. if $2\pi r \geq \lambda$, the attenuation is due to so-called Mie scattering.

The beam attenuation coefficient, $\alpha(\lambda)$, is a measure of the decay of the unscattered light and may be described by the Beer–Lambert law as

$$P_R(\lambda) = P_T e^{-\alpha(\lambda)r} , \quad (1)$$

where $P_T(\lambda)$ is the transmitted optical power, $P_R(\lambda)$ is the received optical power, and r is the path length in the water.

The transmittance of an underwater beam is defined by

$$T(\lambda) = \frac{P_R(\lambda)}{P_T(\lambda)} ; \quad (2)$$

therefore $\alpha(\lambda)$, with units of m^{-1}, or nepers, can be expressed in terms of the transmittance as

$$\alpha(\lambda) = \frac{1}{r} \ln \frac{1}{T(\lambda)} , \quad (3)$$

where $\alpha(\lambda)$ varies with depth and water temperature due to the non-homogeneous nature of sea water.

It is more useful to be able to calculate the maximum expected range in terms of a measured value of attenuation expressed in dB/m. The corresponding expression then becomes

$$r(m) = \frac{1}{\alpha(\mathrm{dB/m})} 10 \log_{10} \frac{P_T}{P_R} . \quad (4)$$

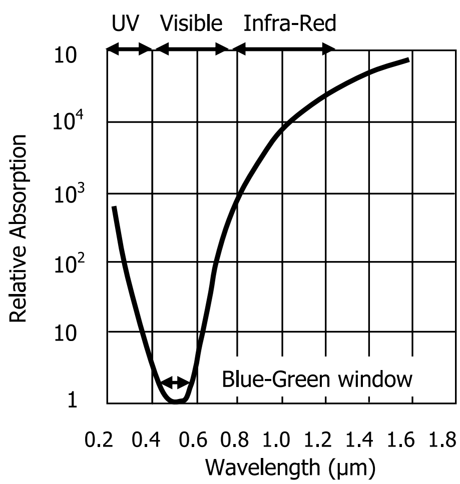

FIGURE 1 Relative absorption spectrum for distilled water [13]

Studies have been carried out on light propagation in water at different wavelengths [3] and, for the red laser diode used in this work, with a wavelength of 660 nm, the attenuation coefficient can vary from 2 dB/m in clear oceanic water up to 6 dB/m or more in shallow coastal water.

2.2 *Practical considerations*

While the longest transmission range can be achieved with a blue–green laser, this type of device is expensive and so a semiconductor red laser diode is used in the prototype. The output radiation of the device is highly divergent; therefore, to achieve a practical design, a high-quality commercial collimator is used. The design is a compromise between the numerical aperture, focal length, communication range, and output beam diameter.

Communication with a modulated laser beam has limitations in terms of range and due to the necessity for line-of-sight alignment, but it has the advantage of being virtually undetectable due to its high directionality and the difficulty of intercepting the transmitted beam. The technique is therefore highly suitable for covert transmission. Another advantage of laser communications is the broad bandwidth available, which makes high data rates possible. This bandwidth may be shared by a large number of time-division multiplexed (TDM) channels, an important feature for speech communications. For this system, the data rate is limited only by the transmitter and receiver bandwidths, while the communication range depends only on the channel characteristics, in this case the clarity of the water. A disadvantage of laser communications is that the beam may be deflected from its straight-line path by reflection and refraction at an interface or by refraction resulting from the presence of a sound velocity profile, which is a function of temperature, salinity, and pressure in the sea.

3 System design

To transmit digitally coded information by means of a semiconductor red laser diode, a driver must be included, for either continuous-mode or pulsed-mode transmission [5]. For this application a pulsed-mode driver is used to transmit digitally coded signals. By considering the operational parameters of the laser diode (Samsung Ltd, SLV71A), the transmitter must supply a 50-mA forward current to produce 5-mW optical output power. The device has a maximum data rate of 1 Mbit/s.

The signals transmitted through the water from the laser diode are picked up by a photodetector. The optical sensitivity and response speed of the photodetector, including its dark current shot noise and thermal noise current, are the dominant parameters of the receiver that significantly affect the performance of the system. External components with low thermal noise specifications were therefore chosen to reduce the noise-equivalent power (NEP) and allow the detection of lower light levels. For the photodetector used here (Centronics, OSD5-5T silicon), the specified NEP is 1.738×10^{-14} W

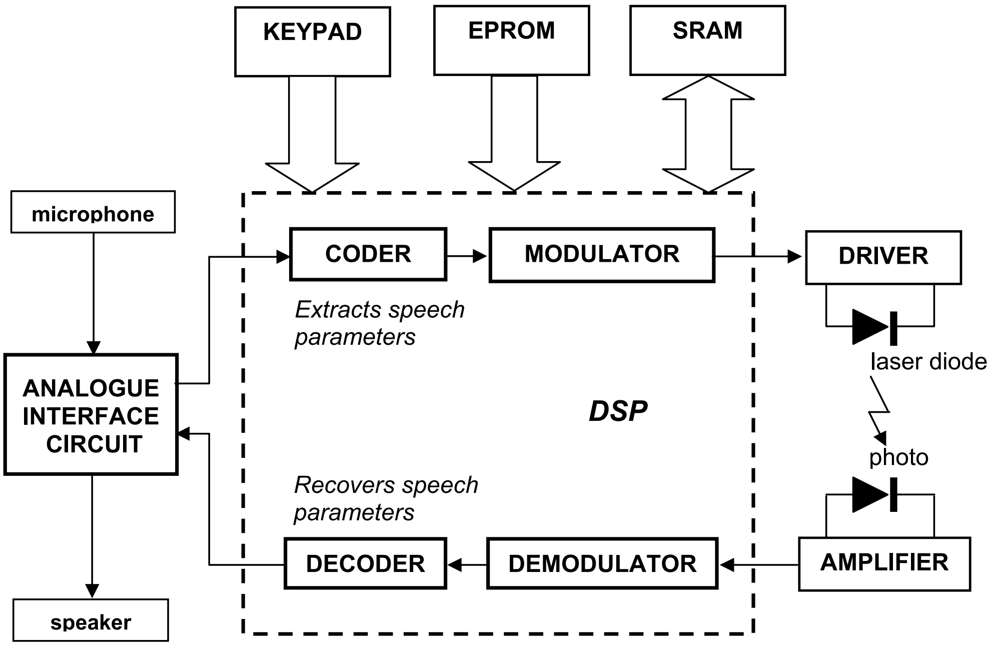

FIGURE 2 Underwater speech communications system with a modulated laser

(−107.6 dBm) and the sensitivity, R_λ, is 0.4 A/W at $\lambda =$ 660 nm.

To achieve real-time digital underwater speech communications, coded speech samples are transmitted and received by implementing a suitable digital modulation and demodulation technique. For this, a digital signal processor (DSP) is essential to implement the complex signal processing computations, which minimises the number of components needed, thus allowing the realisation of a very small, stand-alone, portable system. The system design, shown in Fig. 2, is based on an earlier system designed for speech communications using an underwater acoustic channel [1, 2]. It was configured to transmit speech data at a carrier frequency of 70 kHz, which was the resonant frequency of the omnidirectional, electrostrictive transducer used. For the laser communications system, some of the hardware of the acoustic system has been modified, but the software remains unchanged. The principal hardware components are an analogue interface circuit (AIC, Texas Instruments, TLC320AC01) and a DSP (Texas Instruments, TMS320C31). Peripheral components include erasable programmable read only memory (EPROM), static random access memory (SRAM), and a detachable keypad for programming. The transmitter section first digitises the analogue speech signal from a microphone via the AIC (pre-amplifier, anti-aliasing filter, and analogue-to-digital converter). The digitised speech samples are then coded by the DSP to extract a set of *parameters* (outlined below) that represent the original speech as three-bit symbols, which modulate a laser diode via a driver circuit. The receiver section amplifies the signal received by a photodetector; then the DSP demodulates and decodes it. Finally, the output of the AIC (digital-to-analogue converter, low-pass filter, and amplifier) is fed to a speaker or an earpiece. The rationale of the choice of speech coding and digital modulation techniques, together with details of their practical implementation, is described below.

4 Speech coding

There are several international waveform-coding standards, such as 64-kbit/s pulse code modulation (PCM) and 16-kbit/s adaptive differential pulse code modulation (ADPCM). They produce high-quality speech, but have inadequate spectrum efficiency when applied to bandwidth-limited applications such as satellite communications, digital mobile radio, private networks, and especially underwater communications. There are also problems meeting the sampling rate requirement for speech signals. Since the speech bandwidth is typically 3200 Hz (200–3400 Hz), the minimum sampling rate needed for recovery of a received digital signal back to its analogue form must be at least twice this frequency and typically 8 kHz. The corresponding sampling period of 125 μs is the maximum time allowed to transmit a coded speech sample. Since low-rate transmission is essential for underwater speech communications, it is preferable to use a parametric speech coder, or vocoder, that compresses the speech information for transmission.

Amongst the parametric standards, FS1015 linear prediction coding (LPC), with a bit rate of 2.4 kbit/s, has been widely used for many years [16]. Owing to its simplified speech production model, the speech quality sounds far from natural. Other coding standards use *hybrid codecs* to produce good-quality speech for mobile communications. They include FS1016 code-excited linear prediction (CELP), used for secure military communications at 4.8 kbit/s, and regular pulse excited–long term prediction (RPE-LTP), implemented by Group Special Mobile (GSM) at 13 kbit/s as a standard in Europe. There is also half-rate GSM using vector sum excited linear prediction (VSELP) coding at 5.6 kbit/s [16].

Although not considered a priority in the present design, it is interesting to note that with the parametric speech coding standards it is feasible to accommodate large numbers of time-division multiplexed (TDM) data channels. With the laser diode chosen here, with a maximum data rate of 1 Mbit/s, 76 channels are possible with RPE-LTP and 416 with LPC. By using the LPC, efficient synthesis may be achieved by transmitting *frames* of the speech waveform as a set of *parameters*, which are (i) the amplitude, (ii) a voiced/unvoiced decision and pitch period for voiced sounds, and (iii) a set of so-called linear prediction coefficients [16]. The vocoder adopted here is an improved LPC version, known as LPC-10, which uses the mixed-excitation linear predictive (MELP) algorithm. Its frame interval of 22.5 ms contains 180 voice samples that are transmitted at a sampling rate of 8000 Hz [17]. LPC-10 has been used as a military standard; it is the least complex algorithm with one of the lowest data rates, and so was suitable for our prototype.

5 Digital modulation

5.1 *Transmission of speech data*

Pulse position modulation (PPM) is a suitable technique for underwater speech transmission, because only one carrier frequency is needed, all transmitted pulses have equal amplitude, and phase coherence from pulse to pulse is unnecessary. As this is essentially a digital technique, it may be referred to as digital pulse position modulation (DPPM), which was first used for underwater acoustic communications many years ago [18]. Later, DPPM proved to be effective for transmitting digital information in an optical channel [19]. While the bandwidth efficiency is decreased due to an increase in the required channel bandwidth [20], this is not a serious disadvantage here.

In this application, eight-slot DPPM is implemented for the transmission of quantised speech parameters as three-bit symbols. This choice is based on previous experience with the acoustic system mentioned above. As shown in Fig. 3, three timing intervals are required for the DPPM technique, T_{frame}, T_{symbol}, and T_{slot}. Each speech frame interval is divided into 18 symbols (54 bits), each 1-ms long, and a 4.5-ms dead space, i.e. $T_{\text{frame}} = 22.5$ ms, corresponding to a bit rate of $(1000/22.5) \times 54 = 2.4$ kbit/s. Each symbol is divided into eight data slots, representing all the possible three-bit symbols (000, 001, 010, etc) and two guard band slots, i.e. $T_{\text{symbol}} = 1$ ms and $T_{\text{slot}} = 100$ µs. One of the 10 slots in each symbol interval is occupied by a 70-kHz square waveform, another design parameter from the earlier acoustic version of the system.

5.2 *Detection and decoding*

Since the envelope of the DPPM received signal is used for data decoding, the amplitude shift keying (ASK) detection principle is employed. This incoherent technique is sensitive to multipath propagation but, unlike phase shift keying (PSK) [21], no phase detection is needed and, unlike frequency shift keying (FSK), no band-pass filtering is needed. Accurate synchronisation of the transmitter and receiver is therefore needed only at the slot level, which is why a synchronisation (SYNC) pulse is transmitted at the beginning of each 22.5-ms frame, as shown in Fig. 3, to detect the pulse position accurately. With the system initially in receive mode, the detected signal may be the SYNC signal or DPPM data. To synchronise with the transmitter, the receiver must lock onto the leading edge of the SYNC signal. Then, the decoding process starts.

After band-pass filtering, centred on 70 kHz and with a 20-kHz bandwidth, the input signal is applied to a low-pass filter (envelope detector) to obtain a baseband signal with an approximate duration of T_{slot}. To process the baseband input signal, an eight-bit analogue-to-digital converter is used with a 40-kHz sampling frequency, i.e. 40 samples per symbol interval or four samples per slot interval. The energy in each slot is then computed by the DSP and the result is stored in SRAM. This is done for all the slot intervals until the position of maximum energy is established. The corresponding three-bit code

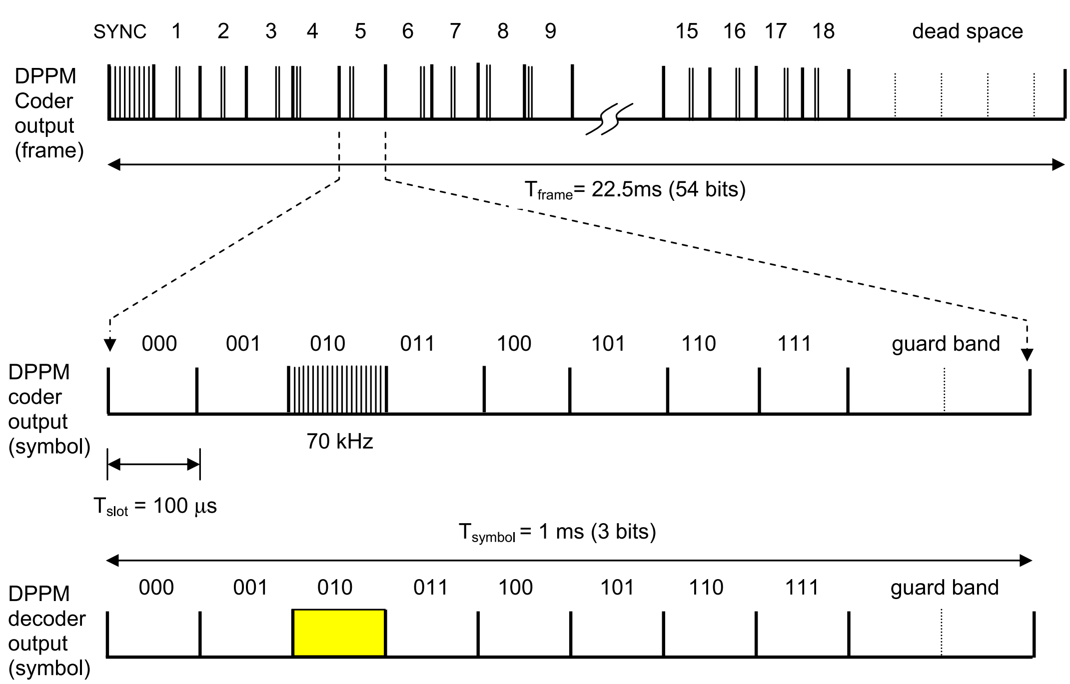

FIGURE 3 DPPM format for transmitter and receiver

is taken as the transmitted LPC parameters, which are used to synthesise the speech.

6 Results and conclusions

For a transmitted optical power of 5 mW, i.e. 7 dBm, the maximum theoretical range at which the laser beam is just discernable corresponds to a power reduction to the NEP (−107.6 dBm), which is defined as having a signal-to-noise ratio of 1 (0 dB), i.e. by 114.6 dB. Equation (4) and Fig. 4 show that a range of about 57 m is achievable for an attenuation of 2 dB/m ('oceanic clarity') and about 19 m for 6 dB/m ('coastal waters'). During testing of the system, the laser diode in the transmitter and the photodetector in the receiver were both placed in watertight enclosures with glass ports, and were perfectly aligned in an indoor water tank (9-m long by 5-m wide by 2-m deep) filled with tap water. By experimentation, it was found that with a transmitted optical power of 5 mW the attenuation in this water was about 12 dB/m, giving a maximum theoretical range of 9.5 m, or slightly longer than the tank. It is clear from Fig. 4 that the expected signal-to-noise ratio is approximately 58 dB at 5 m, reducing by 12 dB/m to 10 dB at 9 m; hence, there was no detection problem during the communications over the ranges selected, which were at one-metre intervals from 1 m to 9 m.

The speech 'input' and 'output' signals are illustrated in Fig. 5. It is evident that the synthesised speech signal at the receiver correlates well with the original analogue speech signal, i.e. the microphone output. Because the system is not capable of measuring bit error rate, its performance was assessed subjectively by applying an intelligibility test, which is the usual practice in speech communications trials, sometimes deciding on a comparison category rating (CCR) [22]. Four listeners were asked to judge the quality of the decoded speech signals at the receiver. As expected, they commented that the subjective quality was poor. However, this was simply due to the LPC-10 speech compression algorithm implemented in this application. The results show that the system performs according to its design specifications; this is considered to be a significant achievement because it appears that there are no other papers anywhere in the literature on the underwater propagation of speech using a laser.

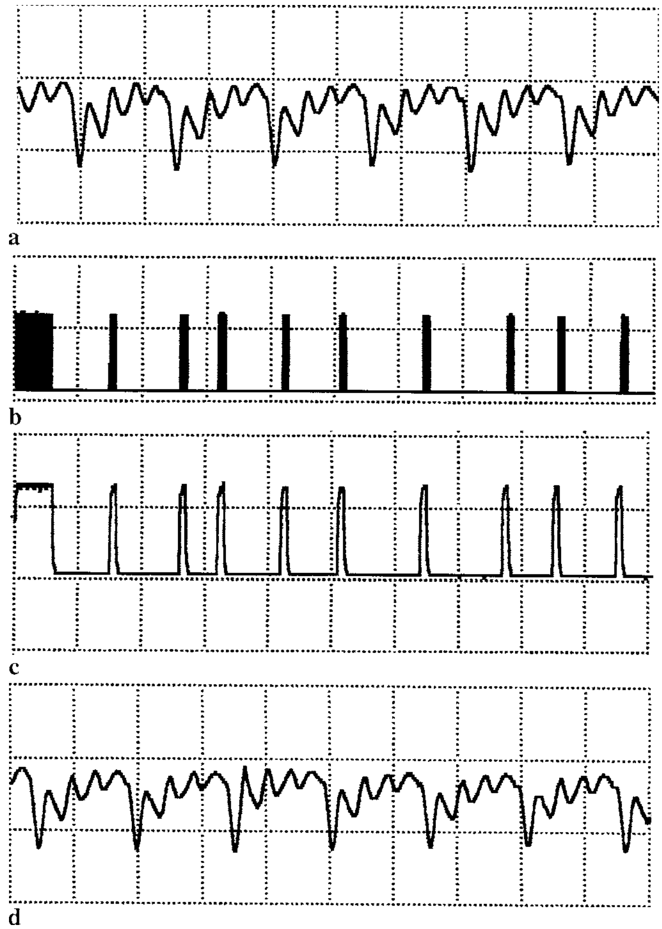

FIGURE 5 (**a**) Analogue speech signal for transmission, (**b**) optical signal transmitted in DPPM format, (**c**) optical signal received in DPPM format, (**d**) synthesised speech signal at the receiver

If the system is further developed for underwater speech communications an improved vocoder algorithm would be adopted. The most suitable is the enhanced version of MELP, called MELPe, which is known as military standard MIL-STD-3005 and NATO STANAG 4591; it can operate at three low rates: 2400, 1200, and 600 bps. The system in its present form is practicable only for a stationary set-up, due to the necessity for optical alignment, although with some modification it is envisaged that it could be used for short-range diver-to-diver speech communications. It could also be used for more conventional data communications, including the short message service (SMS). Further, it could be extended for use in a local area network with fixed transmitters and receivers. Clearly, for long-range communication, a blue–green laser with higher optical emission power must be used. Also, a trial carried out in the sea might reveal changes due to thermoclines or the presence of particulate matter, but such an exercise remains a task for the future.

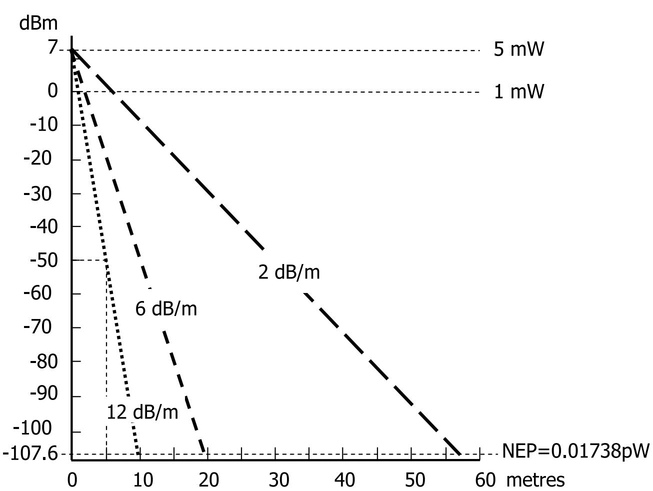

FIGURE 4 Power reduction as a function of range for a transmitted optical power of 5 mW (7 dBm)

REFERENCES

1 B. Woodward, H. Sari, IEEE J. Ocean. Eng. **21**, 181 (1996)
2 H. Sari, B. Woodward, in *Proc. OCEANS '97 Conf. II* (1997), p. 789

3 M.A. Chancey, M.Sc. thesis, North Carolina State University (2005)
4 R.G. Elion, A.H. Elion, *Electro-Optics Handbook* (Marcel Dekker, New York, 1979)
5 J.P. Von der Weid, J.A.P. Da Silva, A.C. Sant'Anna, in *Proc. OCEANS '93 Conf. III* (1993), p. 191
6 R.M. Cagliardi, S. Karp, *Optical Communications* (Wiley, New York, 1995)
7 M. Schroder, H. Barth, R. Reuter, Appl. Opt. **42**, 4244 (2003)
8 F. Schill, U.R. Zimmer, J. Trumpf, in *Proc. Australasian Conf. Robotics and Automation*, Canberra, 6–8 December 2004; www.transit-port.net/uwe.zimmer/publications/publications.html
9 J.W. Giles, I.N. Bankman, in *Proc. Military Communications Conf. (MILCOM 2005)*, vol. 3 (IEEE, Piscataway, NJ, 2005), p. 1700
10 A.D. Weidemann, G.R. Fournier, J.L. Fourand, P. Mathieu, S. McLean, Tech. Rep. 2, Naval Research Laboratory, Stennis Space Center, Mississippi, 25 April 2002
11 Z. Luo, Y. Lu, W. Chen, Proc. SPIE **2889**, 323 (1996)
12 S. Cowen, S. Briest, J. Dumbrowski, in *Proc. OCEANS '97 Conf. II* (1997), p. 1143
13 S.O. Duntley, J. Opt. Soc. Am. **53**, 214 (1963)
14 R.C. Smith, Appl. Opt. **20**, 177 (1981)
15 A. Majunder, J.C. Ricklin, *Free-Space Laser Communications: Principles* (Springer, London, 2005)
16 S.A. Spanias, Proc. IEEE **82**, 1541 (1994)
17 www.MELPe.com
18 S. Riter, in *Proc. IEEE SWIEECO Rec. 22nd Southwestern Conf. Exhib.* (1970), p. 453
19 G. Ling, R.M. Cagliardi, Proc. IEEE Trans. Commun. **34**, 1202 (1986)
20 J.G. Proakis, M. Salehi, *Communication Systems Engineering* (Prentice-Hall, Englewood Cliffs, NJ, 1994)
21 M. Stojanovic, IEEE J. Ocean. Eng. **21**, 125 (1996)
22 K.S. Lee, R. Cox, IEEE Trans. SAP **9**, 482 (2001)