ARTICLE IN PRESS

Neural Networks ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at SciVerse ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

An operationally adaptive system for rapid acoustic transmission loss prediction

Michael McCarron [a], Mahmood R. Azimi-Sadjadi [a,*], Michael Mungiole [b]

[a] Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80523-1373, USA
[b] US Army Research Laboratory, Attn: AMSRD-ARL-CI-ES, 2800 Powder Mill Road, Adelphi, MD 20783-1197, USA

ARTICLE INFO

Article history:
Received 27 January 2010
Received in revised form 26 October 2011
Accepted 9 November 2011

Keywords:
Acoustic propagation
Feedforward neural network
Fusion method

ABSTRACT

An operationally adaptive (OA) system for prediction of acoustic transmission loss (TL) in the atmosphere is developed in this paper. This system uses expert neural network predictors, each corresponding to a specific range of source elevation. The outputs of the expert predictors are combined using a weighting mechanism and a nonlinear fusion system. Using this prediction methodology the computational intractability of traditional acoustic propagation models is eliminated. The proposed system is tested on a synthetically generated acoustic data set for a wide range of geometric, source, environmental, and operational conditions. The results show a significant improvement in both accuracy and reliability over a benchmark prediction system.

1. Introduction

Sound waves traveling through the atmosphere undergo a complicated process of attenuation. The magnitude of attenuation or transmission loss (TL) is dependent on many factors including: atmospheric properties, such as wind, temperature gradients, and nonhomogeneity of atmospheric density; environmental properties, such as ground conditions; and operational properties such as source frequency, elevation, and range from the sensors. Due to the complexity of atmospheric sound propagation a tool which can quickly and accurately predict the TL for a given source and environment would prove to be valuable, finding many applications in military and private sectors, such as situational awareness, surveillance, and calibration of acoustic sensing systems.

A commonly used class of models for propagation of a single frequency source in a stationary, vertically stratified (temperature varies as a function of height) atmosphere are the parabolic equation (PE) models, which result from various approximations to the reduced wave equation (Gilbert & Di, 1994; West, Sack, & Walden, 1991). These models provide a two-dimensional (range and height) estimation of the TL in a given horizontal direction, as specified by an azimuth relative to the source. Propagation effects include spherical spreading, refraction due to vertical temperature and wind velocity variations, and reflections from a flat ground. Some models can accommodate various porous ground types (Gilbert & Di, 1994). Refraction due to vertical variations in the horizontal wind component can also be approximated via the effective sound speed, which adds the vector component of wind in the direction of propagation to the sound speed.

Most PE models use a spatial marching procedure (Mungiole & Wilson, 2006) to solve a parabolic equation which approximates the full wave equation. This method is much faster than numerically solving the reduced wave equation. However, it becomes relatively slow, particularly for longer horizontal ranges or higher source frequencies. To support near real-time applications a faster method utilizing neural networks to approximate the solution of a particular PE model was proposed in Mungiole and Wilson (2006). A method to successfully train the neural network using fewer input parameters than the actual PE model was also proposed.

The method in Mungiole and Wilson (2006) was further improved in Wichern, Azimi-Sadjadi, and Mungiole (2007) by utilizing a bank of expert neural network predictors instead of just a single neural network along with a fuzzy confidence measure-based decision system for selecting the appropriate expert predictors for a given environmental condition. By allowing the prediction system to adapt to given atmospheric conditions, performance for higher source frequencies and horizontal separations are improved.

To support more commonly occurring operational and environmental conditions, the range of parameters of the PE model considered in Mungiole and Wilson (2006), Wichern et al. (2007) needs to be expanded to include: (a) elevated sources e.g., low-flying sources; (b) larger horizontal separation, especially for loud or elevated sources which can travel long horizontal distances; (c) higher source frequencies to account for additional types of acoustic sources; (d) softer ground conditions, e.g., freshly fallen snow, to produce reliable TL predictions over all common ground hardness. Accurate TL prediction over these expanded ranges requires careful design of new neural network-based TL predictors.

* Corresponding author. Tel.: +1 970 491 7956; fax: +1 970 491 2249.
E-mail addresses: michael.v.mccarron@gmail.com (M. McCarron), azimi@engr.colostate.edu, firedem0n@hotmail.com (M.R. Azimi-Sadjadi).

0893-6080/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2011.11.004

Several important factors to consider when developing TL predictors for these expanded ranges are: ground interference, atmospheric regime, effects of high frequency sources, and large horizontal separations between source and receiver.

Elevated sources complicate ground interference effects, hence, significantly increasing the complexity of the PE model due to coherent interference from ground reflections. The inclusion of larger horizontal ranges and higher source frequencies worsens the effects of refraction and atmospheric absorption, further increasing the complexity of the PE model due to increasing the spatial nonlinearities in TL produced by these processes. Softer grounds introduce more nonlinearity into the ground reflected acoustic wave. These sources of additional nonlinearity complicate the PE model making approximation more difficult.

This paper develops a new operationally adaptive (OA) prediction system to improve the accuracy of TL prediction when the range of the input parameters of the PE model are expanded to include these new operational and environmental conditions. This system also employs neural networks to achieve the same computational savings of the previously developed methods (Mungiole & Wilson, 2006; Wichern et al., 2007). However, it utilizes new specialized expert predictors that take into account the elevation of the acoustic source in addition to new fusion methods. These fusion methods include a weighting system which weights the outputs of each expert based on an expert predictor's expected performance for a given source elevation, and a composite fusion method which incorporates the input parameters of the system into the fusion process.

The organization of this paper is as follows: Section 2 provides an overview of the PE model and a review of the expanded range of conditions. Section 3 reviews the overall structure of the proposed OA TL prediction system together with its constituent subsystems and presents the new weighting and fusion methods. Results and analysis of the proposed methods and a benchmark system are presented in Section 4. Section 5 offers concluding remarks on this work and possible directions for future work.

2. Review of the PE model for sound propagation

The crux of our model as described in West, Gilbert, and Sack (1992) is the two-dimensional narrow-angle PE, which approximates the full wave equation for atmospheric sound propagation given by

$$\frac{\partial^2 P(x,z)}{\partial z^2} - \frac{2k_0}{j}\frac{\partial P(x,z)}{\partial x} + (k_{eff}^2(z) - k_0^2)P(x,z) = 0 \tag{1}$$

where x and z are, respectively, the horizontal and vertical coordinates of the sound pressure field in the propagation direction, $P(x,z)$ is the sound pressure at that point, and $j = \sqrt{-1}$. The reference value of the wave number is denoted by $k_0 = 2\pi f/c_0$ where f is the frequency and c_0 is the reference sound speed. We define the effective wavenumber as $k_{eff}(z) = 2\pi f/c_{eff}(z) + j\nu(z)$ where ν is an attenuation coefficient modeling the absorption of sound energy as the acoustic wave propagates higher in the atmosphere (Salomons, 2001), and c_{eff} is the effective sound speed (Wilson, 2003) defined as the speed of sound plus the wind velocity component in the propagation direction.

The PE in (1) can be expressed in terms of dimensionless parameters to reduce the number of parameters required to solve the PE. The method in Mungiole and Wilson (2003) found a subset of 10 dimensionless parameters that are required to successfully train a neural network to approximate the solution of the PE. This reduced parameter set, given in Table 1, will be used in this paper since the accuracy of a function approximator tends to decrease as the dimension of the input vector increases. To express (1) in terms of these dimensionless parameters, the method in Mungiole and Wilson (2006) uses the ambient sound speed c_0, the ambient air density ρ_0, along with the frequency of the source f, to define the dimensionless parameters $\bar{P} = P\sqrt{(c_0/f)}/(\rho_0 c_0^2)$, $\bar{x} = xf/c_0$, $\bar{z} = zf/c_0$, $\bar{k}_{eff} = k_{eff}c_0/f$, and $\bar{k}_0 = k_0 c_0/f = 2\pi$, hence yielding (1) as

$$\frac{\partial^2 \bar{P}(\bar{x},\bar{z})}{\partial \bar{z}^2} - \frac{4\pi}{j}\frac{\partial \bar{P}(\bar{x},\bar{z})}{\partial \bar{x}} + (\bar{k}_{eff}^2(\bar{z}) - 4\pi^2)\bar{P}(\bar{x},\bar{z}) = 0 \tag{2}$$

where the over bar indicates dimensionless parameter.

The location of the acoustic receiver where the TL is computed is defined as $(\bar{x}_{sr}, \bar{z}_r)$, where $\bar{x}_{sr}$ is the dimensionless horizontal separation between source and receiver and $\bar{z}_r$ is the dimensionless receiver height. The sound pressure level in decibels (dB) at receiver location $(\bar{x}_{sr}, \bar{z}_r)$ can be determined from the following equation (Salomons, 2001)

$$\mathrm{SPL} = 20\log_{10}|\bar{p}(\bar{x}_{sr}, \bar{z}_r)| \tag{3}$$

where $\bar{p}(\bar{x}_{sr}, \bar{z}_r) = \bar{P}(\bar{x}_{sr}, \bar{z}_r)\exp(j2\pi\bar{x}_{sr})\sqrt{\bar{x}_{sr}}$ is the actual dimensionless sound pressure. The TL is calculated from the sound pressure level (SPL) by simply subtracting the initial sound pressure of the source, resulting in the relative change in sound pressure between source and receiver (Mungiole & Wilson, 2006). The output value of the PE model using (3), will be used as the target value (desired) during development of the TL prediction systems.

The PE model incorporates the effects of vertical refractions caused by wind velocity and temperature, the effects of the ground on reflected waves, and atmospheric absorption into the calculation of TL (West et al., 1992). The PE model is solved using a spatial marching procedure. The range and elevation plane is first discretized. The TL is then computed at each discrete point starting with the points closest to the source and moving outward. To accurately model the effects of frequency on propagation direction, the spacing between each discrete point is chosen to be inversely proportional to source frequency. This results in smaller spacing between discrete points for higher frequency sources. This enables accurate modeling of abrupt direction of propagation changes due to refraction which occur at high frequencies due to wind or temperature gradients. Although the spatial marching procedure produces accurate solutions, it dramatically increases computation times, especially for long ranges and at high frequencies.

The effects of the ground on reflected waves, and the refraction caused by wind velocity and temperature as a function of elevation and how they are incorporated into the PE are discussed in the next two subsections.

2.1. Ground profile

The model used for the ground surface in the PE model will have a large impact on the calculated TL values, because the ground surface causes wave reflections that interact with the direct acoustic wave as it propagates away from the source. In our formulation, $\bar{Z}_c$ denotes the dimensionless ground impedance using the model described in Wilson (1997) and is given by,

$$\bar{Z}_c = \bar{q}\left[\left(1 + \frac{\gamma - 1}{(1 - j\pi N_{pr}\bar{q}^2/\bar{\sigma})^{\frac{1}{2}}}\right)\left(1 - \frac{1}{(1 - j\pi\bar{q}^2/\bar{\sigma})^{\frac{1}{2}}}\right)\right]^{-\frac{1}{2}} \tag{4}$$

where N_{pr} and γ are the Prandtl Number and the ratio of specific heats for air, respectively. The dimensionless static flow resistivity $\bar{\sigma}$, is defined by $\bar{\sigma} = \sigma {s_p}^2/(f\rho_0\Omega)$, where s_p is the pore shape factor, and Ω is the porosity (Attenborough, 1985). The dimensionless tortuosity (Attenborough, 1985) to porosity ratio $\bar{q}$ is defined as $\bar{q} = q/\Omega$, where q is the tortuosity. The static flow resistivity σ models the hardness of the ground surface with high values

Please cite this article in press as: McCarron, M., et al. An operationally adaptive system for rapid acoustic transmission loss prediction. Neural Networks (2011), doi:10.1016/j.neunet.2011.11.004

Table 1
Range and expanded range of the parameters used to specify the PE model data.

Parameter (units)	Dimensionless definition	Min	Max	Old min	Old max
Frequency (Hz)	f	20	300	20	200
Horizontal separation (m)	$\bar{x}_{sr} = x_{sr}f/c_0$	100	2000	100	900
Source height (m)	$\bar{z}_s = z_s f/c_0$	0	50	0	5
Sound-speed scale (m/s)	$\bar{c}_* = c_*/c_0$	−1.18	0.2	−1.18	0.1
Static flow resistivity (Pa · s/m^2) × 10^6	$\bar{\sigma} = \sigma s_p^2/(f\rho_0\Omega)$	0.005	1.5	0.02	2
Receiver height (m)	$\bar{z}_r = z_r f/c_0$	0	5		
Azimuthal angle (rad)	β	0	π		
Friction velocity (m/s)	$\bar{u}_* = u_*/c_0$	0.05	0.5		
Tortuosity/porosity	$\bar{q} = q/\Omega$	1.54	10		
Roughness length (m)	$\bar{z}_0 = z_0 g/c_0^2$	0.001	0.1		

corresponding to hard ground (i.e., asphalt) and low values corresponding to soft ground (i.e., fresh snow). The porosity Ω models the number of pores in the ground surface, and the tortuosity q models the twisting of the pores, with $q = 1$ representing vertical alignment with the ground surface and $q > 1$ representing twisted pores. This model of the ground is incorporated into the solution of the PE model by imposing a boundary condition on the sound pressure at each range step, that is dependent on $\bar{Z}_c$ (West et al., 1992).

2.2. Atmospheric profile

Refraction, or the bending of sound waves in the near ground atmosphere due to vertical gradients in wind velocity and temperature is an important factor in acoustic modeling. Atmospheric refraction effects are incorporated via the effective sound speed, which is a function of temperature and wind velocity. Vertical variations in temperature and wind velocity can cause the sound to refract differently as a function of height and propagation direction relative to the wind field. To model refraction effects an atmospheric profile is defined, which relates the changes in temperature and wind velocity to elevation through the effective sound speed (Wilson, 2003)

$$c_{eff}(z) = c(z) + u(z)\cos\beta \tag{5}$$

where u is the wind velocity component in the horizontal direction, β is the azimuthal angle (angle between the sound propagation and wind directions), and c is the actual sound speed in air described by

$$c(z) = c_0\left(1 + \frac{T'(z)}{2T_0}\right). \tag{6}$$

It is assumed, for this formulation of the sound speed c, that the effect of humidity is small and can be ignored. Here, T' symbolizes the quantity of a small temperature variation about the ambient value T_0. The near ground temperature T and wind velocity u profiles are formed using the Monin–Obukhov similarity theory (Monin & Obukhov, 1954) derived in Mungiole and Wilson (2006), which describes the wind and temperature profiles (and thus the effective sound speed) in terms of the friction velocity u_* and the sound speed scale c_*. From (5) to (6), we see that the wind velocity and temperature profiles essentially describe the effective sound speed as a function of height.

The ratio of adiabatic coefficients (Wilson, 2003) for sound speed scale and friction velocity will be denoted in this paper, by $A = c_*/u_*$. Adiabatic coefficients are used due to assumptions on how temperature varies as a function of elevation by the Monin–Obukhov similarity theory. This ratio is extremely important in governing the direction of sound refraction in the atmosphere, and illustrates the relative contributions of wind velocity and temperature gradients in the atmosphere as they relate to sound refraction. Typically, values of $A > 0$ occur during calm nighttime conditions, and cause sound to refract downward creating a stable atmosphere. The stable atmospheric regime gives rise to hot spots and shadow regions where TL tends to vary rapidly, at various locations within the sound field. These manifest due to the sound refracting downwards toward the ground, repeatedly reflecting off it, causing coherent interference. Values of $A < 0$ tend to occur when the sun heats the ground during the daytime and causes sound to be refracted upward, creating an unstable atmosphere. As distance increases from the source, more of the source's acoustic energy is refracted up into the atmosphere. Thus, the unstable atmospheric regime tends to create a rapid drop in TL as the horizontal distance from the source increases.

2.3. Extended parameter ranges

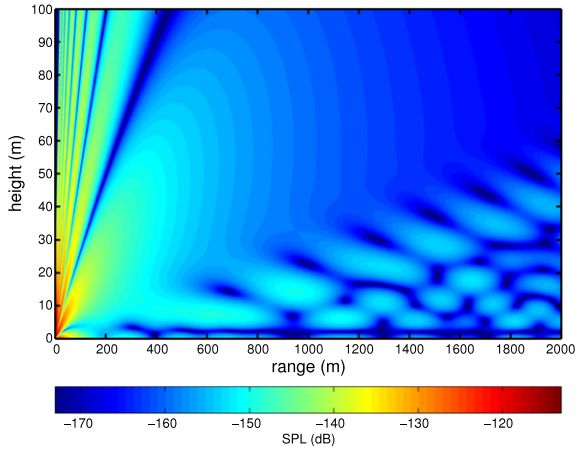
As mentioned before to account for additional atmospheric and ground conditions that are likely to be encountered in practice, the range of some parameters are expanded. Moreover, some sources that are of tactical importance (e.g., UAVs) in battlefield environments require using higher values of source frequency, horizontal range, and elevation. These key parameters with extended ranges, indicated by the italic font in Table 1, include frequency, horizontal separation (range), source elevation, static flow resistivity, and sound-speed scale. Table 1 also includes the dimensionless definition of each parameter and the ranges of each parameter used in this study.

Understanding the PE model and its behavior over these expanded parameter ranges is the key to designing an accurate TL prediction system. Analyzing the physics and geometry of the problem identifies physical phenomena which have the potential to influence the value of TL greatly, such as refraction, atmospheric absorption, and ground interference. It is useful to examine how the TL fields produced by the PE model are effected by these phenomena. This is done by identifying the key parameters and the specific ranges of these parameters which correspond to the conditions that create each of these phenomena, and examining the resulting TL fields. This process reveals valuable clues on how to approach the design of TL predictors.

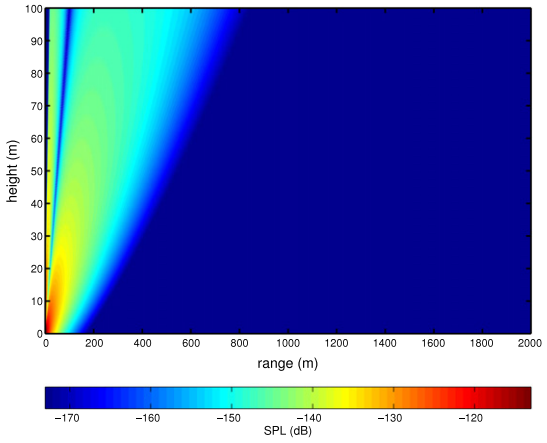
Fig. 1(a) shows a TL field for a high frequency ($f = 300$ Hz) ground source in a downward refracting (stable) atmosphere. The atmosphere causes the acoustic wave to refract downward, repeatedly bouncing off the ground as it propagates horizontally away from the source. This causes interference which manifests as hot spots and shadow regions which are seen as bright bubble-like regions and the dark regions that surround them, respectively. The structure of this field contrasts sharply with the field produced by the same source in an upward refracting (unstable) atmosphere, which is shown in Fig. 1(b). The acoustic wave is refracted up into the atmosphere as it travels horizontally away from the source, resulting in a very large shadow region.

The impact of A on TL is significantly reduced when the expanded source height parameter range is used. Fig. 1(c) shows

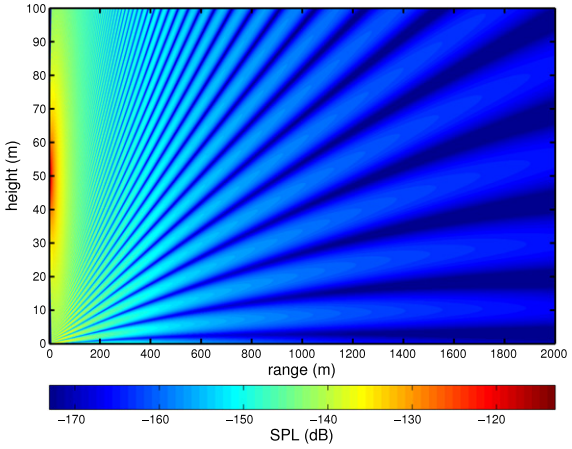
Please cite this article in press as: McCarron, M., et al. An operationally adaptive system for rapid acoustic transmission loss prediction. Neural Networks (2011), doi:10.1016/j.neunet.2011.11.004



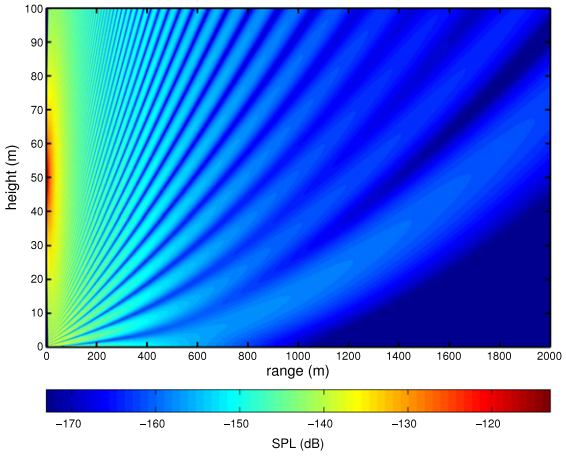
(a) Ground source ($z_s = 5$ m, $f = 300$ Hz) in a stable atmosphere ($A > 0$).



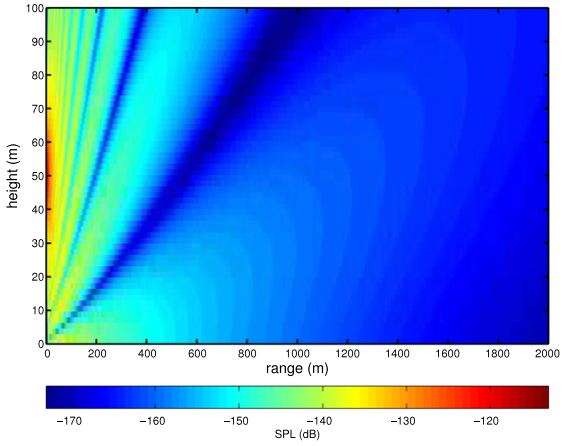
(b) Ground source ($z_s = 5$ m, $f = 300$ Hz) in an unstable atmosphere ($A < 0$).



(c) Elevated source ($z_s = 50$ m, $f = 300$ Hz) in a stable atmosphere ($A > 0$).



(d) Elevated source ($z_s = 50$ m, $f = 300$ Hz) in an unstable atmosphere ($A < 0$).



(e) Low frequency elevated source ($z_s = 50$ m, $f = 60$ Hz) in an unstable atmosphere ($A < 0$).

Fig. 1. Example PE model transmission loss fields.

the TL field produced by an elevated source ($z_s = 50$ m and $f = 300$ Hz) for the same stable atmosphere used in Fig. 1(a). The elevated source causes interference throughout the TL field between the acoustic wave propagating from the source and waves reflected off the ground. This manifests as alternating "spokes" of elevated and depressed TL across the field. Fig. 1(d) shows the TL field produced by the same elevated source for the unstable atmosphere used in Fig. 1(b). By comparing the effect the type of atmosphere has on the TL field for an elevated source vs. a ground source, we see that the atmosphere type does not effect the TL field for elevated sources as strongly as for ground sources. Since A determines the type of atmosphere in terms of refraction, we conclude that positive values of A do not introduce much nonlinearity when the source is elevated. However, a significant amount of nonlinearity is introduced when source height z_s takes

Please cite this article in press as: McCarron, M., et al. An operationally adaptive system for rapid acoustic transmission loss prediction. Neural Networks (2011), doi:10.1016/j.neunet.2011.11.004

values corresponding to elevated sources ($z_s > 5$ m) due to the resulting ground interference.

Many propagation processes which strongly impact TL are dependent on frequency, f, and horizontal separation, x_{sr}. The nonlinearity introduced by these processes tends to dramatically increase as the values of f and x_{sr} increase. These propagation processes include atmospheric absorption, refraction, and ground interference. The rate at which atmospheric absorption occurs is directly proportional to f, thus larger changes to TL occur at high frequencies. The distance traveled also affects the total amount of attenuation due to atmospheric absorption, therefore, large values of x_{sr} also introduce more nonlinearity. The rate of refraction is also proportional to f, with higher frequencies resulting in faster refraction and more nonlinear behavior of TL. Larger values of x_{sr} allow more refraction to occur, increasing the nonlinear effect on TL due to refraction at large horizontal separations. Finally, the nonlinearity introduced by ground interference due to elevated sources is proportional to f, with higher frequencies producing more quickly varying interference patterns. This is demonstrated by Fig. 1(e) and (d), which show a low frequency ($f = 60$ Hz) and a high frequency ($f = 300$ Hz) elevated source in an unstable atmosphere, respectively. Fig. 1(e) shows the slowly varying interference pattern due to this low frequency source, while Fig. 1(d) shows the quickly varying interference pattern due to the higher source frequency. It is apparent that in the latter case, increasing the maximum value of x_{sr} and f, increases the nonlinearity of the PE model, making it more difficult to approximate.

Static flow resistivity, σ, is a parameter of the ground impedance model which describes its resistance to airflow (proportional to how much energy is reflected). Large values correspond to hard surfaces. Reducing the highest value of σ does not affect the relationship between σ and TL significantly because we are still considering hard grounds. Small values of σ correspond to soft surfaces. Reducing the minimum value allows soft grounds such as freshly fallen snow to be considered. These soft grounds tend to add more prominent nonlinearity to the behavior of TL than hard grounds. This nonlinearity increases as we encounter even softer grounds. However, only slightly softer grounds are considered, hence limiting the overall impact to the nonlinearity of the PE model attributed to expanding the range of σ.

Sound-speed scale, c_*, helps to determine the type of refraction present in the atmosphere. By expanding its range to include more positive values we are including slightly more stable atmospheres (downward refracting). It is known that stable atmospheres cause large amounts of nonlinearity for ground sources. However, few combinations of c_* from the expanded range region and u_* exist which satisfy the Monin–Obukov similarity theory. This severely limits the nonlinearity added to the PE model due to the expanded range of c_*.

3. Operationally adaptive TL prediction system

To develop a TL prediction system it is important to consider the high variability in the type of nonlinearity due to the elevation of the source which is seen between the TL fields in Fig. 1(a) and (c), and Fig. 1(b) and (d). This drastic change will negatively impact the successful training of any neural network-based TL predictor that must capture the types of nonlinearity produced by these situations simultaneously.

A major disadvantage of using a single neural network-based TL predictor when wide swings in the type of nonlinearity exist, in the data, is the possibility of a performance bias. This performance bias tends to favor the type of nonlinearity which is better represented in the training data set. Since a large majority of the range of source elevations cause ground interference, a bias may be seen toward the nonlinearity caused by ground interference versus that of no ground interference when all ranges of source elevations are equally represented in the training data set. This bias may degrade the performance of the developed predictor in the critical ground source regions, and must be avoided. Biasing the network's training to emphasize both modes equally may help to alleviate this possible performance degradation but may also negatively impact the performance for elevated source cases, as learning effort will be distributed evenly among the cases. As explained later, to overcome these problems, the proposed OA TL prediction system is designed to adapt to source height (and therefore ground interference), allowing the predictors in the system to learn these situations separately.

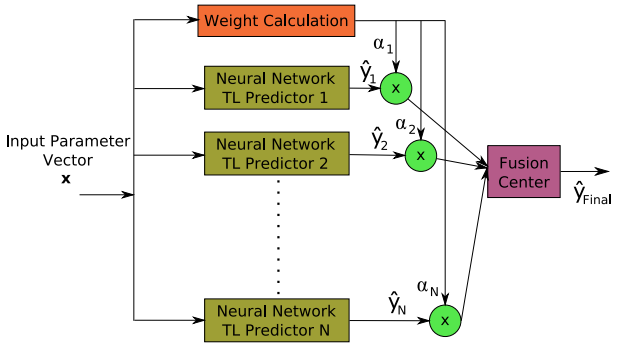


Fig. 2. Block diagram of the OA system with multiple predictors.

To adapt to source height, the OA TL prediction system utilizes a divide-and-conquer strategy to imprint knowledge of the PE model in specific input parameter space regions, namely regions of differing source height, into the structure of the prediction system. If the PE model is viewed as a nonlinear mapping from the input parameter space to the TL space, this can be accomplished by decomposing this mapping into multiple segments each with more "well-behaved" nonlinearity. Each component mapping can then be learned separately by different TL predictors. Choosing the right divisions of the mapping's domain is obviously the key to the success of this approach. The details of the divisions used by the OA system are described in Section 3.2.

With this divide-and-conquer strategy, multiple predictors respond to each input vector, hence providing multiple estimates of the TL value. To produce a single final TL prediction, a fusion system that takes the preliminary predictions and produces a final prediction is needed (see Section 3.3). The general structure of such a system can be seen in Fig. 2. For this application the parameters of the propagation environment along with all the other necessary parameters for TL prediction are supplied to the system through the input parameter vector. This vector, which is formed of the 10 dimensionless PE parameters described in Table 1 and denoted by $\mathbf{x} = [f, \bar{x}_{sr}, \bar{z}_s, \bar{z}_r, \beta, \bar{u}_*, \bar{c}_*, \bar{\sigma}, \bar{q}, \bar{z}_0]^T$, is fed to each OA TL predictor. The preliminary TL estimate produced by the ith OA TL predictor for a given input parameter vector, $\mathbf{x}$, is denoted by $\hat{y}_i$, while α_i represents the multiplicative weight that is applied to the predictor output. The modular nature of this system structure allows new partitions and predictors to be easily incorporated into an already trained system. Finally, the weighted preliminary TL predictions are fed into the fusion system which produces a final TL estimate $\hat{y}_{final}$.

In the following subsections, the structure of the expert neural network predictors and the PE model data sets used to train them are reviewed. This is followed by a description of the proposed weighting and fusion methods.

3.1. Neural network expert predictors

Two-layer back-propagation neural networks (BPNNs) with one hidden layer and one output layer are chosen for the OA

Please cite this article in press as: McCarron, M., et al. An operationally adaptive system for rapid acoustic transmission loss prediction. Neural Networks (2011), doi:10.1016/j.neunet.2011.11.004

neural network TL predictors. The output layer of each network consists of a single neuron with a linear activation function which is commonly used for regression problems due to the large range of output values. The standard sigmoidal activation function was chosen for the neurons in the hidden layer. The size of the hidden layer which produced the most accurate TL predictors was determined experimentally to be 45 neurons. The Levenberg–Marquardt algorithm (Hagan & Menhaj, 1994) is used to train the predictor networks, utilizing mean-square-error (MSE) as the error measure. An early stopping method is used to prevent overfitting.

3.2. PE model data

Information about the PE model needs to be extracted in a suitable format to train the OA system's expert neural network TL predictors. To do this, a set of input/output samples of the PE model is generated. The output of the PE model from (3) defined as d is used as the target (or desired) value during the neural network training. The output of the PE model for a given input vector is a 2-D cylindrically symmetric field of relative TL values as seen in Fig. 1(a)–(e). A moving average (MA) filter is applied to the TL values along the path from the source to the receiver to smooth out sharp peaks and nulls in the TL as a function of horizontal separation which are due to near perfect constructive and destructive interference between direct and ground reflected waves. In practice, in a turbulent atmosphere, sharp extrema in TL as a function of horizontal separation, do not occur. Turbulence causes small fluctuations in the phase of the sound wave preventing perfect constructive and destructive interference between direct and ground reflected waves. Sources that are not pure tones tend to mask the effects of interference at single frequencies. Additionally, horizontal variations in wind, temperature, and ground impedance, which are always present in any real environment, tend to perturb the actual sound field from that of the ideal model prediction. Thus, although shallow interference nulls may occur, they will have a stochastic behavior which cannot be captured by the model. Applying the MA filter, therefore, provides a more reasonable approximation of an actual sound wave by reducing the interference-induced fluctuations in the mapping function for regions of sharp extremes such as those caused by ground interference and the unstable atmospheric regime.

After the MA filter is applied, forty input/output samples are taken uniformly along the line between the source and receiver. To generate enough samples for neural network development, 13,284 TL fields were generated as random vectors with elements that are uniformly distributed over each input parameter's respective range. This, results in a PE model data set of 531,360 input/output samples.

To divide the data based on ground interference cases, subsets of the PE model data must be found which represent different degrees of ground interference. Due to the geometry of the problem it can be devised that ground interference introduces high spatial variability in TL along the line between source and receiver only when source height, z_s, is above 10 m. This leads to the selection of the following data partitioning intervals along z_s:

1. $z_s \in$ [0 m, 3 m]: minimal spatial variability caused by ground interference.
2. $z_s \in$ [2.5 m, 12 m]: minimal to weak spatial variability caused by ground interference.
3. $z_s \in$ [10 m, 40 m]: weak to moderate spatial variability caused by ground interference.
4. $z_s \in$ [25 m, 50 m]: moderate to severe spatial variability caused by ground interference.

These four partitions cover the entire range of source elevations (from ground to low-flying sources) by separating z_s into regions of differing spatial variability introduced by ground interference. The overlap between the partitions allows for each predictor (one for each partition) to gain some knowledge of how the PE model behaves when transitioning from one ground interference region to the next. This helps to reduce highly erroneous outputs (outliers) which may result from insufficient information in the transition regions.

Each PE model subset is then used to train one of the OA TL neural network predictors. To train these neural networks each subset needs to be divided into a training, validation, and testing set. Due to the additional samples along the horizontal separation for every combination of the other inputs caused by taking multiple samples from each TL field, care must be taken when creating these subsets to avoid improper fitting of the PE model data. The training set is used during the training of the neural network to capture the PE model mapping. The testing set is used to evaluate the performance of the network on novel data, giving an estimation of the performance of the trained system. The validation set is used to improve the generalization (performance on new data) of the neural network.

The validation set improves generalization performance in two ways. First, during neural network development different random initializations of the network weights are chosen resulting in several trained neural networks. The network with the best performance on the validation set is then used for subsequent testing. The validation set is also used during the training process for early stopping. If the neural network training algorithm is run for too many epochs, the parameters of the network become fine-tuned to the training data. To prevent this overfitting problem, the performance on the validation set is evaluated after every training epoch. Training is stopped when performance fails to improve on the validation set after five epochs.

If the training and validation sets are not constructed carefully the network will lose its ability to generalize on new data. To avoid this, the training and validation sets are formed such that each group of forty input vectors, differing only in horizontal separation, are in the same set. Thus, for every sample in the training set, all of the samples in the validation set have a different combination of the other nine inputs, excluding horizontal separation.

The fusion of the preliminary TL predictions produced by the OA TL predictors into a final TL prediction is presented next.

3.3. Weighting and final prediction fusion

The fusion system takes the output of the expert neural network predictors and produces $\hat{y}_{final}$, the final TL prediction. The most straightforward way to create a fusion system is to use a single neural network and use the output of each predictor as a separate input. This fusion method will be called *standard* fusion and will be used to benchmark the proposed fusion methods.

3.3.1. Weighted fusion

The goal of fusion is to map the inputs to the fusion system to the desired TL value. However, the neural network predictors can produce very similar values for many different input parameter vectors. If these input vectors to the fusion system have different desired TL values then the fusion system has to map the same inputs to different desired values, which cannot be done using the standard fusion network. Thus, to overcome the problems of standard fusion, the outputs of each predictor must be modified so that similar outputs with different desired TL values can be separated by the fusion neural network. In this way, the fusion neural network can map each case independently, resulting in a more accurate fusion.

Please cite this article in press as: McCarron, M., et al. An operationally adaptive system for rapid acoustic transmission loss prediction. Neural Networks (2011), doi:10.1016/j.neunet.2011.11.004

One way to accomplish this goal is to apply a 0–1 weighting to the output of each predictor for a given input vector. The predictor which produces the most accurate TL prediction is weighted one, while the other outputs are weighted zero. Clearly, if we assume that the desired TL value, d_j, for the jth input vector $\mathbf{x}_j$ is available, the index of the predictor that produces the most accurate TL prediction is then chosen using,

$$B = \arg\min_i |d_j - \hat{y}_{ij}|, \quad \forall i \in [1, N] \tag{7}$$

where $\hat{y}_{ij}$ is the output of the ith predictor for $\mathbf{x}_j$ and $N = 4$ for the partitioning mentioned in Section 3.2. In this case, the 0–1 weighting scheme is applied to the outputs of each predictor to create the weighted input to the fusion neural network, for input vector $\mathbf{x}_j$, as

$$\hat{\mathbf{y}}_j = [\alpha_1 \hat{y}_{1j}, \alpha_2 \hat{y}_{2j}, \alpha_3 \hat{y}_{3j}, \alpha_4 \hat{y}_{4j}]^T \tag{8}$$

with

$$\alpha_B = 1, \qquad \alpha_{i \neq B} = 0, \quad i \in [1, N] \tag{9}$$

where $\hat{\mathbf{y}}_j$ is the weighted input to the fusion network for the jth input vector and $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \alpha_3, \alpha_4]^T$ is the weight vector.

Now, during operation, when a new input parameter vector $\mathbf{x}_j$ is applied to the system, the "most accurate predictor" is selected using the *maximum a posteriori* (MAP) estimator which maximizes the *a posteriori* conditional PDF, $p(B|z_s)$ or equivalently the joint PDF $p(B, z_s)$. That is, we have

$$\hat{B}_{MAP} = \arg\max_B p(B|z_s). \tag{10}$$

This estimate is then used to generate the coefficients in the weight vector, $\boldsymbol{\alpha}$, using $\alpha_{\hat{B}_{MAP}} = 1, \alpha_{i \neq \hat{B}_{MAP}} = 0, i \in [1, N]$. An estimation of the conditional distribution $p(B|z_s)$ is constructed (see e.g., Scott (1979)) using pairs of $\{\hat{y}_{ij}, d_j\}_{j=1}^M$, from the training data set, where M is the number of training samples.

Note, the reason for using conditional distribution $p(B|z_s)$ is due to the operational partitions discussed in Section 3.2. Additional input parameters can also be used along with z_s such as f or x_{sr}, but there is a tradeoff between the size of the measurement vector and the MAP estimate's accuracy due to the finite number of data samples available to estimate $p(B|z_s)$. Finding the ideal combination of input parameters and preliminary predictions is a subject for future research.

This weighting process successfully improves the usefulness of the inputs to the fusion network in separating some cases with the same prediction value but different desired values, resulting in a more accurate $\hat{y}_{final}$. Despite improving accuracy, better performance can still be achieved if the desired output is also a function of the inputs to the fusion neural network.

3.3.2. Composite fusion

The performance can further be improved by designing a fusion neural network and giving the network visibility to the input space, since the desired value is a function of the input vector. This guarantees the existence of the mapping from the fusion neural network's inputs to the desired value. The fusion neural network can now separate expert predictor's outputs and adjust error independently, hence significantly improving performance. To do this a composite input vector is created as $\hat{\mathbf{y}}_j = [\hat{y}_{1j}, \hat{y}_{2j}, \hat{y}_{3j}, \hat{y}_{4j}, \mathbf{x}_j^T]^T$, for input parameter vector $\mathbf{x}_j$. Alpha weighting is not used here as this was applied in the weighted fusion system to help simplify the mapping from the expert predictor's outputs to the desired value in an effort to remove ambiguity when similar expert outputs had different desired values. Since all ambiguity is removed when using the composite input vector, this is no longer needed.

Table 2
Performance comparison of the OA and single neural network TL prediction systems on the testing data set.

System	Error (dB)	% Outliers (>3 dB)
Single NN	5.11	46.2
OA-Std	4.46	36.8
OA-Weight	4.21	38.0
OA-Comp	3.23	25.6

Table 3
Performance comparison of the standard and weighted fusion methods.

System	$\hat{B}$	Error (dB)	% Outliers (>3 dB)
OA-Simple	$\hat{B}_{MAP}$	4.29	39.0
OA-Weighted	$\hat{B}_{MAP}$	4.21	38.0
OA-Simple	B	2.97	20.2
OA-Weighted	B	2.94	20.3

4. Results and discussion

This section presents the results of the OA system using the fusion methods presented in this paper, along with those of a single BPNN (Mungiole & Wilson, 2006), which is used as a benchmark. The performance of each TL prediction system was evaluated on the testing set using two measures namely the RMS error and the percent outliers for errors with magnitude greater than 3 dB.

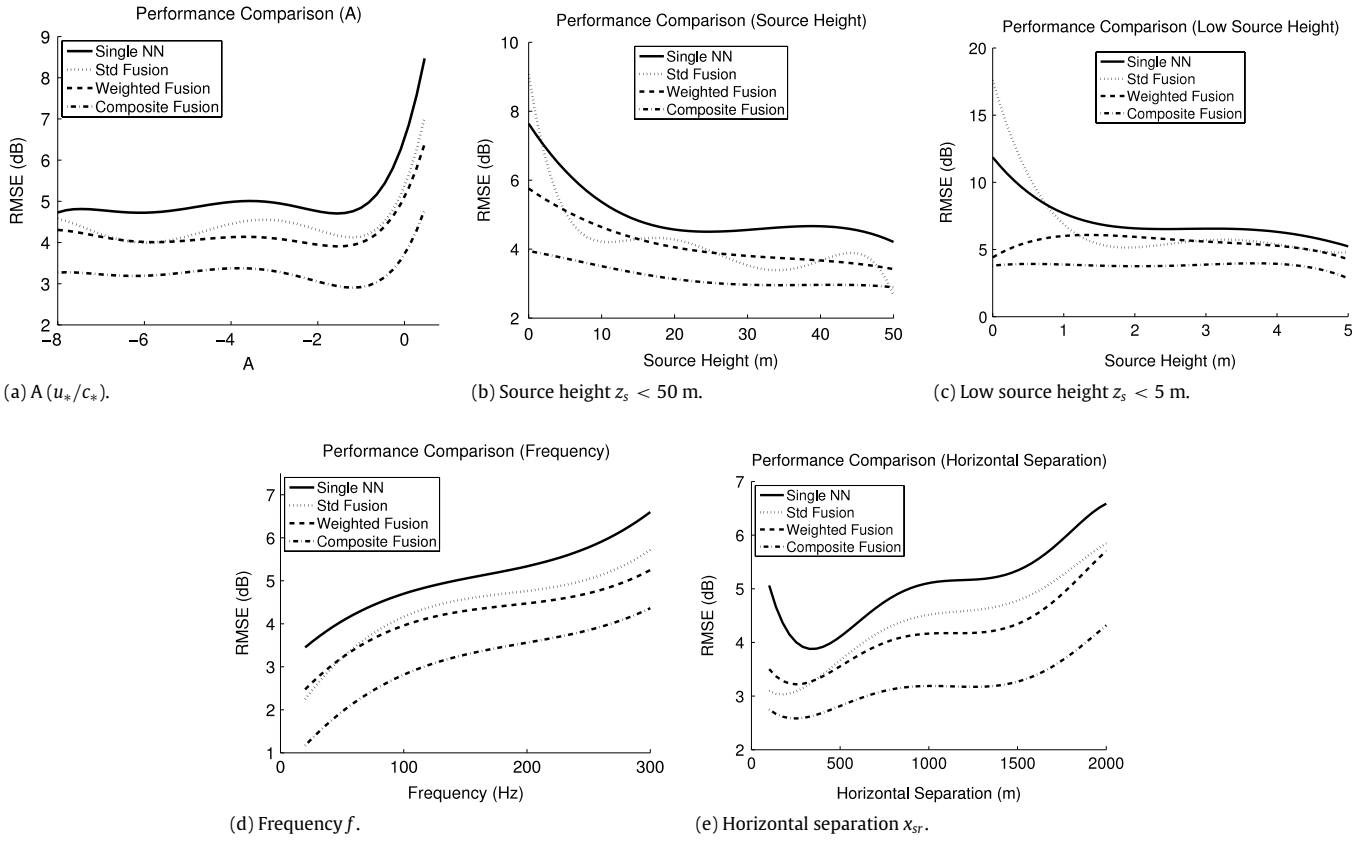
Table 2 shows the values of these performance criteria for the OA system and the single BPNN TL predictor. The benchmark single BPNN TL predictor resulted in an RMS error of 5.1 dB and 46% outliers. The OA TL prediction system with standard fusion reduces the RMS error to 4.46 dB and the percent outliers to 36.8%. By using the 0–1 weighting, the RMS error was reduced to 4.21 dB while the % outliers slightly increased to 38% compared to the benchmark system. This may be attributed to the 0–1 weighting of the preliminary predictions, which results in the fusion neural network losing the ability to account for the outliers. The composite fusion method reduced RMS error to 3.23 dB and outliers to 25.6% compared to the benchmark. These results are indeed acceptable from a real operational point of view, as the PE model is an approximation of actual acoustic propagation. Further attempts to improve the approximation of the PE model may not necessarily improve performance on actual data.

To evaluate the effectiveness of the weighted fusion network it is illuminating to compare the performance of the weighted fusion system and a simple fusion system which simply selects the preliminary prediction $\hat{y}_{\hat{B}}$ as the final TL prediction $\hat{y}_{final}$. In this manner, the actual performance gain due to the fusion neural network can be examined. Table 3 presents the performance results for the simple fusion system compared to the weighted fusion system when (a) the MAP estimation $\hat{B}_{MAP}$ is used (rows 2–3 of Table 3); or (b) $\hat{B}$ is chosen as the actual value B, i.e. knowing which expert predictor is the most accurate (rows 4–5 of Table 3).

Regardless of the choice for $\hat{B}$, the difference in performance between the simple fusion system and the weighted fusion system is slight. This shows that the weighted fusion neural network does little to improve the output of the best predictor. However, using the composite fusion instead of the weighted fusion, the RMS error is reduced from 4.2 to 3.2 dB (see Table 2) and percent outliers from 38% to 25.6%. This is a significant improvement as these values are much closer to those achieved when the actual value of B is assumed (not possible in real situations) to be known (see last two rows of Table 3).

Since the performance values in Table 2 are averaged over the entire testing set, it is useful to examine the trend of RMS error as a function of input parameters to get a better idea of how the systems

Please cite this article in press as: McCarron, M., et al. An operationally adaptive system for rapid acoustic transmission loss prediction. Neural Networks (2011), doi:10.1016/j.neunet.2011.11.004



(a) A (u_*/c_*). (b) Source height $z_s < 50$ m. (c) Low source height $z_s < 5$ m.

(d) Frequency f. (e) Horizontal separation x_{sr}.

Fig. 3. RMS error versus various input parameters.

perform for various ranges of PE model input parameters. Fig. 3(a) shows the RMS error as a function of A, the ratio of sound speed scale to friction velocity, for the single BPNN TL prediction system and the OA TL prediction system with standard fusion, weighted fusion, and composite fusion network. Note that values of $A < 0$ correspond to upward refracting, unstable atmospheres. Over this range, while the performance of each system remains relatively constant, the lowest RMS error is provided by OA system with composite fusion. Values of $A > 0$ correspond to stable, downward refracting atmospheres. The additional nonlinearity introduced from downward refracting waves bouncing off the ground and coherently interfering with themselves degrades performance of all the systems in this region. As can be seen, however, in both atmospheric regimes the OA systems performed better than the single BPNN benchmark. Amongst the OA systems, the composite fusion neural network system performs significantly better than the others.

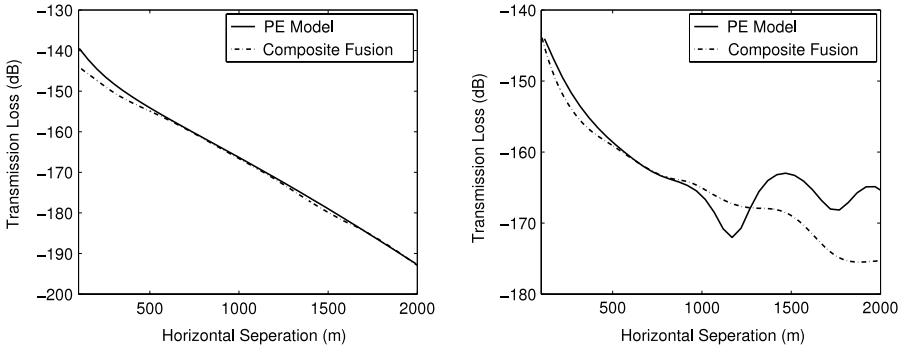
Fig. 3(b) and (c) show RMS error as a function of source height, z_s, for each system over source heights $z_s < 50$ m and source height corresponding to ground sources ($z_s < 5$ m), respectively. The single BPNN system and the OA system with standard fusion both performed very poorly in the ground source region with an RMS error of around 7.5 dB as seen in Fig. 3(c). The single BPNN performs better than the OA system with standard fusion for the lowest source heights which suggests the expert neural network predictors produce similar outputs with different desired values for many inputs from this region, resulting in poor fusion performance. The weighted fusion helped to reduce error at low source heights to around 5.75 dB by improving the separability of predictor outputs generated by input vectors from the ground source region and other regions. The OA composite fusion system further reduced the error to 4 dB in this region which is significant due to the importance of ground sources in many military applications. The error generally decreases for all systems as source height increases. This is attributed, in part, to the reduced impact A has on TL for elevated sources, which results in an easier to capture mapping.

Fig. 3(d) shows RMS error as a function of the frequency of the source. The RMS error of each TL prediction system has a nearly linear relationship of increasing error with increasing frequency. This is caused by propagation effects whose impact on TL scale with frequency. These include atmospheric absorption, interference due to ground reflections, and refraction. The OA system with composite fusion network, however, outperformed the other TL prediction systems for all frequencies, and provided significant improvement in the critical region of 50–200 Hz which includes the range of frequencies typically emitted by ground vehicles (engine firing and tire noise).

Fig. 3(e) shows RMS error as a function of horizontal separation x_{sr}. The single BPNN system has relatively large errors of around 5dB for the lowest range of horizontal separations (i.e. $x_r < 200$ m). This is partly due to the inability of the MA filter to smooth out the trend of TL at the smallest horizontal separations as significantly as for other horizontal separation values since the MA filter window only overlaps a smaller number of TL values. The error in this region is reduced significantly to less than 3 dB by the OA system with the composite fusion network. The error tends to increase with horizontal separation due to the increasing impact on TL of atmospheric absorption and refraction as horizontal separation increases.

Fig. 4(a) and (b) show the output of the PE model and the OA system with composite fusion for two different combinations of input parameters. These plots are provided to demonstrate the type of errors that may be seen by the OA system with composite fusion. Fig. 4(a) is generated by choosing input parameters that fall into the lower RMS error regions as seen in Fig. 3(a)–(e),

Please cite this article in press as: McCarron, M., et al. An operationally adaptive system for rapid acoustic transmission loss prediction. Neural Networks (2011), doi:10.1016/j.neunet.2011.11.004



(a) Typical output. (b) High RMS error case.

Fig. 4. Example cases of PE model And OA system output.

while Fig. 4(b) is generated with input parameters from high RMS error regions. The low RMS error input vector is $[f = 20, x_{sr} \leq 2000, z_s = 40, c_* = -0.8, \sigma = 1.3 \times 10^6, z_r = 2.5, \beta = 0, u_* = 0.25, q = 2, z_0 = 0.01]$ and the high RMS error input vector is $[f = 300, x_{sr} \leq 2000, z_s = 3, c_* = 0.05, \sigma = 0.1 \times 10^6, z_r = 2.5, \beta = \pi, u_* = 0.1, q = 4, z_0 = 0.1]$. Fig. 4(a) shows that the PE model and OA system with composite fusion agree closely, with little error after 500 m of horizontal separation. Fig. 4(b) shows low error until around 850 m horizontal separation at which point the PE model's output starts to vary quickly as horizontal separation increases. This rapid variation in the PE model's output is due to downward refracting sound waves bouncing off the ground causing destructive and constructive interference. The OA system with composite fusion was unable to completely capture this nonlinear behavior resulting in errors of up to 10dB. It is difficult to completely characterize the type of error that the OA system produces, however, these cases should give some hint as to the types of errors that may be encountered.

5. Conclusion

This paper introduced an OA acoustic TL prediction system which utilizes multiple expert neural network predictors specializing on different ranges of source height to incorporate knowledge of the PE model into the prediction process. Two new fusion methods are presented which combine the outputs of the OA system's expert neural network predictors. The first is a 0–1 weighting method which applies a weight of one to the best predictor for a given input while applying a zero weight to the other predictors. The best predictor for a given input is determined by a MAP estimator. The second proposed fusion method uses a composite input vector that combines the OA system's input vector with the outputs of the bank of predictors to generate the final prediction using a fusion neural network. The proposed fusion systems were extensively benchmarked against the OA system with standard fusion and a single BPNN predictor on a synthetically generated PE model data set. These results showed that the proposed systems improve the accuracy and reliability of PE model approximation over standard fusion in terms of both RMS error and percent of outliers, while retaining near real-time capabilities important in many situational awareness and tactical decision making applications. The OA system with composite fusion network resulted in the best overall performance with an RMS error of 3.2 dB and 25.6% outliers, a significant improvement over a benchmark BPNN with RMS of 5.1 dB and 46.2% outliers. The process of developing these systems also exemplified how similar fusion methods can be applied to other mixtures of experts to improve decision fusion performance in other applications.

Acknowledgments

This research was supported by the DoD Center for Geosciences/Atmospheric Research at Colorado State University under Cooperative Agreement W911NF-06-2-0015 with the Army Research Laboratory.

References

Attenborough, K. (1985). Acoustical impedance models for outdoor ground surfaces. *Journal of Sound and Vibration*, *91*, 521–544.
Gilbert, K. E., & Di, X. (1994). A fast Green's function method for one-way sound propagation in the atmosphere. *Journal Acoustic Society of America*, *94*, 2343–2352.
Hagan, M., & Menhaj, M. (1994). Training feedforward networks with the Marquardt algorithm. *Neural Networks, IEEE Transactions on*, *5*(6), 989–993.
Monin, A. S., & Obukhov, A. M. (1954). Basic laws of turbulent mixing in the atmosphere near the ground. *Akad. Nauk SSSR Geofiz. Inst. Tr.*, *24*, 163–187.
Mungiole, M., & Wilson, D.K. 2003. A nondimensional parameterization for sound propagation in the atmosphere. US Army Research Laboratory Technical Report. (ARL-TR-2950).
Mungiole, M., & Wilson, D. K. (2006). Prediction of outdoor sound transmission loss with an artificial neural network. *Applied Acoustics*, *67*, 324–345.
Salomons, E. (2001). *Computational atmospheric acoustics*. Norwell, MA: Kluwer Academic Publishers.
Scott, D. W. (1979). On optimal and data-based histograms. *Biometrika*, *66*(3), 605–610.
West, M., Gilbert, K., & Sack, R. A. (1992). A tutorial on the parabolic equation (PE) model used for long range sound propagation in the atmosphere. *Applied Acoustics*, *37*, 31–49.
West, M., Sack, R. A., & Walden, F. (1991). The fast field program (FFP). A second tutorial: application to long range sound propagation in the atmosphere. *Applied Acoustics*, *33*, 199–228.
Wichern, G., Azimi-Sadjadi, M., & Mungiole, M. (2007). Environmentally adaptive acoustic transmission loss prediction in turbulent and nonturbulent atmospheres. *Neural Networks*, *20*(4), 484–497.
Wilson, D. K. (1997). Simple relaxational models for the acoustical properties of porous media. *Applied Acoustics*, *50*, 171–188.
Wilson, D. K. (2003). The sound-speed gradient and refraction in the near-ground atmosphere. *Journal of the Acoustical Society of America*, *113*, 750–757.

Please cite this article in press as: McCarron, M., et al. An operationally adaptive system for rapid acoustic transmission loss prediction. Neural Networks (2011), doi:10.1016/j.neunet.2011.11.004